



VNiVERSIDAD
D SALAMANCA

**FROM ETHICS TO AGENCY:
PARTICIPATORY DESIGN OF A
TEACHER TRAINING COURSE FOR
AI IN EDUCATION**

Doctoral Programme in
Education in the Knowledge Society

DOCTORAL DISSERTATION
ANA ISABEL MOUTA COSTA

DIRECTORS
EVA MARÍA TORRECILLA SÁNCHEZ, PHD
ANA MARÍA PINTO LLORENTE, PHD



VNiVERSiDAD D SALAMANCA

Doctoral Dissertation

FROM ETHICS TO AGENCY: PARTICIPATORY DESIGN OF A TEACHER TRAINING
COURSE FOR AI IN EDUCATION

PHD STUDENT

Ana Mouta

Signed by the directors:

Professor Eva María Torrecilla-Sánchez

Professor Ana María Pinto-Llorente

Doctoral Thesis

Authorisation from the directors for the presentation of the Doctoral Thesis

Dr. Eva María Torrecilla-Sánchez, from the area of Research Methods and Educational Diagnosis, and Dr. Ana María Pinto-Llorente, from the area of Didactics and School Organisation, both affiliated with the Department of Didactics, Organisation, and Research Methods at the University of Salamanca.

Certify:

That Ana Isabel Mouta Costa has completed the Doctoral Thesis, titled “**From Ethics to Agency: Participatory Design of a Teacher Training Course for AI in Education**”, under their supervision, and that it meets the necessary quality and originality criteria to qualify for the degree of Doctor from the University of Salamanca. They also certify that the doctoral candidate is opting for the mention of International Doctorate.

And for the record, it is signed by the directors of the Doctoral Thesis for all pertinent purposes in Salamanca, on July, 2025.

ACKNOWLEDGEMENTS

To my supervisors, I extend my warmest thanks. To Professor Eva María Torrecilla Sánchez, for her unwavering support in advancing the empirical studies, for her on-the-ground guidance, and for challenging me to channel intellectual enthusiasm and passion into the discipline of academic inquiry and the process of becoming a researcher.

To Professor Ana María Pinto Llorente, for being a constant source of stimulus, challenge, and support in just the right measure, making us want to become better each day in our work. I am grateful for her rigour, intellectual brilliance, and empathetic understanding of both research and life itself.

I also thank Professor Francisco José García-Peñalvo for his consistent cooperation and thoughtful diligence that helped move our work forward. His support was always timely and light.

Likewise, I received prompt and efficient help from the Doctoral Administration Office, especially with all bureaucratic procedures. These small gestures can make a world of difference at the right moment, and they are never forgotten.

To Emanuele Bardone, for the privilege of spending time alongside a Socratic, deeply humanistic thinker, where looking toward the future of research is a sociohistorical commitment to all that has come before, while always anticipating the radical encounter with new possibilities for thinking about Education, society, and our shared destiny.

To Paulo Vicente, for facilitating my entire academic journey with the opportunities he opened up to me, inviting me into the team at the Social Observatory of AI and Digital Data.

To my fellow travellers. The dear Lorena Camusso and Miriam Pina, with whom I shared this doctoral journey, step by step, through small gestures, words of encouragement, and joyful notes.

To Pedro Falcão and Daniel Silver, for knowing just when to be my right hand in thinking through the research. Pedro will always be the truest of friends.

To all the professors and researchers, especially Norma Lemus, who agreed to take part in this study, with the wholehearted generosity of those who not only contribute to this space of collective agency, but who also know they are supporting a researcher along a path that was not always easy, nor always swift.

To my Anas: Paulino, Paleta, Gomes, Couto, and Roseira. Each of you, at different times, took care of me during this beautiful season of my life, when I had the privilege of studying what moves me deeply. Special thanks to Couto for being a light for both my soul and my grounding. And special thanks to Roseira, with whom the seeds of curiosity for these themes began to sprout during long beach walks back in 2018.

To Wayne Holmes, whom I can only admire for embodying the very spirit of what it means to be a researcher: an open, bold, sharp intellectual peer, capable of turning actions into the highest expression of what is woven in the realm of ideas and in the closed circuits of academic inquiry. A true agitator of our collective agency.

Finally, to those who hold me in unconditional love and listening: my mother and my father. To the most playful, understanding, and loving of boys, my son Álvaro, and to the greatest companion I could have had on this journey, Eduardo Pais, who recognised in my drive to question, and in every decision and path I chose to take, the very reason to offer me his enduring support from behind the scenes.

Abstract

Artificial Intelligence is increasingly being integrated into educational settings, yet its ethical implications and impact on pedagogical agency remain underexplored. This thesis investigates the ethical challenges and agency-related concerns in education through an educational design research process, with the aim of developing a teacher training course for K-12 educators, designed through their own voices.

The study begins with a systematic literature review (2011–2022), conducted using PRISMA guidelines, which maps the current state of research on AI in education. This phase identifies substantial gaps in ethical frameworks, teacher-specific guidance, and the preservation of educational agency. Building on this foundation, the research adopts a participatory futures methodology, using the Delphi Method to co-construct eight future scenarios. These scenarios explore the socio-technical imaginaries shaping AI's pedagogical implications, including issues of equity, assessment, student voice, and professional autonomy.

Subsequent research phases engage teacher educators through iterative focus groups, exploring how AI alters agency dynamics, subjective, intersubjective, and collective, within educational contexts. Findings reveal a pressing need to move beyond dominant technosolutionist narratives and instead support teachers in reclaiming their roles as ethical and relational agents. These insights inform the co-design of a professional development course, which integrates dialogic, experiential, and reflective learning practices. The course is hosted on a custom-designed Canva platform and structured around a three-layered framework of educational agency, offering educators conceptual and practical tools to critically engage with AI.

By foregrounding the symbolic, relational, and ethical dimensions of education, this thesis argues that responsible AI integration must not only be technically sound but also aligned with the core purposes of education: subjectification, qualification, and socialisation. It proposes that sustaining teacher agency requires special attention to the preservation and care of the educational lexicon, one that sustains complexity, openness, ethical discernment, as well as desire and memory in the face of algorithmic pressures. For it is through desire that alternative imaginaries of socio-technical systems and comprehensive educational ecosystems are made possible.

This dissertation contributes four main outcomes: (1) a comprehensive ethical mapping of AI in education, (2) a participatory ethical dilemma toolkit, (3) a conceptual framework of agency in AI-mediated education, and (4) a context-responsive, agency-centred professional

development course for K–12 educators. Together, these outcomes constitute a theoretically grounded and empirically informed contribution to ongoing scholarly and professional efforts aimed at cultivating educational environments in which decisions regarding the use of AI, and the conditions under which it is integrated, are co-constructed through dialogic, participatory processes that uphold educational purpose, human agency, and the democratic ethos of schooling. It counters the depoliticising and deprofessionalising tendencies of technocratic models by supporting teachers in critically engaging with AI, resisting unreflective automation, and challenging algorithmic normativisation.

Keywords: Education, Artificial Intelligence, Artificial Intelligence in Education, Collective Agency, Sense of Agency, Teacher training, Educational Design Research, Futures Studies

Resumen

La Inteligencia Artificial está siendo cada vez más integrada en entornos educativos, pero sus implicaciones éticas y su impacto sobre la agencia pedagógica siguen siendo insuficientemente explorados. Esta tesis investiga los desafíos éticos y las preocupaciones relacionadas con la agencia en la educación mediante un proceso de investigación basada en el *Educational Design Research*, con el objetivo de desarrollar un curso de formación para profesores de Educación Primaria y Educación Secundaria diseñado a partir de sus propias opiniones.

El estudio comienza con una revisión sistemática de literatura entre 2011 y 2022, realizada según las directrices PRISMA, que mapea el estado actual de la investigación sobre la IA en la educación. Esta fase identifica vacíos significativos en los marcos éticos, en la orientación específica para docentes y en la preservación de la agencia educativa. Sobre esta base, la investigación adopta una metodología de futuros, utilizando el método Delphi para construir de manera colectiva ocho escenarios futuros. Estos escenarios exploran los imaginarios sociotécnicos que configuran las implicaciones pedagógicas de la IA, incluyendo cuestiones de equidad, evaluación, voz estudiantil y autonomía profesional.

Las fases posteriores de la investigación involucran a formadores de docentes mediante grupos focales iterativos con el fin de explorar cómo la IA transforma las dinámicas de la agencia subjetiva, intersubjetiva y colectiva en contextos educativos. Los hallazgos revelan una necesidad urgente de superar las narrativas tecnosolucionistas dominantes y, en su lugar, apoyar a los docentes en la recuperación de su rol como agentes éticos y relacionales. Estas ideas informan el diseño conjunto de un curso de desarrollo profesional que integra prácticas de aprendizaje dialógicas, experienciales y reflexivas. El curso se aloja en una plataforma personalizada basada en Canva y se estructura en torno a un marco tridimensional de agencia educativa, ofreciendo a los docentes herramientas conceptuales y prácticas para involucrarse críticamente con la IA.

Al poner en primer plano las dimensiones simbólicas, relacionales y éticas de la educación, esta tesis sostiene que la integración responsable de la IA no solo debe ser técnicamente sólida, sino también coherente con los fines fundamentales de la educación: subjetivación, cualificación y socialización. Se propone que la sostenibilidad de la agencia docente requiere una atención especial a la preservación y al cuidado del léxico educativo, un léxico que sostenga la complejidad, la apertura, el discernimiento ético, así como el deseo y la memoria frente a las presiones algorítmicas. Es a través de este deseo que pueden emerger nuevos imaginarios de los sistemas sociotécnicos y de ecosistemas educativos más integrales.

Esta tesis ofrece cuatro aportes principales. En primer lugar, un mapeo comprensivo sobre la IA en la educación y la ética. En segundo lugar, un *toolkit* de dilemas éticos aplicados a la IA en la educación. En tercer lugar, un marco conceptual de agencia en contextos educativos mediados por IA. Y, en cuarto lugar, un curso de formación docente contextualizado y centrado en la agencia. En conjunto, estos resultados constituyen una contribución teórica y empíricamente fundamentada a los esfuerzos académicos y profesionales que buscan cultivar entornos educativos en los cuales las decisiones sobre el uso de la IA, y las condiciones bajo las cuales se implementa, se construyan colectivamente mediante procesos dialógicos y participativos que sostengan el propósito educativo, la agencia humana y el horizonte democrático de la escuela. Esta investigación contrarresta las tendencias despolitizadoras y desprofesionalizadoras de los modelos tecnocráticos al apoyar al profesorado en su compromiso crítico con la Inteligencia Artificial, en su resistencia a la automatización irreflexiva y en el cuestionamiento de la normativización algorítmica.

Palabras clave: Educación, Inteligencia Artificial, Inteligencia Artificial en Educación, Agencia Colectiva, Sentido de Agencia, Formación Docente, Investigación de Diseño Educativo, *Futures Studies*

CONTENTS

CHAPTER ONE Contextualisation and Structure of the Thesis.....	18
1.1. Introduction.....	19
1.1.1. AI as a Socio-technical System: Imaginaries, Intelligence, Instituting, Archiving	19
1.1.2. Artificial Intelligence in Education.....	20
1.1.3. Challenges of Using AI in Learning Environments.....	23
1.1.4. Coverage of AIED Ethics in Research Policy and Training.....	27
1.1.5. Framing the Shift from Ethics to Agency	29
1.1.6. Agency as a Basis for Ethics: Accountability and Meaningful Reasoning.....	31
1.2. Research Design	33
1.2.1. Research Questions	34
1.2.2. Methodological Framework: Educational Design Research.....	34
1.2.3. Research Phases and Thesis Structure	36
CHAPTER TWO State of the Art.....	38
2.1. Introduction.....	39
2.2. Methodology.....	40
2.2.1. Planning	40
2.2.2. Conducting and Analysing.....	42
2.3. Discussion.....	62
2.3.1. SLR Implications for Rethinking Education Ethics: Using a Theoretical Model.....	62
2.3.2. Uncovering Research Blind Spots on AIED Ethics.....	65
2.3.3. Recommendations for Educators on AIED Ethics.....	66
2.3.4. Potential Limitations of This Research.....	67
2.4. Conclusion	68
CHAPTER THREE Empirical Research I: Futures Studies	70
3.1. Introduction.....	71
3.1.1. Arguments for a Participatory Approach to the Ethics of AIED Use.....	71
3.1.2. Futures Studies.....	73
3.1.3. Historical Overview of Futures Studies: Theory and Method	73
3.1.4. Historical Overview of Futures Studies Application to Education.....	75
3.1.5. The Delphi Method: Theory, Evolution, and Educational Use.....	76
3.2. Methodology.....	78
3.2.1. Expert Group Constitution	78

3.2.2. Rounds Implementation	82
3.3. Results Per Round.....	86
3.3.1. Iteration 1	86
3.3.2. Iteration 2	88
3.3.3. Iteration 3	89
3.4. Discussion.....	90
3.5. Conclusion	96
CHAPTER FOUR Empirical Research II: Participatory Collectives	98
4.1. Introduction.....	99
4.1.1. Ethical and Meaningful Pedagogical Standards through CPD	99
4.1.2. Teacher Agency	100
4.2. Methodology.....	102
4.2.1. Planning	102
4.2.2. Implementation	106
4.2.3. Data Analysis Procedures	106
4.3. Results.....	109
4.3.1. RQ1: Structuring Dimensions for the Training Programmes	109
4.3.2. RQ2: Aims and Objectives Prioritised by Educators.....	111
4.3.3. RQ3: Design Considerations for the Training Programmes	113
4.3.4. RQ4: Implementation Strategies	116
4.4. Discussion.....	117
4.5. Conclusion	124
CHAPTER FIVE Empirical Research III: The Shifting Nature of Agency in Education with AI.....	127
5.1. Introduction.....	128
5.1.1. The Concept of Human Agency.....	130
5.1.2. Human Agency and Education	132
5.1.3. Education and AI	133
5.2. Methodology.....	135
5.2.1. Research Question.....	135
5.2.2. Participants, Implementation, and Data Analysis	136
5.3. Results.....	137
5.4. Discussion.....	140
5.4.1. Agency Levels.....	140

5.4.2. Cross-Level Patterns and Interdependencies	145
5.5. Conclusion	149
CHAPTER SIX Course Design and Evaluation	151
6.1. Introduction: Situating the Course in the Research Trajectory	152
6.2. Methodological Positioning for a Research-Based Pedagogical Intervention	153
6.2.1. Assessment in an Age of Automated Scaling: A Pedagogical Defense of Relational Learning	155
6.2.2. Reframing Beyond SCORM and Personalisation: Rethinking the Role of Technology in Course Delivery	157
6.3. Pedagogical Framework and Course Design	159
6.3.1. Aligning Course Design with Key Research Findings	159
6.3.2. Design and Implementation Aspects.....	160
6.4. Course Piloting through Walkthrough and Micro-evaluation	170
6.4.1. Walkthrough Stage.....	172
6.4.2. Micro-evaluation Stage	184
6.5. Conclusion	195
CHAPTER SEVEN Conclusions, Implications, and Future Directions	197
7.1. Conclusion	198
7.2. Limitations of the Current Research.....	207
7.3. Future Directions for Research and Development.....	208
CHAPTER EIGHT Dissemination of Ph.D. Research	211
8.1 Publications in Peer-Reviewed Journals.....	212
8.2. Publications in Conference Proceedings	212
8.3. Conference Presentations.....	213
8.4. Journal Reviewer	214
8.5. Advisory Board Member	214
8.6. Research Network.....	215
8.7. Publications Resulting from Research Network.....	215
8.8. Research Stays	215
CHAPTER NINE References	217
APPENDICES	252
Appendix A Published Manuscript in International Journal of Artificial Intelligence in Education	253
Appendix B Published Manuscript in Education and Information Technologies	294

Appendix C Published Manuscript in Digital Society321
Appendix D Published Manuscript in Education and Information Technologies345
Appendix E Course Guidebook and Session Plans (Sessions 1–8)391
Appendix F Extended Abstract in Spanish413

FIGURES

Figure 1. Overview of the EDR Research Trajectory.....	37
Figure 2. PRISMA 2020 Flow Diagram for the SLR Process: Papers Identification.....	44
Figure 3. Number of Articles on AIED Ethics Between 2011-2022	46
Figure 4. Countries with Publication in AIED Ethics.....	47
Figure 5. Publications on AIED Ethics Main Subject Areas	48
Figure 6. Impact and Uncertainty Grid for AIED Futures Scenarios	90
Figure 7. Process of Code Creation and Testing.....	107
Figure 8. Number of References Per Item	109
Figure 9. Items Clustered by Coding Similarity	111
Figure 10. Thematic Flowchart of Subjective, Intersubjective, and Collective Levels.....	148
Figure 11. Modular Course Layout on Canva: Opening Slide View.....	162
Figure 12. Sample Session Plan Demonstrating Applied Pedagogical Framework	163
Figure 13. Example of the 7th Session Entry on the Course support Platform	169
Figure 14. Iterations of the Systematic Design Cycles	171
Figure 15. Distribution of Responses to the First Question: Most Striking Theme	186
Figure 16. Distribution of Responses to the Second Question: Theme for Further Exploration	187
Figure 17. Distribution of Responses to the Third Question: Most Relevant Activity	188

TABLES

Table 1. SLR Planning Phase.....	42
Table 2. Quality Assessment Questions.....	45
Table 3. Reasons for the Inclusion of Additional Papers.....	45
Table 4. AIED Technologies, Applications, Targets and the Ethical Challenges.	51
Table 5. Percentage of Papers on AIED Cultural and Inclusiveness Ethical Challenges.....	54
Table 6. Cultural and Inclusive Features of AIED.	54
Table 7. Percentage of Papers on AIED Ethical Challenges on Emotions.....	55
Table 8. Capacity Building on AIED Ethics.....	59
Table 9. Ethical AI and AIED Guidelines and Frameworks	62
Table 10. Experts’ Sociodemographic and Professional Profiles	80
Table 11. Experts’ Professional Profiles.....	81
Table 12. Delphi Third Round: Tables with Hypothetical Scenarios Sent to Experts	85
Table 13. Delphi First Round Results: Frequencies Per Category	86
Table 14. Teachers’ Sociodemographic Profile.....	104
Table 15. NVivo Framework Matrix	108
Table 16. Optimising Teacher Training: Planning Dimensions	110
Table 17. Framework for Agency in Education: Teachers’ Perspectives on AI and Human Agency	138

Since the raison d'être of machines lies in output, or rather, in maximum output, they, or in fact each one of them, require surrounding worlds that ensure this maximum. And what they require, they conquer.

— Günther Anders, *Wir Eichmannsöhne: Offener Brief an Klaus Eichmann*

CHAPTER ONE

Contextualisation and Structure of the Thesis

1.1. Introduction

1.1.1. AI as a Socio-technical System: Imaginaries, Intelligence, Instituting, Archiving

Artificial Intelligence (AI) is not merely a technological achievement; it is a socio-technical system deeply rooted in human culture, imagination, and evolving semantics over time. At their core, AI systems are products of "imaginaries", highlighting how envisioning and imagining are fundamental to society. These imaginaries thrive independently of the actual feasibility of technologies. They reflect multiple facets of human endeavours to construct possible futures and address complex problems across several domains. Far from being a recent phenomenon, such imaginaries have been a persistent feature of human history, shaping everything from ancient myths of automata (where the concepts of agency and overcoming were embedded) to contemporary aspirations for superintelligent machines. This continuity presents AI not only as a technological artifact but also as a reflection of deeply rooted societal desires, fears, values, and governance structures. As a significant force in contemporary politics, these imaginaries not only have the power to influence the design of technology but also to guide public spending, determine who benefits from technological advancements, and decide who may be excluded from these benefits (Jasanoff et al., 2007).

Despite the theoretical and conceptual contradictions surrounding the concept of "intelligence", it has served as a leitmotif in the creation and development of systems that are claimed to embody our aspirations and demonstrate our mastery. Pasquinelli (2023) argues that AI cannot be conceived as an imitation of biological or human intelligence, nor does it replicate individual organisms. Therefore, AI is not a "showing" or "presentation" of human intelligence. It is not also a product of computational advancement: AI crystallises collective intelligence, aligning with Malabou's (2021) conceptualisation of intelligence as plastic, continuously shaped by socio-political, technological, and cultural shifts rather than existing as a fixed entity.

Historically, the concept of intelligence has transitioned through multiple metamorphoses: from its early measurability in Binet's psychological tests, through the epigenetic paradigm that prioritises adaptability, and now into an era where automation and AI act as "intelligent" agents. The epistemological evolution of intelligence, as framed by Piaget's (1952) epigenetic approach, emphasises that intelligence develops through interaction with the environment. When we consider socio-technical systems, AI not only extends human cognition but also reshapes the very conditions in which intelligence is generated, assessed, and institutionalised.

As Bourdieu's (1999) concept of field suggests, intelligence – whether conceived as biological, artificial, or hybrid – is always embedded in power relations. AI technologies, while often perceived as neutral or objective, embody sociopolitical biases, determining which forms of knowledge are privileged and which are marginalised. Malabou (2021), drawing on this perspective, urges us to recognise AI not as an autonomous agent but as a site of contestation, where human and machine intelligences co-construct each other within specific institutional and ideological frameworks.

In the context of AI, rather than asking what intelligence is, the focus needs to be on how it happens, who enacts it, and through which technologies it is archived and distributed. In a Deleuzo-Guattarian (1995) sense, intelligence is part of an ongoing process of subjectivation and individuation; it can take multiple, fluid forms depending on who is instituting it, how it is being archived, and what power relations are embedded in its formation. This highlights intelligence as pragmatic and situated; relational and collective (unfolding within networks of interaction, whether among humans, machines, institutions, or environments); and technologically mediated, meaning it cannot be conceived independently of the infrastructures that enable it. In this way, AI must be understood as an assemblage, a socio-technical ecology in which human, institutional, and algorithmic agencies are intertwined.

The question is no longer whether AI possesses intelligence but rather how intelligence, as a historically contingent and plastic phenomenon, is being reshaped by the very systems we have created. In a sense, we might argue that these technologies do indeed reflect the state of our intelligence, as they depict its displacement, conceptually, theoretically, and most importantly organically. Furthermore, for Pasquinelli (2023), AI represents a crystallisation of collective intelligence, emerging from the collaboration of our bodies and minds within society. This raises questions about how we organise collectively and produce our collective agency.

1.1.2. Artificial Intelligence in Education

Artificial Intelligence can broadly be understood as a technique for recognising patterns extracted from material data through multi-dimensional analysis (Pasquinelli, 2023). Since its emergence in the late 1950s, AI has been seen both as a promise and a threat, expected to ease everyday tasks while also potentially exerting decision-making power over humans. Today, AI functions as an autonomous, adaptive, and interactive software system, capable of making complex decisions through data-based perception, interpretation, and reasoning (Dignum, 2021).

Artificial Intelligence in Education (AIED) research began in the 1970s and developed gradually before experiencing rapid growth in the past decade, particularly in Western countries. This has led to a surge in research interest. A comprehensive literature review covering 1993–2020 identified a range of AIED applications, including school management, student learning, teacher support, and lifelong education (Zhang & Aslan, 2021). According to Luckin et al. (2016), AI in education offers a wide range of functions, from AI teaching assistants to personalised tutoring for learners in all subjects. On the administrative side, AI systems can align institutional objectives with policy, support teachers in their day-to-day tasks, respond to student interests, and foster lifelong learning (Miao et al., 2021).

AI-driven lifelong learning assistants will be available to offer guidance, personalised recommendations, and learning progress tracking, while flexible learning environments will enable students to study at their own pace and preferred location (Luckin et al., 2016). School management systems powered by AI can streamline operations such as admissions, timetabling, attendance monitoring, homework tracking, resource allocation, and inspections (Miao et al., 2021), while also enhancing communication between stakeholders.

These technologies are marketed as having numerous benefits: identifying and nurturing talent, reducing teacher workload, and addressing learner diversity (Mousavinasab et al., 2018; Miao et al., 2021). AI is also positioned as a tool to predict underachievement, support professional transitions, and increase access to high-quality, affordable education, particularly for disadvantaged groups. In terms of teaching and learning, AI applications include personalised support and intelligent tutoring systems (ITSs) that adjust to learner needs (Khosravi et al., 2022; Luckin et al., 2016).

These systems are capable of processing large datasets to analyse individual learning preferences, strengths, and difficulties. ITSs employ pedagogical agents (PAs) to scaffold self-regulated learning, providing prompts and feedback to help learners manage cognitive, emotional, metacognitive, and motivational processes (Dever et al., 2024). They can identify areas of struggle and offer targeted explanations, hints, or feedback to improve understanding (Wang et al., 2023). Such systems may take the form of virtual dialogues, interactive simulations, or step-by-step problem-solving tools, and they refine their approaches over time by learning from user interactions (Khosravi et al., 2022).

AI can also predict future learning outcomes based on past performance, helping educators identify both struggling and high-performing students. This enables timely interventions and personalised support, with recommended resources and activities tailored to individual learning needs and pace (Miao et al., 2021). AIED can also provide insights into how students acquire

different skills, using embedded assessments, adaptable content, and real-time monitoring of learning progress and influencing factors (U.S. Department of Education, Office of Educational Technology, 2023).

In addition to ITSs, AI also facilitates low-pressure, conversational environments for language learners through chatbots. For example, AI-powered chatbots such as Andy English Chatbot allow learners to practice speaking and receive feedback without fear of judgement, supporting both fluency and self-confidence (Fathi et al., 2024). Language learning platforms like Duolingo illustrate the broader potential of AIED for lifelong language acquisition. These tools have demonstrated their effectiveness through high validity and practicality in written and oral assessments (Sudina & Plonsky, 2024).

AI technologies are widely advertised, promoted, and researched as tools that enhance professional development and teaching practice in multiple ways. Generative AI (GenAI), for instance, is promoted as enabling experiential learning through simulations with AI-generated mentors and assessors, creating scalable and interactive training environments (Mollick et al., 2024). These simulations are said to support pre-service teacher training by offering dynamic, adaptive scenarios that are otherwise difficult to replicate in real-world contexts (Dai, 2024; Xu et al., 2024). In addition, platforms are designed to incorporate teachers' assessment criteria, while AI agents assist with subject-specific scenarios and pedagogical strategies tailored to the proficiency levels of future educators (Liu, 2023). AI also creates new opportunities for teachers through the provision of personalised resources and planning tools that adapt to instructional needs (Copur-Gencturk et al., 2024).

AI-driven adaptive assessments are marketed as providing responsive evaluation by tailoring to difficulty levels of a student's ability, thereby avoiding frustration or boredom. These tools can also assess engagement and emotional states by analysing language patterns and interactions, giving teachers the insight needed to adjust teaching approaches and improve learners' emotional readiness (Luckin et al., 2016). Such capabilities move learning beyond the traditional "stop-and-test" model and offer new ways to reduce achievement gaps, increase teaching effectiveness, and address shortages and turnover in the teaching workforce (Luckin et al., 2016). AI also supports timely professional development and provides tools for parents to support their children's education, with learning materials updated in line with current trends, research, and standards.

The COVID-19 pandemic intensified debate over AI's role in online assessments and remote learning (García-Peñalvo et al., 2021). In Romania, for example, AI-based educational platforms saw increased adoption between 2019 and 2020 (even in less developed regions) as

a response to pandemic-induced disruptions (Pantelimon et al., 2021). These debates continue today, questioning whether AI can deliver equitable learning during crises. In the future, questions may arise regarding AIED's suitability in contexts of forced migration, as seen in UNHCR initiatives to provide digital access, relevant online content, and education to displaced populations (UNHCR, 2016).

Despite widespread enthusiasm for AI-driven teaching and learning, there is still limited evidence supporting its transformative impact on education (Holmes et al., 2022). Williamson (2024) extends this critique, asserting that the promotion of AI in education assumes that these systems will perform as intended, with any issues encountered during their development or implementation being addressed through technical refinements or ethical regulations. Back in 2019, Richards and Dignum affirmed the need for a systematic examination of the values and ethics that justify the use of these technologies in education, considering the pedagogical approaches they can foster and their societal impact.

1.1.3. Challenges of Using AI in Learning Environments

Using Biesta's (2020) framework of subjectification, socialisation, and qualification, requires questioning the educational purposes underlying AIED. At the pedagogical level, some AI tools may conflict with socio-constructivist learning principles, which can result in a fragmented or superficial integration into classroom practices. While they are marketed as personalised, many AIED tools create individualised paths that converge toward a standardised goal, potentially overlooking the diversity of learner experiences and contexts. This approach may reflect broader political and economic agendas that prioritise mass education at the expense of developing individual agency and capacity.

On a subjective level, AI can significantly influence or interfere with human experiences such as decision-making, moral reasoning, deliberate action, personal development, and ownership of projects. In educational settings, for instance, Learning Analytics platforms offer descriptive and predictive insights into student engagement and learning trajectories. The Council of Europe (2022) warns that these tools may inadvertently undermine the agency of both teachers and students. In fact, research has shown that interacting with automated systems can reduce users' sense of control (Berberian et al., 2012; Obhi & Hall, 2011). This sense of agency is key for students to develop ethical reasoning skills, which in turn enable them to critically engage with complex societal challenges and foster accountability and moral responsibility (Bandura et al., 1999).

Questions persist about the cognitive impact of AI-based educational systems, as there is currently insufficient evidence to determine their long-term effects on learning and cognition (Council of Europe, 2022). Natural Language Processing (NLP) tools provide automated feedback on grammar, syntax, and coherence, but they often prioritise quantity over quality and have been found to be inconsistent in error detection (Liu & Kunnan, 2016; Miao et al., 2021). Similarly, tools like Photomath offer students access to mathematical content but may present interpretation challenges, especially for users with limited prior knowledge, which could negatively affect students' self-perception in mathematics (Gaona et al., 2022). For this reason, transparency and explainability in AI systems are crucial, along with creating opportunities for discussion about assessment and decision-making.

ITSs have faced criticism for insufficient inclusivity concerning learners with special needs or diverse backgrounds (Alrakhawi et al., 2023). These systems risk amplifying educational disparities and introducing cultural biases (Miao et al., 2021). GenAI may address some of these gaps by supporting students with hearing or visual impairments and, also, by enabling learning in a student's own language. Some researchers also propose that GenAI could use conversations to flag psychological, social-emotional, or learning issues, but evidence of safety or effectiveness is limited, and results would need expert review (UNESCO, 2023a).

In addition to these concerns, there are significant privacy issues related to the collection of sensitive data, which could potentially reveal individuals' health, ethnic identities, or sexual orientations (Tundrea, 2020). Another ethical issue concerns the current discourse on AI ethics, which often confines privacy to the boundaries of legality, privileging compliance with regulatory frameworks (e.g., GDPR or FERPA) over a deeper ethical engagement with the lived experience of privacy itself. This legalistic framing often obscures dimensions of privacy such as anonymity, intimacy, and secrecy (O'Shea, 2019), qualities that are key for cultivating ethical subjectivity and personal growth in educational spaces. AI systems, especially those built on dashboards, tracking tools, and automated profiling, rarely make room for rupture, surprise, or silence, the unpredictable conditions under which ethical subject-formation can unfold. When the "ideal learner" is defined by algorithmic coherence, there is limited space for unpredictability and discontinuity (Biesta, 2013), for error as formative, or for ethical development grounded in uncertainty and reflection. In such environments, dissent, creativity, resistance, and questioning are easily pathologized as "deviations" rather than valued expressions of autonomous thought and individual agency. Legal privacy protections do not safeguard these symbolic and existential dimensions. They cannot ensure that learners experience anonymity as a protective veil for experimentation, or that secrecy and intimacy

remain available as spaces of becoming, particularly in socio-technical ecosystems that equate visibility with value and standardisation with fairness. To think ethically about AI in education, we must move beyond compliance and ask: What does it mean for students to be truly private? What role does opacity play in ethical learning? And how do we preserve the conditions under which learners can choose not to be seen, tracked, or predicted?

At an intersubjective level, agency is shaped through interactions between students, teachers, and other stakeholders, including families. AI-based systems can alter power dynamics in the classroom, potentially undermining teachers' authority and sense of professional agency. The Council of Europe (2022) has noted that these technologies may not only limit students' autonomy but also reduce teachers' ability to make pedagogical decisions. Additionally, AI-driven systems designed to support teachers risk reinforcing inadequate pedagogical practices or even automating poor teaching methods, thereby disempowering both educators and parents. Pedagogical agents, introduced over two decades ago to replicate human interaction and emotional engagement, have increasingly integrated affect (a critical component of learning), as positive emotions have been shown to enhance both learning experiences and academic performance (Dobrosovestnova & Hannibal, 2020). Despite their potential benefits, however, PAs have also raised concerns about the development of virtual dependencies between students and AI agents (Hudlicka, 2016).

At a collective level, AI in education raises ethical and institutional concerns related to policy, collective decision-making, political engagement, and democratic participation. AI-driven school administration relies on extensive learning analytics, raising serious privacy, trust, and fairness concerns. Mood analysis and activity logs pose risks related to political profiling and ethnic identity (Tundrea, 2020). These challenges also question ethical boundaries, as the collected data often extend beyond individual learners to include their peers and families. Moreover, different educational and governmental stakeholders, such as private organisations, developers, government agencies, research centres, universities, and schools, hold varying perspectives on AI ethics, creating tensions in the regulation and implementation of AI in education (Popenici & Kerr, 2017; Tuomi, 2018).

In addition to these socio-political and ethical considerations, the ecological dimension of AI systems in education must not be overlooked. Every AI system is inherently entangled with ecological systems, from the extraction of raw materials and energy consumption during training, to the waste and emissions generated through deployment and disposal. This environmental entanglement adds a layer of complexity to ethical evaluation, demanding that

educational institutions and policymakers consider not only the human implications of AI use, but also its sustainability and long-term impact on the planet.

Finally, the rise of Generative AI tools presents additional challenges: it concentrates technological power and deepens digital inequality, outpaces national regulation, and often relies on data collected without consent. This technology also enables the creation of convincing deepfakes and other forms of disinformation, underscoring the need for strong safeguards, critical engagement, and clear governance (UNESCO, 2023a). Furthermore, it offers interactive, adaptive learning experiences and are increasingly integrated into daily life. As GenAI systems become more embedded in human routines, they create a sense of shared agency between users and AI, raising critical ethical and practical questions that require further examination (Godwin-Jones, 2024). As they mediate what is seen, recommended, flagged, or evaluated, they foster a form of shared or proxy agency that disperses responsibility across human and non-human actors. In such conditions, the locus of ethical reflection becomes diffuse. When pedagogical decisions are increasingly filtered through algorithmic systems, the opportunity for shared reasoning, public engagement, and contestation is diminished. Schools risk shifting from being democratic spaces of experimentation and imagination to sites of silent automation, driven by opaque systems and abstract metrics.

In response, it is key to foreground the relational axis as a core component in the intersection of AI and education. "Entangled pedagogy" sees technology not as a separate or neutral tool, but as inseparable from pedagogy, ethics, and context. Using technology in education always involves considering what is taught, how it's taught, with what values, and who is helped or harmed by it (Fawns, 2022). The ethical tensions provoked by AI are not marginal: they shape the very conditions under which educational meaning is made and transmitted. These tensions begin in language and ripple outward through symbolic and epistemic processes. First, there is the opposition between history and testimony. While AI can simulate or reconstruct historical facts, it cannot testify. Testimony is personal, embodied, and ethically grounded, it bears the weight of experience. Without it, history risks losing its affective and moral anchoring. Second, there is the contrast between information and memory. AI can process and recall data at scale, but memory is more than recall; it is interpretive, emotional, and narrative. Meaning does not reside in pattern recognition alone but emerges from resonance within human experience. Third, the tension between networking and bonding. AI can connect individuals through optimised networks, but it does not forge bonds: the unpredictable, the encounter that arises through alterity, vulnerability, and care.

In this context, this research challenges the depoliticising and deprofessionalising effects of technocratic approaches, encouraging teachers to resist uncritical automation and algorithmic normativisation.

1.1.4. Coverage of AIED Ethics in Research Policy and Training

Given the challenges noted above, the present study attempts to explore how to conceptualise the ethics of using AIED in learning environments. Although the ethics of general AI has been the focus of a large number of studies and regulations (Dignum, 2021; Jobin et al., 2019; Nguyen et al., 2022; Whittaker et al., 2018), AIED impact has not been adequately documented in the research literature (Holmes et al., 2019; Hrastinski et al., 2019; Zawacki-Richter et al., 2019). Furthermore, systematic educational policies for AIED are still a mirage. By 2021, countries like China, India, Italy, Kenya, Malta, Singapore, South Korea, Spain, and the United States were still debating AIED in their policies, but only five of them included it in the context of their AI policies (Schiff, 2021). In 2024, government experts, representatives from international organisations, private sector companies, and civil society organisations were still convening to develop a legal instrument for the education sector that complements the Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law (Council of Europe, 2024).

With regard to the use of AI in schools, before GenAI became available to the general public in 2022, capacity building focused primarily on its technical component (and not on its human dimension) and it was almost only aimed at secondary and tertiary education in computer science courses. In fact, there was a lack of teacher and parent education and limited training opportunities for the general public (Miao et al., 2021). For educators who had no experience with AI, training considered simple aspects of social impact, related to human interactions, and work-related issues (Kowch, 2019). Conceptual AI and learning, along with the ability of existing learning theories (e.g., sensemaking, decision-making, and self-regulation) to integrate new constructs, were largely overlooked (Gasevic et al., 2023). Following the introduction of GenAI, concerns about teacher training increased significantly. UNESCO's AI competency framework for educators responded by establishing a three-stage progression that encompasses fostering a human-centric approach, addressing AI ethics, understanding AI fundamentals and applications, integrating AI into teaching practices, and leveraging AI for professional development (UNESCO, 2023b). A practical example of this was the AI4T (2024a) initiative, an Erasmus+ K3 project led by France, Slovenia, Italy, Ireland, and Luxembourg, aimed at providing AI education specifically for teachers and school leaders.

The initiative included understanding the ethical implications of AI use in educational settings as a key component of the professional development pathways shared across the participating countries. Despite these efforts, many proposals remained insufficient, primarily focusing on content-related issues and generic pedagogical integration, while failing to adequately address the andragogic needs of teachers as a central element of training in education ethics. In fact, there is broad consensus that it is urgent to support teachers and students with the competencies required for the responsible and ethical use of GenAI tools. This viewpoint underscores the importance of moving beyond superficial adoption, emphasising the need for focused professional development and ethics-based training (García-Peñalvo et al., 2024).

In response to these policy gaps, initiatives like the “Safe AI in Education Manifesto” have emerged to promote ethical and responsible use of AI in education. The manifesto outlines human-centred principles and offers practical tools, such as an integration checklist, to guide educators and institutions in evaluating and adopting AI technologies (Alier Forment et al., 2024). Complementing these efforts, theoretical-practical approaches have been employed to develop training models that integrate ethical reflection on Generative AI within computer engineering education, using case studies based on real coursework. García-Peñalvo et al. (2025) specifically examine this approach, highlighting Generative AI tools developed to be part of the learning process, which serve both as support for the educational programme and as objects of analysis to evaluate their compliance with the SAFE-AI principles outlined in the “Safe AI in Education Manifesto”. Their findings demonstrate that the experience gained through studying a concrete use case in the educational context supports the viability of this methodology in bringing theoretical concepts closer to the professional reality of students.

As AIED becomes part of learning environments, its ethical framework must clearly reflect the ethics of education. This includes considerations such as teacher expectations, the allocation of resources and expertise, gender and ethnic biases, behaviour and discipline, the accuracy and validity of assessments, the quality of knowledge, teacher roles, power dynamics between teachers and students, and pedagogical approaches such as constructivism (Holmes et al., 2021). More specifically, AIED must address the accuracy and validity of assessments, define what constitutes valuable knowledge, support educators' roles and agency in choosing pedagogies suited to their learners, and remain sensitive to diverse approaches to pedagogy (Porayska-Pomsta et al., 2023). In this process of AI integration into classrooms, education is expected to be preserved as a democratically-run space for the formation of public thought, language and concepts about social, economic, political, cultural, ethical and caring life

(Lynch, 2022). This implies that education ethics preserves and enhances human agency, which in turn supports accountability, behaviour regulation, and ethical reasoning (Bandura, 1999).

1.1.5. Framing the Shift from Ethics to Agency

This doctoral research began with a focus on ethics as the primary lens for investigating the implications of Artificial Intelligence in education. Early engagement with the field revealed broad international agreement on core principles such as transparency, fairness, non-maleficence, responsibility, and privacy, as outlined in documents reviewed by Jobin et al. (2019). However, despite their normative appeal, these principles often translate into abstract and decontextualised frameworks, what Floridi and Cowls (2019) term “checklist ethics”, that remain detached from the affective and relational dynamics of educational practice.

To move beyond this abstraction, virtue ethics, particularly Shannon Vallor’s (2016) emphasis on cultivating techno-moral virtues such as empathy, humility, and techno-moral wisdom, offers a more situated and character-based approach. Yet while virtue ethics enriches reflection on individual dispositions in contexts where AI technologies are used, it does not comprehensively address the collective and procedural dimensions of ethical norm debate. Specifically, it offers limited guidance on how ethical claims could be negotiated within diverse educational communities.

Habermas (1991, 1998) provides a complementary perspective with his discourse ethics by shifting the focus from internal dispositions to the procedural justification of norms. Rather than relying on tradition or authority, he argues that ethical validity depends on open, inclusive, and rational dialogue among all those affected. His theory of communicative action distinguishes between instrumental action, oriented towards achieving results, and communicative action, aimed at fostering mutual understanding. In this view, ethics emerges through dialogue, as participants engage in reasoned discussion to resolve disagreements. This framework resonates with the dialogic nature of educational spaces.

However, discourse ethics also encounters limitations. It presupposes ideal conditions such as equality of voice, freedom from coercion, and rational deliberation, which are often challenged in everyday school life. Furthermore, the increasing use of AI systems in educational practices introduces new forms of opacity and asymmetry that complicate communicative interaction. These concerns align with Habermas’s (1985) concept of the colonisation of the lifeworld, which becomes increasingly relevant in the context of AI in education. As algorithmic systems mediate processes such as assessment, progression, and

behaviour management, they bring strategic and administrative logics into domains that depend on human understanding, cultural practices, and ethical judgement.

While discourse ethics offers a consistent framework for thinking about the collective justification of norms, it remains difficult to translate into pedagogical action without further elaboration. Therefore, this research contends that discourse ethics must be grounded in a developmental approach that builds the conditions for ethical dialogue over time, through agency. This move from ethics to agency was not solely a conceptual repositioning prompted by the limitations of existing normative frameworks. It was also driven by critical empirical developments, particularly the rapid evolution of AI technologies and their growing influence in educational contexts.

These overlapping pressures, conceptual, socio-technical, and practical, led to a reframing of the research focus: from ethics as a theoretical lens to agency as a more actionable and pedagogically meaningful construct. These tensions can be more clearly understood through three key inflection points. The first was empirical and socio-technical: with the accelerating development and diffusion of GenAI through large language models (e.g., ChatGPT, Gemini, Copilot), it became untenable to conceptualise AI solely as a neutral “tool” for integration. These systems increasingly exhibit autonomous behaviours, generate content unpredictably, and shape decision-making processes, often blurring lines between human and machine agency. The new landscape called for a reframing: AI technologies were no longer neutral enablers of learning, but actants in Latourian terms: systems that shape practices, values, and even forms of subjectivity (Latour, 2005).

This evolution renders many existing EdTech integration frameworks (such as TPACK, SAMR, or DigCompEdu) insufficient. These models primarily focus on how teachers should use technology, rather than interrogating how such technologies may restructure pedagogical relationships, authorship, epistemic authority, and educational agency, which are all core challenges introduced by GenAI. As scholars like Metzinger (2019) and Zuboff (2019) have argued, AI now forms part of a broader socio-technical infrastructure that calls for democratic accountability and ethical reflexivity, not merely technical proficiency.

A second inflection point arose from a conceptual review of ethical frameworks typically applied to AI and education (Mouta et al., 2024a). Three key limitations were identified. One includes epistemic ambiguity, referring to the fact many AI ethics frameworks collapse into principle-based lists devoid of grounding in ethical theory. This creates a “checklist morality”, where ethical considerations are performed rather than deeply interrogated. Furthermore, such models rarely clarify whether they are drawing on deontological, consequentialist, or virtue-

based traditions. This lack of clarity frequently leads to ethical relativism or tokenistic applications of ethics that avoid addressing underlying value tensions. In education, where values must be explicit and negotiated, this ambiguity is particularly problematic. Another limitation concerns fragmented competencies: existing frameworks often present ill-defined or overlapping categories and fail to distinguish between personal, interpersonal, and systemic dimensions. They tend to suggest that ethical understanding can be “taught” in isolated modules (e.g., UNESCO’s AI Competency Framework for Teachers/for Students). However, ethical capacity develops relationally and over time; it is dialogical, embodied, and closely tied to human flourishing (Coeckelbergh, 2020; Nussbaum, 2011).

A third and equally significant inflection point was practical. Despite the proliferation of AI ethics guidelines, many fail to distinguish between policy, pedagogy, and school-level practice, often overlooking educators’ lived experiences, institutional constraints, and affective responses to technological change. As a result, ethics is frequently reduced to compliance rather than enacted as a critical, participatory, and transformative process.

Confronting these limitations, this research re-centres agency (defined through Bandura’s (2001) components of intentionality, forethought, self-regulation, and self-reflection) as the developmental and pedagogical core of the course design. This framework makes three critical contributions. First, in terms of epistemic grounding, rather than assuming a universal or abstract ethical code, the course design adopts a situated and dialogic ethics (Habermas, 1991), allowing ethical reasoning to emerge from relational encounters and pedagogical reflection. Teachers engage with ethical ambiguity not by following prescriptive checklists, but by negotiating meaning, values, and responsibilities within educational specific contexts. Second, for conceptual coherence, the course avoids fragmented taxonomies by embedding competencies into interrelated domains of agency (subjective, intersubjective, and collective) which align with Biesta’s (2013) educational functions of qualification, socialisation, and subjectification. Third, with regard to implementation depth, the professional development course is not presented as a generic curriculum but as a participatory prototype designed to operate across pedagogical, strategic, and policy levels.

1.1.6. Agency as a Basis for Ethics: Accountability and Meaningful Reasoning

Agency is fundamentally tied to ethics, as it is the foundation upon which accountability is built. From an ethical standpoint, agency is not just about the ability to act but also about the capacity to reflect on one’s actions, anticipate consequences, and take responsibility for them. It also refers to how structural changes produce new conditions that stimulate responses from

individuals or groups, who then actively reshape those same structures (Archer, 2015). Bandura (1999) argues that weakening agency disrupts moral judgment because accountability relies on an individual's ability to see themselves as an agent in their own choices. Further, agency plays a crucial role in ethical AI governance. As Bandura (2006) highlights, reflecting on the values and social structures that shape technological power is critical in managing its influence on human agency. Without considering agency, ethical frameworks risk losing depth, as accountability may become unclear between human actors and machine actants unless there are clear mechanisms in place to define responsibility. This also overlooks the opportunity to thoughtfully position socio-technical systems within a continuum of proxy and distributed agency.

In human-machine interactions, a robust sense of agency ensures that individuals can take ownership of their decisions, attribute responsibility appropriately, and engage in ethical reasoning. In fact, agency may be diminished through automation that bypasses human intervention, opaque AI decision-making, or learned dependence on technological systems (Berberian et al., 2012; Moore, 2016; Obhi & Hall, 2011). In these situations, ethical responsibility becomes diffuse, making it difficult to determine who is accountable for actions and outcomes. If AI systems mediate, influence, or even override human decisions in educational settings, they risk detaching students, teachers, and institutions not only from their ethical responsibilities but also from their common ground and collective agency. This detachment can manifest in over-reliance on AI-driven feedback or assessments, diminished critical engagement with learning processes, and an erosion of professional judgment in teaching. This is particularly critical in education, where ethical considerations shape not only learning experiences and students' moral development but also how they construct meaning during their academic journey.

To address this, the thesis advances an agency-based framework, developed through a participatory design process, which provides conceptual coherence by responding to the common conflation of values, competencies, and outcomes in existing educational models. This approach articulates a triadic structure of educational agency: (1) subjective agency, referring to the capacity for individual reflection, ethical deliberation, and self-regulation; (2) intersubjective agency, which encompasses the ability to engage relationally, co-construct meaning, and participate in dialogical reasoning with others; (3) and collective agency, understood as the power to influence or transform institutional and socio-technical structures through informed collaboration and action. Collective agency reinforces schools as democratic spaces that foster personal growth and meaningful participation in society (Dewey, 1916).

Therefore, the ethics of education should ensure that AI systems go beyond merely enabling qualification, and instead actively support education as a symbolic process that recognises alterity and enables individuals, groups, and institutions to engage in imagination, intention, memory, interpretation, meaning-making, experimentation, and the pursuit of bold new beginnings.

In this sense, this thesis positions agency not just as a subject of concern in AI ethics, but as a critical precondition for ethical action itself. Strengthening agency in education means fostering environments where students and educators manage their own development and ensure the agency of their relationships and collective experiences. It means maintaining control over decisions, understanding AI's role in learning, and critically engaging with technological systems: actively and collectively shaping the narratives through which these systems are designed, built, supplied, and used. Ensuring that AI enhances rather than diminishes agency is a fundamental ethical imperative, as it is key for maintaining accountability, moral reasoning, collective intelligence thresholds, meaning-making, and the human-centred values that education seeks to uphold.

1.2. Research Design

This research project is rooted in a dual ambition: to respond meaningfully to an urgent and underexplored educational need, and to do so through a participatory, ethically engaged design process that prioritises agency as both a conceptual and pedagogical cornerstone. What is most innovative about this work is not just the introduction of AI into teacher training (an effort that gained momentum globally after the pressure created by the rise of Generative AI in 2022), but the deliberate and critical focus on agency. This theme, while increasingly recognised as central to AI ethics (Holmes et al., 2022; Holmes, W., & Tuomi, I., 2022), is rarely and unclearly addressed in existing teacher education initiatives. This gap was a key finding of the preliminary research and forms the conceptual core of the course developed.

A key contribution of this thesis lies in its use of educational design research (EDR) as a form of risk-taking in the Biestan sense: affirming education as an unpredictable, relational, and ethically situated process (Biesta, 2013). Rather than aiming to test a fixed model under experimental conditions, the project traces the iterative development of a prototype course and accompanying online platform that together constitute a formative educational response to a rapidly evolving and conceptually unstable field. In a context where technological, ethical, and educational norms are constantly shifting, the aim is not to suspend meaning, but to design a

stable and contextually relevant framework that remains open to dialogue, change, and reflection.

1.2.1. Research Questions

In response to the conceptual questions outlined above around ethics, agency, pedagogy, and socio-technical entanglements, this study is guided by the following research questions:

1. What are the ethical considerations associated with the integration of socio-technical AI systems in educational settings?
2. How can a conceptual framework on AI, ethics, and educational agency be co-designed to inform pedagogical practices that preserve and strengthen human agency in educational environments where AI is present?
3. How can a training course for K-12 teachers on the ethics of education with AI be designed and implemented using an Educational Design Research approach?

These questions reflect a shift from treating AI as a neutral tool toward a more agency-oriented interrogation. They inform both the research and course design processes, ensuring that decisions remain anchored in critical dialogue, educational ethics, and collective agency.

1.2.2. Methodological Framework: Educational Design Research

Considering the repertoire of concerns and challenges presented in the introduction and the research questions, this doctoral dissertation advocates a research approach that is participatory, iterative, and practice-oriented. To achieve this, educational design research has been selected, a methodology particularly suited to contexts where practical solutions need to be developed for emerging challenges (Kelly, 2006).

EDR combines scientific inquiry with the systematic design and implementation of educational interventions, ensuring that research is both theoretically robust and practically relevant (McKenney & Reeves, 2018). Unlike purely quantitative research approaches that prioritise hypothesis testing under strictly controlled conditions, EDR embraces the inherently fluid, organic, and context-bound nature of educational practice. This flexibility is particularly critical in emergent domains such as the integration of AI socio-technical systems in education, where technological capabilities, pedagogical possibilities, and ethical implications are still in flux. In particular, the rise of Generative AI introduces novel forms of agency into educational contexts, prompting the need for renewed negotiations of meaning, purpose, and pedagogical practice. In such a setting, fixed-variable, outcome-driven methodologies risk oversimplifying the nuanced and contested realities of classroom practices and the enactment of meaning,

especially when they prematurely attempt to evaluate interventions using experimental designs before understanding their ethical and societal dimensions.

EDR is not without its critics. Brown (1992) identified several methodological vulnerabilities within design research, including inadequate attention to sampling and response bias, researcher bias, confounded variables, unmanageable volumes of data with limited strategies for transforming data into evidence, poor scalability, and insufficient focus on dissemination and diffusion. Yet it is important to recognise that such challenges are not unique to design research. Similar methodological tensions have historically emerged in the evolution of experimental and correlational research traditions in education and psychology. As Shadish et al. (2002) note, even traditional experimental designs required decades of refinement following the foundational work of Fisher in the early twentieth century.

EDR does not reject the value of quantitative approaches but rather precedes and informs them. It can be conceived as a necessary first step in the research continuum, helping to articulate the questions, variables, and ethical considerations that may later be tested through more controlled empirical methods. Through iterative cycles of design, implementation, reflection, and revision, EDR allows researchers and practitioners to collaboratively shape interventions that are contextually meaningful and theoretically robust, interventions that may later be subjected to experimental scrutiny. In this way, this methodology plays a foundational, hypothesis-generating role, clarifying what should be tested and under what ethical and practical conditions such testing would be justifiable.

This is particularly critical in the case of AI in education, where the ethical stakes are profound and unresolved. Educational interventions that involve students' data, algorithmic bias, and dimensions of individual or collective agency carry ethical risks that cannot be fully anticipated or mitigated through metrics alone. While ethical principles are expected to inform both qualitative and quantitative research, their enactment in real-world educational settings is deeply performative, situated, and relational.

Crucially, deploying AI-based interventions in classrooms without prior iterative exploration of their pedagogical relevance, potential risks, and ethical implications, in close collaboration with teachers, would represent a fundamental ethical misstep. Treating such technologies as neutral tools to be integrated into quantitative studies, without first understanding their contextual value and consequences through participatory design, can overlook the relational nature of education and the agency of those most affected. Following current ethical discourse, it is advisable that practice in this domain begin with iterative, dialogic engagement rather than with experimental rollout.

For this reason, EDR is particularly well-suited to performative ethical inquiry: it allows researchers and participants to co-construct ethical meaning in action, exploring not just what works, but what matters, and for whom, in which contexts, and under what constraints. These reflective spaces are key precursors to any eventual experimental design, ensuring that what is tested is not only effective but also collectively negotiated: a translation of collective agency, which is fundamental to any autonomous and democratic educational institution.

The co-construction of perspectives and solutions grounded in contextualised pedagogical environments ensures that AI's role is framed through the lens of human agency, and education ethics. By fostering a continuous dialogue between theory and practice, this study aims to develop a training course for K-12 educators that directly responds to their needs in integrating AI meaningfully.

1.2.3. Research Phases and Thesis Structure

This doctoral thesis has been developed as a collection of articles that document the research conducted throughout the study. While several chapters are based on published work, some sections present original contributions that are yet to be published. The following chapters are organised into three main parts that document the different research phases.

This doctoral thesis is organised into four main parts, each corresponding to a distinct phase of the research process.

Part 1, titled “Theoretical Foundations”, comprises Chapter 2, which presents a Systematic Literature Review (SLR). This chapter maps the current landscape of AI adoption in education, highlights key ethical gaps in existing research, and establishes the conceptual and empirical foundation for the study.

Part 2, “Empirical Research”, encompasses Chapters 3 to 6.

Chapter 3, Futures Studies, employs the Delphi Method to engage experts in constructing a set of future scenarios that surface anticipated ethical challenges in AIED.

Chapter 4, Participatory Collectives, presents findings from iterative focus groups with teacher educators, providing insight into how educators conceptualise and respond to ethical and agency-related tensions in AI-mediated educational contexts. These insights inform the development of a framework for structuring professional development grounded in ethics, agency, and educational purpose.

Chapter 5, The Shifting Nature of Agency in Education with AI, explores how teachers perceive and negotiate agency through qualitative analysis of focus group data, ultimately

advancing a three-layer conceptual framework of agency (subjective, intersubjective, and collective) in educational contexts involving AI.

Building on these findings, Chapter 6, Course Design and Evaluation, outlines the co-design and preliminary evaluation of a research-based professional development course for K-12 teachers. The course is structured to support critical engagement with AI through dialogic, experiential, and reflective learning practices.

Part 3, Conclusions, includes Chapter 7, which synthesises the study's key findings, discusses their implications for educational policy and practice, and proposes future directions for research in the field.

Finally, Part 4, Research Dissemination and Engagement, presented in Chapter 8, details the dissemination of research outputs through academic publications, and conference presentations. It also presents the research outputs, the networks and groups through which the findings were disseminated, including policy fora concerned with the ethics and governance of AI in education.

The chapters outlined above build upon one another to provide a comprehensive exploration of the ethics of education in the context of AI socio-technical systems, spanning theoretical foundations, empirical investigations, and future implications.

Figure 1 below graphically illustrates the EDR journey, to support understanding of how each finding builds on the previous one, leading up to the design and deployment of the course in a micro-explorative evaluation.

Figure 1

Overview of the EDR Research Trajectory.



CHAPTER TWO

State of the Art

This section presents the results of a systematic literature review conducted in 2021 and supplemented in 2022 using the PRISMA methodology. Serving as the foundation for this research, the review was later published in the “International Journal of Artificial Intelligence in Education” and is included in full as Appendix A. The SLR aimed to systematically map the extant literature to understand the scope of AI adoption in education, identify key themes, and highlight ethical considerations that had been largely overlooked in the research landscape from 2011 to 2022. At the time, there was a paucity of studies addressing the ethical implications of AI in educational settings despite the increasing integration of AI technologies into learning environments. Through this approach, the review not only provided a comprehensive synthesis of the state of the art in AI in education but also highlighted critical gaps requiring further research and development, particularly regarding the ethical challenges educators and learners face when AI technologies are integrated into pedagogical practices. Exploring these gaps through the SLR was key to informing the subsequent stages of the thesis, which further elaborate on teacher training needs in a way that preserves both epistemic and collective agency.

2.1. Introduction

The growing integration of AI in education has led to a wide range of applications aimed at enhancing learning and institutional processes. A literature review of AI technologies in education from 1993 to 2020 identified various applications, including school management, student learning, teacher support, and lifelong education. These technologies encompass chatbots, intelligent tutors or agents, machine learning, personalised learning systems, and visualisation tools (Zhang & Aslan, 2021). While these systems offer innovative ways to support learning, concerns about their broader implications for agency and ethical considerations persist. Ethical reasoning, in this context, provides a critical framework for addressing concerns and guiding sustainable transitions, particularly in education, where the “when, how, and to what end” of AI’s pedagogical use must always be questioned (Holmes et al., 2021).

Despite the increasing presence of AI in education, global AI policies often fail to address its ethical implications. An analysis by Schiff (2021) found that these policies primarily frame education as a means to ensure a steady supply of AI talent, rather than considering its ethical dimensions. Additionally, there is a lack of training opportunities for teachers, parents, and the general public on the responsible use of AI in education (Miao et al., 2021).

A systematic review of AIED research from 1970 to 2020 highlighted five key areas of focus: 1) adaptive learning and personalisation, 2) deep learning and machine learning algorithms in education, 3) human-AI interaction in learning, 4) AI-generated data in educational contexts, and 5) AI in higher education. Notably, ethical concerns around AIED remain largely underexplored (Bozkurt et al., 2021).

To explore the critical gaps in the literature, the following sections will systematically address the key questions identified in this study, including the evolution of research on ethics in AIED from 2011 onward, the leading authors and contributing countries in AIED ethics, and the interdisciplinary origins of AIED ethics research. Additionally, the review will examine the main technologies, targets, and applications in AIED, considering their associated challenges, cultural inclusivity, emotional monitoring, and ethical concerns. Through this detailed mapping and analysis, the review will also evaluate current efforts in capacity building on AIED ethics and assess existing principles, regulations, and frameworks guiding the field.

2.2. Methodology

Following the discipline of a systematic literature review, this research screens the most relevant data from the existing literature, responding to pre-defined eligibility criteria (Ramírez-Montoya & García-Peñalvo, 2018). The next lines acknowledge the nature of the debates on AIED ethics and how ethics are screened in AIED studies.

2.2.1. Planning

As shown in Table 1, the planning phase began with four mapping questions (MQs) and a set of four inclusion and exclusion criteria to globally screen what has been published in AIED and ethics over the last 10 years. To ensure a comprehensive review considering the global scope of this study, the researchers' language proficiency was taken into account and multiple languages were included. Due to the time frame for completing the review, it was not possible to add more than four languages, as translation would depend on resources not immediately available. To allow for a more in-depth analysis, five research questions (RQs) were added. These research questions are intricately connected to the themes and concerns raised in the introduction.

They provide the necessary direction and focus to fulfil the paper's primary motivation, which is a thorough examination of the shortcomings in the literature, limited to the chosen timeframe, regarding the ethical aspects of AIED and their impact on educational practices. These questions offer a structured approach to identifying current applications and the challenges they may bring (RQ1), disparities in terms of data ownership, access,

implementation quality and contextualisation (RQ2), emotion surveillance (RQ3), focus and targets of capacity building for an ethical use of AIED (RQ4) and available regulations to safeguard the compliance of ethical principles in AIED design and implementation (RQ5). As stated in the introduction, the knowledge obtained from these RQs would be beneficial for promoting ethical responsibility through dialogic practice, as it addresses several unresolved aspects in AIED use, including the balance between access and quality (of implementation), the tension between universality and context, the interplay of cognition and emotions, and the normative guidance provided by targeted training and regulations.

The PICOC method (Petticrew & Roberts, 2008) was used to accurately structure this review. The deliberate use of broader keywords in this SLR was necessary because including research question terms directly in the search strings limited the results. Broader keywords were chosen to encompass the terminology variations used by authors when addressing the same research questions. This approach allowed for a comprehensive exploration of the AIED ethics field and the flexibility to fine-tune the search based on the results obtained. Considering the comprehensive focus of the research, Scopus and Web of Science were the selected databases as they correspond to all-inclusive digital libraries that cover a wide range of academic disciplines within the scope of this study; they were also selected for having logical expressions and both full-length and field-specific articles.

Table 1*SLR Planning Phase*

GOAL	TO UNDERSTAND HOW ETHICS HAS BEEN COVERED IN AIED STUDIES
PICOC	<ul style="list-style-type: none"> • Population (P): AIED and Ethics studies • Intervention (I): Conceptual models and projects on ethics for using AIED • Comparison (C): No comparison • Outcomes (O): Frameworks and projects/programmes of AIED and ethics • Context(C): k-12 Education, Higher Education, LLL
MAPPING QUESTIONS	<p>MQ1. What has been the evolution of the number of research articles concerning ethics in AIED since 2011?</p> <p>MQ2. Who are the key authors in AIED ethics?</p> <p>MQ3. Which countries exhibit the most significant productivity in terms of disseminating conceptual or theoretical approaches, frameworks, or interventions in the realm of AIED ethics?</p> <p>MQ4. From which subject areas are coming the main studies in AIED ethics?</p>
RESEARCH QUESTIONS	<p>RQ1. Which AIED technologies, targets and applications are being considered and what challenges do they pose in general?</p> <p>RQ2. How is AIED considering cultural differences and inclusion?</p> <p>RQ3. How does AIED monitor emotions and what ethical challenges may be at stake?</p> <p>RQ4. How is capacity building on AIED ethics being covered?</p> <p>RQ5. What principles, regulations and frameworks are there for AIED?</p>
SOURCES & STRING SEARCH	<p>SCOPUS</p> <p>(TITLE-ABS-KEY (artificial AND intelligence) AND TITLE-ABS-KEY (education) OR TITLE-ABS-KEY (learning) OR TITLE-ABS-KEY (teaching) AND TITLE-ABS-KEY (ethics)) AND PUBYEAR > 2010 AND PUBYEAR < 2022 AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "French") OR LIMIT-TO (LANGUAGE, "Portuguese")) OR LIMIT-TO (LANGUAGE, "Spanish"))</p> <p>WoS</p> <p>Artificial Intelligence in education (All Fields) or Artificial Intelligence in learning (All Fields) or Artificial Intelligence in teaching (All Fields) and ethics (All Fields) and English OR French OR Portuguese OR Spanish (Language)</p>
INCLUSION CRITERIA	<p>IC1. Publication about AIED and ethics</p> <p>IC2. Data range 2011-2022</p> <p>IC3. English, French, Portuguese, Spanish</p> <p>IC4. Peer-reviewed</p>
EXCLUSION CRITERIA	<p>EC1. Publication about AIED or ethics but not both</p> <p>EC2. Data range before 2011 or over 2022</p> <p>EC3. Other language than English, French, Portuguese, Spanish</p> <p>EC4. Not peer-reviewed</p>

2.2.2. Conducting and Analysing

To manage the final cluster of articles, the results were gathered into a spreadsheet ([cf. https://docs.google.com/spreadsheets/d/1u6ArXbQ5w4bxhHbKAcrUCZLUSACK0AoG/edit?pli=1#gid=1005218435](https://docs.google.com/spreadsheets/d/1u6ArXbQ5w4bxhHbKAcrUCZLUSACK0AoG/edit?pli=1#gid=1005218435)), which presents the analysis in several stages (Figure 2, below). In the identification phase, 250 publications were retrieved from Scopus and 249 from WoS. The duplicate papers were eliminated, and a selection of 410 papers was made based on their title and abstract. For the preliminary screening phase, the alignment of the paper's topic with the research goal, the PICOC elements (population, intervention, comparison, outcomes, and

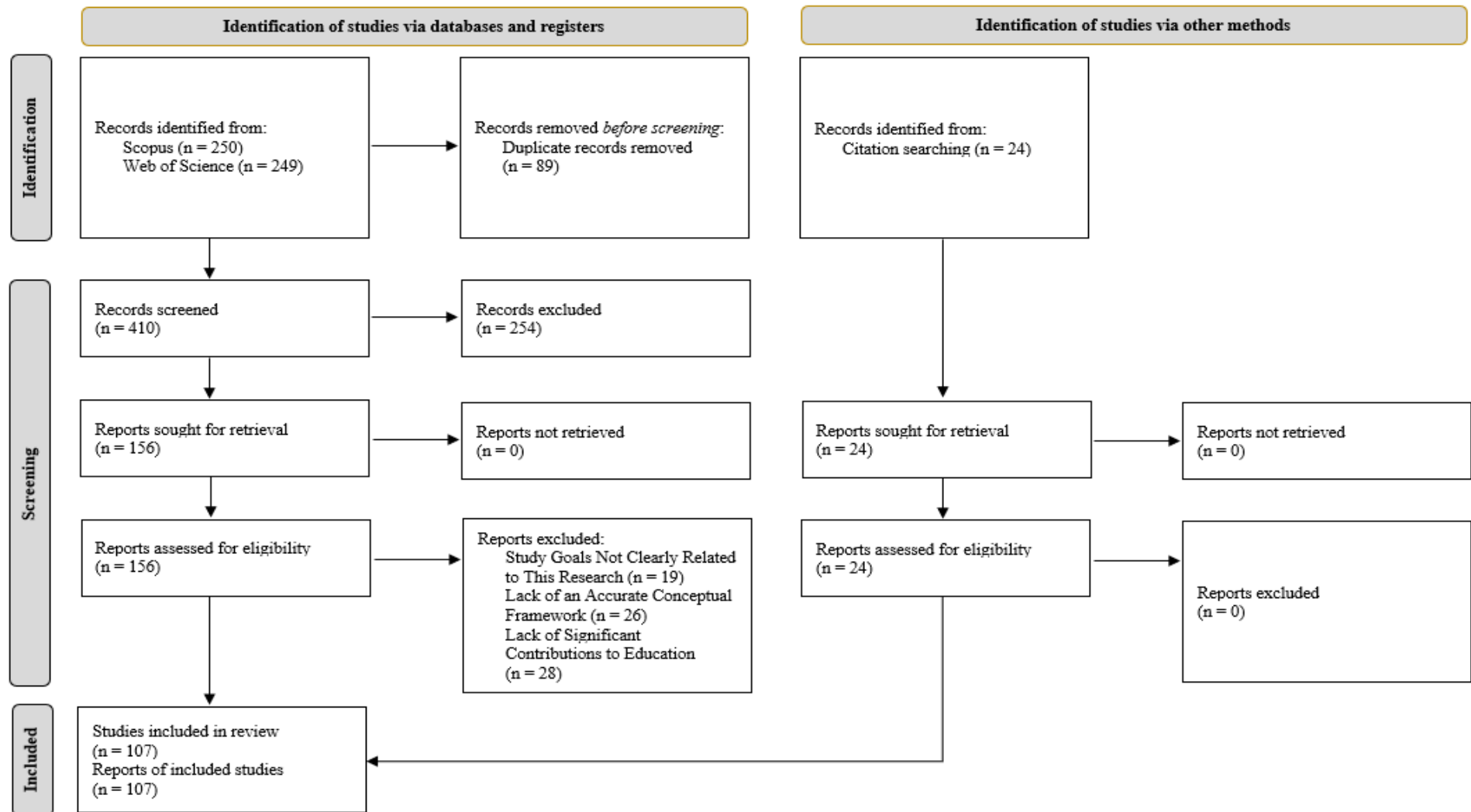
context), the paper's implicit (e.g., using terms such as “AIED challenges”, “FATE”) or explicit relation to AIED ethics, and the fulfilment of the inclusion/exclusion criteria (Table 1, above) were taken into account. If in doubt, the entire article was read. This preliminary screening phase (abstracts) enabled the gathering of 156 contributions.

After this stage, a quality assessment screening was conducted for several purposes, including ensuring relevance (i.e., that the selected papers directly align with the research goals, objectives, and topic), assessing methodological rigor (i.e., whether the selected papers meet specific quality criteria), evaluating reliability (using a predefined scoring system to measure the assessment's reliability), ensuring a high-quality and relevant sample selection, and verifying data integrity (i.e., that the selected papers are reliable and meet established quality standards). In the eligibility stage, the retrieved articles were thoroughly read and analysed according to the five-criteria quality assessment checklist presented in Table 2, below.

Three options were available for each criterion: Yes (1 point), Partially (0.5 point), and No (0 point). Five points were awarded to items that fully met the defined quality parameters. All papers with four or more points were considered; since four is the cut-off point, all papers with a lower score were excluded from the final sample and from the corresponding RQs. This quality assessment phase allowed the collection of 83 papers. Finally, 24 articles were added to this cluster: they were first or second references, chosen whenever they engaged in relevant work surrounding the RQs, particularly on topics requiring a deeper understanding (e.g., AIED and cultural differences) or further examples (e.g., AIED technologies in use). The reasons for choosing these papers are presented in Table 3, below. As a result of this process, 107 papers make up this sample.

Figure 2

PRISMA 2020 Flow Diagram for the SLR Process: Papers Identification



Note. Adapted from “The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews”, by M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., 2021, *BMJ*, 372, n71 (<https://doi.org/10.1136/bmj.n71>). Copyright 2021 by BMJ Publishing Group.

Table 2*Quality Assessment Questions*

QUALITY ASSESSMENT QUESTIONS	SCORE CRITERIA
Is it a theoretical/conceptual approach or research/study in AIED ethics?	Yes/No/Partial
Do the study's goals clearly relate to AIED ethics?	Yes/No/Partial
Does the theoretical/conceptual analysis or research meet those goals?	Yes/No/Partial
Are the relations between concepts, data, interpretation, and conclusions made clear?	Yes/No/Partial
Does the paper explicitly refer its contribution to think about/question the future of education?	Yes/No/Partial

Table 3*Reasons for the Inclusion of Additional Papers*

PAPERS NOS	REASONS FOR PAPERS INCLUSION	EXPLANATION
1, 2, 4, 7, 8, 10, 12, 16, 18, 19, 20, 21	Directly related to ethics and AI	These papers, directly related to ethics and AI, serve as the core material for examining ethical considerations within the field of education.
3, 6, 11, 24	Relevance for RQ1	These papers offer insights into the technologies, targets, and applications in use, as well as the challenges they may pose.
13, 15, 23	Relevance for RQ2	These papers offer insights into AIED, cultural differences and inclusion.
9, 14, 17	Relevance for RQ3	These papers offer insights into AIED and emotions monitoring.
5	Related to ethics and AI + author specialist in this field	This paper combines the connection to ethics and AI with the expertise of the author.
22	Indirectly related to ethics and AI (children moral development)	This paper is indirectly related to ethics and AI, shedding light on children's moral development within the context of AI.

2.2.2.1. SLR Mapping Analysis

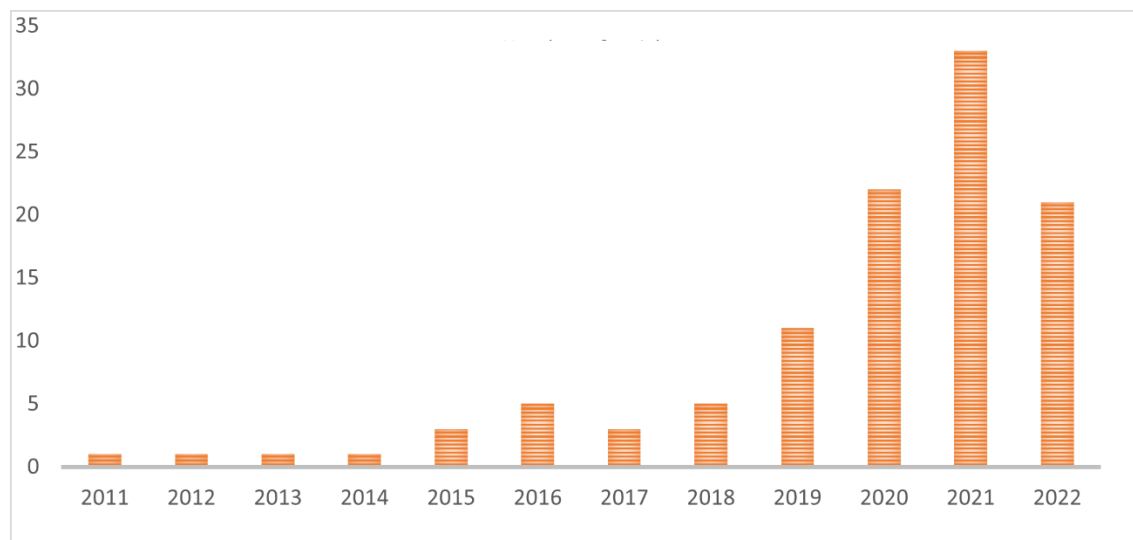
MQ1. What has been the evolution of the number of research articles concerning ethics in AIED since 2011?

As Figure 3, below, shows, between 2011 and 2018, the number of publications on this subject exhibited some fluctuations. This period can be seen as a foundational phase, where discussions around the ethical dimensions of AI in education began to take shape. However, starting in 2019, there was a noticeable shift in the trajectory. AIED research began to experience exponential growth, although there was a small retraction in 2022, with a remarkable surge in the number of articles in 2021. This explosion in publications can be attributed to several factors. Firstly, AI technologies continued to advance rapidly, permeating various educational settings. As these technologies gained traction, there was a corresponding surge of interest in their ethical implications. Additionally, the global adoption of AI in education became more widespread, with researchers and educators worldwide focusing on the

ethical aspects that should accompany its integration. This broader reach led to a surge in research outputs, as experts from various geographies contributed to the discourse. Moreover, the spotlight on AI ethics in mainstream media and entertainment further fuelled academic research and discourse in the field of AIED.

Figure 3

Number of Articles on AIED Ethics Between 2011-2022



MQ2. Who are the key authors in AIED ethics?

Identifying key contributors allows us to recognise those who play a significant role in shaping the discourse within this domain. It provides valuable context to comprehend the methodologies employed in these studies and the perspectives from which they were elaborated. Furthermore, it opens the door to assisting researchers in tracking experts in this specific field of study. In this SLR, Holmes, W. has contributed the most publications with 4, followed by Tuomi, I. with 3. Additionally, there are some authors who have two publications each, including Cukurova, M., Dignum, V., Luckin, R., Mouta, A., Pinto Llorente, A. M., Shum, S. B., and Torrecilla Sánchez, E. M.

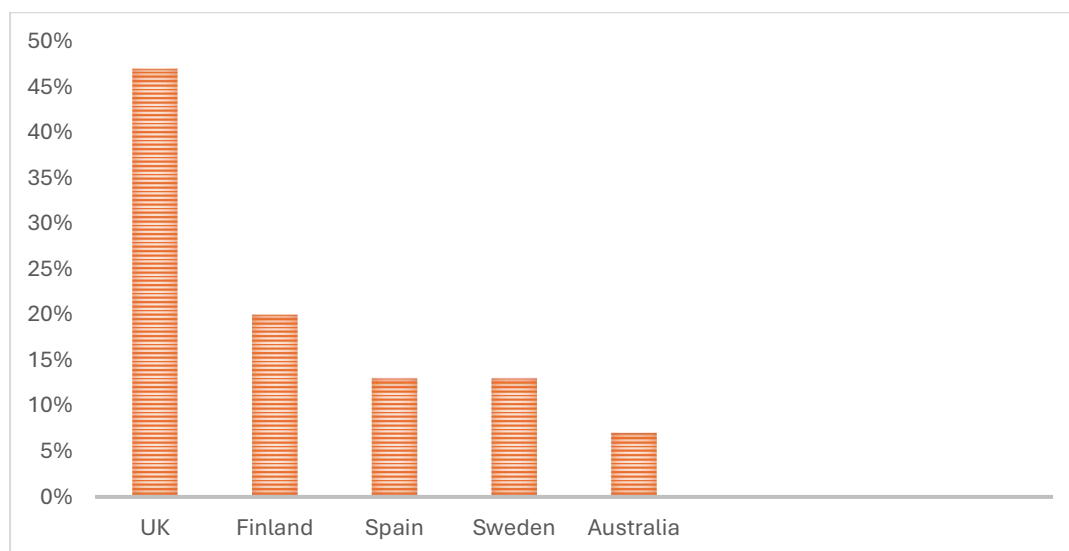
MQ3. Which countries exhibit the most significant productivity in terms of disseminating conceptual/theoretical approaches, frameworks, or interventions in the realm of AIED ethics?

Including this mapping question addresses the need to explore how cultural differences are considered in AIED ethics. Figure 4 shows that 47% of the publications originate from the United Kingdom, with Finland contributing 20%. Spain and Sweden together account for 13%, while Australia contributes 7%. Although most AIED publications during the research period

are attributed to the USA, it is noteworthy that, when the ethical dimension is introduced, a European discourse becomes predominant. This may reflect Europe's long-standing traditions of ethical and philosophical inquiry, which emphasise human rights, individual autonomy, and the responsibilities of technology developers and educators. These values are reflected in the EU's General Data Protection Regulation (GDPR), with its strong focus on data privacy and ethics. The concentration of research in AIED ethics within Europe suggests the presence of regional perspectives and priorities in addressing ethical considerations in educational technology. This underscores the importance of global dialogue that accounts for how AIED's ethical implications are culturally situated and interpreted across regions.

Figure 4

Countries with Publication in AIED Ethics



MQ4. From which subject areas are the main studies in AIED ethics coming?

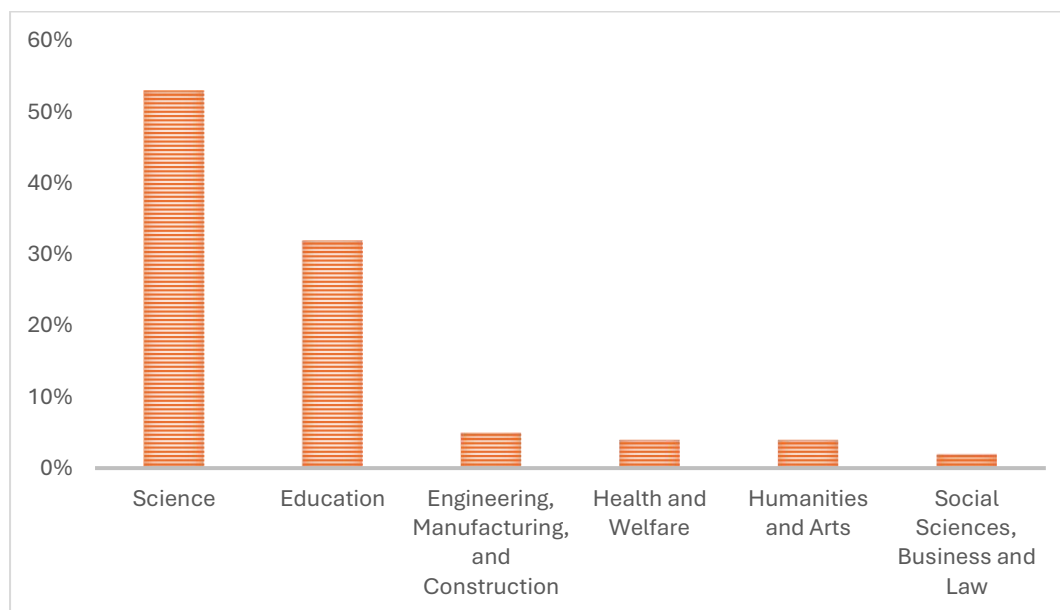
The analysis of subject areas relied on the International Standard Classification of Education (ISCED), developed by UNESCO (2012). This classification system encompasses 25 fields of education organised into 9 broad groups. The main areas that ethical discussions surrounding the use of AIED relate to are shown in Figure 5, below: Science (e.g., computing, life sciences, mathematics) takes the lead, with over 50%. It is followed by Education where 32% of articles may be found. Engineering, Manufacturing, and Construction (e.g., architecture, engineering, manufacturing) follow with 5%. Health and Welfare (e.g., medicine, nursing, social work) and Humanities and Arts (e.g., languages, literature, history, philosophy, performing and visual arts, cultural studies) each cover 4% of the total number of articles.

Finally, Social Sciences, Business and Law (e.g., economics, journalism, political science, psychology, sociology) collectively have a total of 2% of the articles.

The diverse distribution of ethical discussions across academic domains in AIED underscores the need for a comprehensive ethical framework that addresses the unique challenges and concerns within each field. The dominant discourse in science, particularly computing, comprising over 50% of ethical discussions, highlights the central role of technology and data-driven approaches in AIED. While science, notably computing, dominates the discourse, it's crucial to broaden the conversation to ensure that other areas are not overlooked. The scarcity of ethical discussions in humanities and arts suggests a potential gap in recognising the human-centred dimension of learning and pedagogy.

Figure 5

Publications on AIED Ethics Main Subject Areas



2.2.2.2. SLR Research Analysis

The results obtained after analysing and studying the articles supporting this SLR are presented below according to the research questions.

RQ1. Which AIED technologies, targets and applications are being considered and what challenges do they pose in general?

AI technologies and the overall ethical challenges they pose have been the second most studied topic of research on the ethics of AIED between 2011 and 2022. Table 4, below, summarises 36% of the studies in this SLR, presenting AIED technologies, their targets, the

primary educational purposes, and potential ethical challenges. In terms of targets, AIED technologies are being applied to a wide spectrum, including school institutions (K-12) and universities, as well as individual students and teachers. This highlights the broad scope of AIED's influence within the educational ecosystem. AIED applications are designed to serve a multitude of educational purposes. The claims include personalised learning, metacognition, support for learning disabilities, writing enhancement, STEM and programming promotion, virtual schooling, and collaborative learning. These applications are tailored to meet diverse educational needs and objectives.

The risks associated with the use of such technologies are frequently overlooked in the literature, revealing a substantial gap in understanding how automated measurement processes could potentially compromise well-being (Burr et al., 2020). For instance, while profiling higher education students may seem promising in terms of career advancement, it can also be viewed as manipulative, as it fails to acknowledge individuals' freedom of expression and choice (Tundrea, 2020). Some authors draw attention to the potential conflict between the obligation of higher education institutions to fulfil their mandate of ensuring learning and their fiduciary duty to act ethically (Prinsloo & Slade, 2017).

While doubts persist among neuroscientists regarding the potential of certain technologies to improve learning, the risks of inaccurate results or unintended consequences linked to electroencephalography (EEG) remain significant. Despite these concerns, EEG sensors were already incorporated into headbands for detecting students' brain activity (Miao et al., 2021). Therefore, AI has faced criticism for being "dehumanising" in this context, as it often promotes prescriptive teaching with minimal interaction and automated pathways, ultimately diminishing students' agency (Miao et al., 2021).

Another worrisome issue pertains to children aged four to ten who perceive social robots as trustworthy. For example, a talking doll might influence children to reconsider their moral judgments (Williams et al., 2018). This raises further concerns about whether these robots are primarily profit-driven. Additionally, questions regarding privacy emerge as AI delves into children's socio-emotional and cognitive characteristics, as well as aspects of their home environment. Furthermore, transparency issues arise regarding who has access to this data (Mohammed & Watson, 2019).

The significance of robust assessment tools has been underscored by the COVID-19 pandemic. Gudiño Paredes et al. (2021) conducted a mixed-methods study to examine the impact of remote AI-powered proctored exams on the learning processes and academic

integrity of online graduate students. This technology offers facial recognition, audio analysis, eye movement tracking, and object/face detection in the surroundings. While the results indicated a substantial reduction in dishonesty, students reported feeling compelled to cheat under observation, lacking internal motivation or a personal reflective process. They also expressed concerns about privacy and anxiety during assessments conducted in this manner.

Regarding teachers, the possibility of letting these AI technologies automate ineffective pedagogical practices (based in data incompleteness and bias, for example) is also a reason for concern. In the long run, AI-based education may have the potential to disempower teachers (Miao et al., 2021). Another concern revolves around the development of learners' cyborg identities as they engage with AI, the changing relationship between humans and AI systems in a posthuman hybrid dynamic and how it may impact teachers and students (Adams et al., 2022). Finally, most studies assessing the effects of technology are typically carried out by the creators of that technology, often affiliated with commercial entities and involving a limited number of participants (Holmes & Tuomi, 2022).

These insights shed light on the transformative potential of AI in education, recognising how important it is to consider the ethical dimensions that come into play when implementing these technologies.

Table 4
AIED Technologies, Applications, Targets and the Ethical Challenges.

STUDIES	AIED TECHNOLOGIES	TARGETS	PURPOSE	ETHICAL CHALLENGES
Bucea-Manea-Tonis et al., 2022; Chounta et al., 2022; Gallastegui et al., 2021; Holmes & Tuomi, 2022; Kitto & Knight, 2019; Latham & Goltz, 2019; Leonidis et al., 2012; Li et al., 2021; Miao et al., 2021; Pammer-Schindler & Rosé, 2022; Prinsloo & Slade, 2017; Qin et al., 2020; Serholt et al., 2017; Shum, 2018; Tuomi, 2018	Ambient classroom tools, linking the physical & digital worlds (pervasive & multimedia computing, sensor networks, robotics, agent-based software)	Schools (K-12) & Universities	Fees, attendance, retention and dropouts, school inspections, teachers' performance; monitor students' behaviour, provide related data to the classroom's services and applications	Utilitarianism & Deontology: data collected not only related with the learners, but with family; individuals' conflicting claims; interests of private organisations vs. public authorities, schools, universities; AIED instructional implications
Abboud et al., 2020; Adams et al., 2022; Belpaeme et al., 2018; Burr et al., 2020; Córdova & Vicari, 2022; Dobrosovestnova et al., 2020; Ghotbi & Ho, 2021; Gudiño Paredes et al., 2021; Holmes & Tuomi, 2022; Hood et al., 2015; Lameris & Arnab, 2022; Li et al., 2021; Miao et al., 2021; Mohammed & Watson, 2019; Mouta et al., 2019; Pammer-Schindler & Rosé, 2022; Papa & Jackson, 2021; Renz & Vladova, 2021; Serholt et al., 2017; Smakman et al., 2021; Tundrea, 2020; Ungerer & Slade, 2022; Williams et al., 2018; Zhang & Aslan, 2021	ITS, AI-supported reading, smart & social robots, teachable agents, virtual and AR, AI-enabled collaborative learning, emotional facial recognition & regulation, sensing technology (gaze tracking)	Students (K-12, University, Master's and Ph.D.)	Personalised learning; meta-cognition & self-esteem improvement; learning disabilities support; improvement of writing before assessment; promotion of STEM subjects and computer programming; virtual school (impossibility to go); tutors (any part of the world) & peers (collaborative learning)	Human rights & Justice: surveillance causing anxiety; reduction of human interactions; human replacement; students' privacy (activity logs screen political views or sexual orientation), safety, trust, and fairness; sense of agency; self-efficacy; student freedom of choice (e.g., AI career guidance); children' moral development (e.g., social robots, talking dolls); posthumanist ethics of AIEDK-12
Adams et al., 2022; Belpaeme et al., 2018; Coghlan et al., 2021; Ghotbi & Ho, 2021; Holmes & Tuomi, 2022; Leonidis et al., 2012; Li et al., 2021; Miao et al., 2021; Mohammed & Watson, 2019; Tuomi, 2018	Adaptive assessment, automated writing evaluation, AI-driven discussion forum, AI-teaching assistants, "dual teacher"	Teachers	Assessing; helping learners stick to the curriculum; expertise feedback; asynchronous discussion; personalised content & activity; bureaucratic tasks	Justice & Human rights: data extrapolation; no creativity assessment; quantity preferred to relevance; inability to prevent "deep fake" assignments; incorrect correlations deriving from bias, posthumanist ethics of AIEDK-12
Holmes & Tuomi, 2022; Marras et al., 2022; Pammer-Schindler & Rosé, 2022; Pham & Wang, 2017; Westman et al., 2021	Multimodal mobile learning system for MOOCs	Lifelong Learning	Real time feedback; career guidance; speculative AI-assisted LLL assistant	Justice: learning opportunities imbalance among learners; inaccurate data; data ownership, data privacy and consent

RQ2. How is AIED considering cultural differences and inclusion?

Although “the greatest good for the greatest number” is a concern when introducing these technologies, what can be considered good in an educational context is a challenging question. Equity, cultural and interpersonal differences, which are the leitmotiv of 17% of the papers under analysis, are an issue when it comes to massive technologies that may not easily adapt and respond to the specifics of the context and the people who are meant to use them, potentially undermining the fairness and fundamental rights of individuals. Table 5, below, demonstrates that out of the 18 articles included in the SLR, there are four distinct categories of ethical challenges pertaining to cultural and inclusion responsiveness.

The “intercultural” challenge has to do with these systems’ capacity to accommodate different cultural background and learners’ values. Approximately 53% of the papers within this subject area highlight that AI solutions for educational purposes may not sufficiently consider cross-cultural variations. Another challenge has to do with “cultural realism” and the difficulty of representing the diversity of learners’ particular physical characteristics. Some authors defend that AIED is best suited to Western, educated, industrialised, wealthy, and democratic nations (Nye, 2015; Ogan et al., 2015). For example, limited broadband access will leave people out of data sets and AI will be unintentionally biased against them (Miao et al., 2021). There's also the issue of reported misidentifications from facial recognition software in relation to darker skin tones.

This imperatively requires the choice of suitable workaround solutions or the omission of facial recognition (Coghlan et al., 2021). Sanusi & Olaleye (2022) discuss the role of cultural competence and ethics in AI education, and how these factors influence students' learning of AI. It also touches upon the disparities between rural and urban students in terms of cultural and ethical competence. Holmes and Tuomi (2022) discuss AIED colonialism, which includes the adoption of single products in state education systems, language biases, and the imposition of specific pedagogical approaches. This colonialism varies to an extent but often results in well-funded Global North AIED tools overshadowing locally sensitive alternatives.

The “inclusion” challenge refers to the effort of guaranteeing that AIED features recognise learners’ neurological, physiological, and psychological diversity and become accessible and accurate to all. Although it's a challenging and early-stage process, robots, and empathic intelligent learning environments (ILEs) are already being designed to recognise emotional patterns and respond accordingly (Pham & Wang, 2017). This is especially beneficial for children with special needs. For instance, socially assistive robots proved more effective in

reducing repetitive and stereotyped behaviours in autistic children compared to interactions with people (Costa et al., 2018). In the case of autism spectrum disorder, educational games are being used to enhance children's ability to distinguish emotions in a simulated learning environment, aiming to facilitate their transition between the virtual and real-world contexts (Epp & Makos, 2013).

Another study showed that children engaging in a 30-minute daily interaction with a caregiver and a social robot over one month improved their attention skills when the robot was not present (Johnson & Lester, 2016). The robot encouraged emotional storytelling, perspective-taking, and tailored the difficulty of activities based on past performance. These studies demonstrate that empathic robots can create more engaging and fearless learning experiences, especially for K-12 students (Johnson & Lester, 2016). While these advancements offer promising perspectives for inclusion in special education, challenges arise in the context of online proctoring. Technology may not be adequately prepared to interpret the behaviours of neurodiverse individuals or those with disabilities, potentially leading to false positives for cheating (Coghlan et al., 2021).

Finally, these systems may not be able to address issues of harassment, bullying, or discrimination that may occur in online educational environments and ensure that AIED technologies have mechanisms in place to prevent or respond to such incidents. The European project ACACIA (Restrepo et al., 2019), funded in part by Erasmus+, features a chatbot named Artemisa designed to address sexual harassment and recruit volunteers to promote diversity and tolerance at the Peruvian National University of San Marcos. Like other chatbots, it presents accessibility challenges for users, underscoring the need for training in accessibility, tolerance, and diversity acceptance to mitigate biases.

Table 5*Percentage of Papers on AIED Cultural and Inclusiveness Ethical Challenges.*

STUDIES	ETHICAL CHALLENGES	PERCENTAGE OF PAPERS
Coghlan et al., 2021; Dieterle et al., 2022; Holmes & Tuomi, 2022; Miao et al., 2021; Nye, 2015; Ogan et al., 2015; Pinkwart, 2016; Roll & Wylie, 2016; Sanusi & Olaleye (2022); Timms, 2016	Cross-cultural differences not being respected and tailored by AI	53%
Mohammed & Watson, 2019	Cultural realism not being incorporated into AI design and development	5%
Coghlan et al., 2021; Costa et al., 2018; Epp & Makos, 2013; Johnson & Lester, 2016; Pham & Wang, 2017; Radford et al., 2021; Rello et al., 2016	Inclusiveness not being taken into account (e.g., Special Educational Needs and Disabilities)	37%
Restrepo et al., 2019	Responsiveness to victims of violence in schools not being considered	5%

Despite earlier challenges, there is growing recognition of the need for culturally inclusive research in AIED (Nye, 2015). Table 6 showcases research underscoring the need for AI systems that account for social and cultural contexts. For example, enculturated agents can adapt their interactions to align with users' cultural backgrounds and preferences, improving communication and learning outcomes (Mohammed & Watson, 2019). Empathic robots and intelligent learning environments (ILEs) can also be customised for students with special educational needs, including autism, hearing or speech impairments, and dyslexia (Epp & Makos, 2013; Johnson & Lester, 2016; Pham & Wang, 2017; Radford et al., 2021; Rello et al., 2016). Chatbots offer additional support in sensitive contexts, such as aiding victims of sexual harassment by providing anonymity, accessibility, and responsiveness (Restrepo et al., 2019).

Table 6*Cultural and Inclusive Features of AIED.*

STUDIES	FOCUS
Mohammed & Watson, 2019	Cultural roots of teaching and learning factored into ILEs
Mohammed & Watson, 2019	Enculturated conversational agents
Epp & Makos, 2013; Johnson & Lester, 2016; Pham & Wang, 2017; Radford et al., 2021; Rello et al., 2016	Empathic robots and ILEs for Special Education
Restrepo et al., 2019	Chatbots for victims of sexual harassment (e.g., European project ACACIA)

RQ3. How does AIED monitor emotions and what ethical challenges may be at stake?

Since AIED is primarily justified as a response to learner holistic needs, this section aims to understand how the affective dimension is addressed alongside the cognitive and performance dimensions. Regarding ethical considerations, only 4% of all papers analysed in

this SLR address the particular ethical concerns that may arise from AIED's handling of emotions. The main concerns include negative feelings related to surveillance, direct correlations between behaviour and emotions, a lack of respect for affective privacy, and the induction of emotions (Table 7, below). The biggest percentage of papers related to this topic concentrate on the emotional outcomes of affective surveillance.

Automated monitoring aimed at improving productivity and well-being has been linked to increased stress and anxiety (Burr et al., 2020) and heightened nervousness during evaluation (Gudiño Paredes et al., 2021). Despite advances in AIED, a contextual understanding of emotions remains lacking, often limited to the "signalling paradigm" of linking emotions to behavioural cues (Dobrosovetsnova et al., 2020). Although positive emotions, motivation, and academic achievement drive AI integration into education, there is excessive optimism about these systems. Studies show that affective agents can enhance learning, yet most educational robots focus solely on expressing positive emotions. Since learners also experience complex and negative emotions, it is crucial to model these ambivalent traits accurately (Dobrosovetsnova et al., 2020).

According to Hudlicka (2016), another potential area of concern pertains to how interactions with an agent can jeopardise the privacy of an individual's emotional experiences. Agents within the AIED context also possess the capability to induce or manipulate emotions, and virtual relationships with these agents may blur the lines between reality and fiction, potentially resulting in psychological challenges. These studies emphasise that the discussions regarding the management of emotions in educational technology play a crucial role within the larger ethical conversation in this domain.

Table 7

Percentage of Papers on AIED Ethical Challenges on Emotions.

STUDIES	ETHICAL CHALLENGES	PERCENTAGE OF PAPERS
Burr et al., 2020; Gudiño Paredes et al., 2021	Human rights: emotional consequences of affective monitoring	40%
Dobrosovetsnova et al., 2020	Utilitarianism and Justice: diversity of emotions not being recognised by AI	20%
Hudlicka, 2016	Human rights: affective privacy (the right to keep one's feelings to oneself)	20%
Hudlicka, 2016	Human rights: emotion induction (the deliberate process of changing someone's feelings)	20%

RQ4. How is capacity building on AIED ethics being covered?

Table 8 presents the 41% of studies in this SLR covering AI ethics education and the targets, content, skills, and delivery methods of such education. Through a systematic policy review, Schiff (2021) found that education for AI includes training AI professionals (e.g., computer scientists), preparing the workforce for AI, and broader public AI literacy. Many authors stress the need for ethics in Engineering education to bridge the gap between technology and society (Antoniou, 2021; Dignum, 2020; Dignum, 2021; Hoeschl, 2017; Park et al., 2021; Williams et al., 2020).

Qualitative data reveals that US Information and Computer Science students often fail to consider the ethical implications of AI design for privacy and well-being without explicit guidance (McDonald and Pan, 2020). Another exploration in the US analysed 31 standalone AI ethics classes and 20 AI/Machine Learning technical courses, revealing both commendable practices and notable omissions, such as accessibility, diversity in the AI workforce, and sustainability (Garrett et al., 2020). In 2020, research conducted across 12 Australian universities indicated a lack of ethics education in Computer Science courses or a focus solely on micro-ethical concepts linked to professionalism and industry standards (Gorur et al., 2020).

The University of North Carolina piloted an AI ethics course for Computer Science students, focusing on explicit ethical agents, and suggested the value of prototyping/hands-on approaches and challenging students to employ diverse ethical approaches and summarise their implications (Green, 2021). Similarly, the University of Oulu in Finland introduced an AI ethics pilot course, covering a range of AI applications, legislative and ethics aspects, and the pros and cons of AI applications (Tuovinen & Rohunen, 2021). Future implementations were recommended to incorporate unexpected ethical issues, including methods like case study analyses and role reversal between defenders and opponents.

Studies in various fields of health, including anatomy, psychiatry, and clinical psychology, highlight the growing importance of ethics and education in the context of AI adoption. While AI holds promise for improving diagnoses and prognoses, it must align with medical epistemological frameworks to address emerging ethical and clinical concerns (Gauld et al., 2021). Research with clinical psychology Master's students at the University of Basel has shown that despite some familiarity with AI/ML tools, they require education to assist patients in making informed choices regarding mental health AI/ML applications, taking into account issues like privacy, equality, and discrimination (Blease et al., 2021).

Other authors suggest that when preparing curricula for AI education, it is crucial to consult with students to understand their needs and their perceived readiness for AI-related topics. To facilitate this, Karaca et al. (2021) have developed the MAIRS-MS, a reliable tool for assessing students' readiness for AI and its applications in medicine.

Schiff's (2021) study found that ethical training for the implementation of AI in the education sector was almost absent. Ng et al. (2023) conducted a SLR that includes thematic and content analysis of 49 publications from 2000 to 2020. The review highlights that AI teaching and learning primarily focused on computer science education at the university level. The findings presented in this review emphasise the significance of educating individuals in AI literacy and AI ethics. So, there are not as many programmes reporting training on ethical issues around AI, and even fewer on best practices for ethical integration of AI resources. Additionally, there are no reported tools to assess pedagogical practices using AI. Even in training projects conducted in schools, there is a noticeable absence of emphasis on the ethical considerations associated with AIED.

This training still addresses shortcomings in countries where AI is already a reality in the classroom, such as China. Its unsystematic nature, non-intentionality and the lousy quality of its supporting materials show there is still a long way to go (Gong et al., 2020). When not absent, it appears that training specifically for the use of AIED is only being provided to teachers, school administrators, and researchers. Loftus and Madden (2020) emphasise the importance of ensuring that students also understand the datafication of their own lives and learning processes, and advocate for placing students at the heart of the construction of AI-powered models, which again draws attention to the importance of sense of agency.

According to Dignum (2021), the digital age is no longer compatible with the separation of STEM from humanities, arts, or social sciences. In fact, arts seem to be a great platform to promote AI education, considering it an expression medium and its role fostering empathy, diversity, and inclusion in the AI pipeline (Srinivasan & Uchino, 2021). Xu (2020) also suggests that all ethical challenges in introducing AIED must be considered from a humanistic educational perspective. Following this principle, some examples of AI ethics training in schools were conceived. Research from Ottenbreit-Leftwich et al. (2022) focuses on introducing AI education to K-12 students and explores the potential starting point for teachers to teach AI and computer science concepts.

It suggests that AI ethics can be a compelling entry point. Teachers showed more confidence in discussing AI ethics with their students, which led to engaged discussions. The

research aims to lay the groundwork for elementary AI education by considering students' ideas, experiences, and ethics as essential components for curriculum design in K-12 education.

Nevertheless, another topic of concern has been exposed. Although 42 studies were dedicated to the topic, only four of them covered the impact assessment phase of these training programmes. A study by Lee et al. (2021) considered the impact of a summer workshop on AI literacy on middle school students, mainly from underrepresented groups in STEM. Certain benefits included a notable improvement in students' grasp of AI and its potential biases, enhanced adaptability to future AI-related employment, a better understanding of the consequences of their actions on others, improved capacity to discuss ethical AI issues with their families, and the ability to leverage their family's resources for self-improvement.

Another study (Kong et al., 2023) highlights the potential of AI literacy education for senior secondary students, emphasising that programming knowledge is not a prerequisite for understanding AI concepts. The results suggest that with sufficient learning time and project-based pedagogy, senior secondary students can develop AI literacy, although there may be challenges in comprehending complex AI ethical principles, which require further guidance and time. The work of Lucic et al. (2022), presents a course at the University of Amsterdam aiming at providing students with a comprehensive understanding of FACT-AI topics and algorithmic harm through lectures, paper discussions, and a reproducibility project. Students engage with the open-source and research communities, creating a public code repository.

The course emphasises the importance of reproducibility and received positive feedback from students who appreciate the critical perspectives gained and insights into AI research. The course successfully motivates students and promotes critical thinking in AI. Finally, using a phenomenographic approach, Yau et al. (2022) studied 28 in-service teachers from 17 secondary schools in Hong Kong after implementing an AI curriculum. They identified six categories of teacher conceptions related to teaching AI, including technology bridging, knowledge delivery, interest stimulation, ethics establishment, capability cultivation, and intellectual development. The study presents a hierarchical outcome space that illustrates the range of surface to deep conceptions held by teachers. It offers insights into cultivating both technical and non-technical teachers' competence in AI education, aiding teacher educators and policymakers in enhancing AI education for K-12 students.

Table 8*Capacity Building on AIED Ethics*

STUDIES	TARGETS	TRAINING CONTENT AND SKILLS	DELIVERY METHODS
Antoniou, 2021; Bogina et al., 2021; Dignum, 2020; Dignum, 2021; Garrett et al. 2020; Goldsmith et al.; 2020; Gorur et al., 2020; Green, 2021; Herzog et al., 2022; Lucic et al. (2022); McDonald & Pan, 2020; Miao et al., 2021; Schiff, 2021; Sjöden, 2020, Tuovinen & Rohunen, 2021; Williams et al., 2020	Future AI professionals, especially Engineering	Distributed nature of AI applications (human-agent interaction); distributed learning entities (autonomous & self-organising AI); inclusion/diversity in design and in AI workforce; governance, socio-technical, legal and economic models; prediction models to understand AI impact on every profession; multiple competencies	Seminars, summer schools, challenge-based learning engineering ethics courses, embedded lectures; hands-on approaches and case-studies
Dignum, 2021; Miao et al., 2021; Reisach, 2021; Xu, 2020	General Public (Including vulnerable and older people)	AI literacy; imagination and humanities; opportunities for on-the-job training; diverse and inclusive AI; training of ethicists among citizens; capacity to adapt to rapid change	AI online media courses
Ali et al., 2019; Bates, 2011; Blease et al., 2021; Chiu et al., 2022; Charisi et al., 2020; Dignum, 2021; Gauld et al., 2021; Javed et al., 2022; Karaca, 2021; Katznelson, 2021; Khurana, M., 2020; Kong et al., 2023; Lazarus et al., 2022; Lee, 2014; Lee et al., 2021; Loftus & Madden, 2020; Miao et al., 2021; Ng et al. (2023); Ottenbreit-Leftwich et al., 2022; Shih et al., 2021; Srinivasan & Uchino, 2021; Zhang et al., 2023	Students (k-12, TVET & Higher Education)	STEM, humanities, arts and social sciences together; medical courses – philosophy of ethics, cognitive neuroscience, computational psychiatry, clinical practice; informed choices about mental health through AI apps; perceived AI readiness (e.g., MAIRS-MS); ethical dilemmas concerning the use of AI in medical practice; agency; meaningful content connecting to prior knowledge; technical concepts and processes, ethical and societal implications, and career futures in the AI era	Mandatory modules, lectures or elective courses; constructivist hands/minds-on experiences w/robots; project-based pedagogy; cognitive levels, course content, and disciplines (BAG model)
Bogina et al., 2021; Chounta et al., 2022; Gary, 2019; Holmes et al., 2021; Luo, 2019; Miao et al., 2021; Mouta et al., 2021; Ottenbreit-Leftwich et al., 2022; Sjöden, 2020; Williams & Breazeal, 2020; Yau et al., 2022	Teachers, School Deans, Researchers (k-12 & Higher Education)	Equitable AI research; personal emotional support; AIED history, philosophy of technology use and adoption; usefulness of AI in education; application of AI-powered EMIS/LMS; computational thinking; human interaction, higher-order thinking, and human values; content that teachers feel confident to teach	Non-traditional educational approaches; participatory methods

RQ5. What principles, regulations and frameworks are there for AIED?

The concerns described in the previous RQs have justified the adoption of guidelines for the use of AI in education. The first attempts to adopt general ethical usage guidelines were unsuccessful, as it was quickly realised that the specifics of the education sector required a specific framework. So, educational actors are now requested to produce workable ethical frameworks to tackle AI potential risks, improving educational institutions and student outcomes (Weber, 2020). However, it is crucial to develop consistent terminology and scope in formal standardisation efforts, especially in the context of Information Technology for Learning, Education, and Training, within and across standardisation bodies (Mason et al., 2020).

Specific discussions on AIED in education policies are taking place in countries like China, India, Italy, Kenya, Malta, Singapore, South Korea, Spain and the United States. Although these nine countries are discussing some version of AI for education, only four or five are discussing it in depth beyond superficial mentions (Schiff, 2021).

As indicated in Table 9, below, which pertains to the 8% of articles discussing this subject, some ethical frameworks show tensions and gaps concerning the ethical advancement and application of learning analytics, while others are not appropriate for use in the field of education. For example, the GDPR is too complex when used in education (Kitto & Knight, 2019) and it can be difficult to apply these guidelines in sensitive cases where the primary concern is the safety of people (Al-Omran et al., 2019).

As of 2019, there were no specific policies or regulations regarding AIED, despite efforts to ensure trustworthy AI (Holmes et al., 2018). In response to this absence, guidelines such as the "Beijing Consensus on Artificial Intelligence and Education" (UNESCO, 2019a) and "The Ethical Framework for AI in Education" (The Institute for Ethical AI in Education, 2021) have been introduced. The UNESCO Consensus on AI in education was developed through the collaboration of multiple stakeholders, including government ministers, international representatives, and experts. It outlines seven key principles for UNESCO's member states: prioritise AI in education policies to achieve SDG 4 goals, support AIED-enhanced pedagogies when benefits outweigh risks, promote AI tools for lifelong and personalised learning, base policies on evidence, provide comprehensive AI training for teachers, cultivate critical skills for the AI era, and encourage equitable, transparent, and ethical use of AIED, with a focus on gender equality (UNESCO, 2019a). "The Ethical Framework for AI in Education" (The Institute for Ethical AI in Education, 2021) appears to be the most sophisticated and up-to-date

tool to closely monitor AI technologies used in the various stages of AI adoption in education, from pre-procurement to implementation and impact evaluation. Launched in 2018 by the University of Buckingham, this framework became publicly available in 2021. It resulted from a two-and-a-half-year effort, which included collecting interviews from policymakers, academics, philosophers, ethicists, industry experts, and educators to establish a consensus regarding the integration of AI into the education sector.

The framework addresses ethical design, privacy, equity, transparency, and accountability concerns, considering the sector's specific needs. For example, in relation to privacy, it states that even if an institution is required to continuously assess students, it must also establish safe spaces where learners are not assessed. When it comes to autonomy and agency, the framework draws attention to the actions that should be taken when the AI system predicts an unfavourable outcome. It also underscores the importance of ensuring that AI systems are designed to benefit learning without leading to addictive behaviours. This framework is the first step in addressing the ethical challenges of AIED and ensuring it is educationally sound from the ground up.

Finally, Miller and Tuomi (2022) emphasise the importance of sense-making when envisioning the future. They highlight the significance of employing various forms of anticipation to perceive the world from different perspectives. This approach is rooted in a shift in the field of futures studies towards an ontological perspective, which allows us to reconceptualise the future as a point of origin. It positions anticipation and anticipatory processes about the future as integral aspects of the present. This challenge could be integrated into the ongoing enhancement of frameworks related to AIED and educational ethics, fostering broader, more creative, and innovative processes that aren't limited by predefined objectives.

Table 9*Ethical AI and AIED Guidelines and Frameworks*

STUDIES	GUIDELINES/ FRAMEWORKS	PURPOSE
Al-Omran et al., 2019	Software Engineering Code of Ethics and Professional Practice; IEEE Code of conduct; ACM Code of Ethics	Preparing stakeholders on tech ethical considerations
Kitto & Knight, 2019	General Data Protection Regulation (GDPR)	Safeguarding individuals' rights when producing personal data in daily activities
UNESCO, 2019b	UNESCO's Rights, Openness, Access and Multi-stakeholder Governance)	Guaranteeing human rights and emerging ethical issues in the application of AIED
European Commission, 2019	Ethics guidelines for trustworthy AI - High-Level Expert Group on AI	Assuring human agency, accountability and transparency; technical safety; diversity, non-discrimination and fairness; societal and environmental well-being
Dignum, 2021	ART principles for responsible AI	Anticipating AI societal impact; accountability, responsibility, and transparency
UNESCO, 2019a	Beijing Consensus on Artificial Intelligence and Education	Achieving SDG 4; beneficial AIED-empowered pedagogies; ubiquitous lifelong and personalised learning; evidence-based policies; auditable uses of AIED
The Institute for Ethical AI in Education, 2021	The Ethical Framework for AI in Education	Ensuring AIED provides the greatest possible advantage to the learner

2.3. Discussion

The following sections delve into various aspects concerning the implications of the findings from this SLR. They encompass the integration of a theoretical model for analysing pedagogical implications based on the collected data, the identification of research gaps, recommendations for teacher training in AIED ethics, and a critical examination of potential limitations in this research.

2.3.1. SLR Implications for Rethinking Education Ethics: Using a Theoretical Model

The introduction of new technology in the realm of education provides an opportunity not only to reconsider the specific ethical consequences it introduces but also to contemplate the broader ethics associated with education itself. The ik-model (Mouta et al., 2015) represents an extended framework building upon the TPACK model, encompassing the human element within educational technology integration. This framework comprises four pedagogical dimensions: "technologies and resources", "content knowledge", "learning processes and strategies", and "educational actors and their relationships". Subsequently, this theoretical framework is applied to assess the potential consequences of the aforementioned results.

2.3.1.1. Technologies and Resources Dimension

RQ1 and RQ3 emphasise the critical importance of creating inclusive AI systems in education, underscoring the significance of diversity in decision-making. Yet, ethical discussions on AI in education are primarily Western-centric and STEM-dominated, as indicated by MQ2, MQ3, and MQ4, potentially resulting in biases and hindering inclusive AI system development. RQ5 stress the importance of collaboration in evaluating ethical frameworks for AIED, highlighting the need to balance student interests, meaningful innovation, and overcoming resistance to change in education. Furthermore, it points out the challenges of applying these frameworks across diverse educational contexts and levels. In light of these findings, three key outcomes emerge: (1) funding for research and educational programmes including underrepresented areas (e.g., arts, health, humanities), populations and groups, exemplified by AIED DEIA mentoring fellowships (“Call for Fellow Nominations”, 2023); (2) systematic collaboration among AI researchers, developers, and practitioners, drawing from school fieldwork to create technically robust and culturally sensitive AI systems.

This would be an opportunity to build on teachers' own ethical development, by creating opportunities for them to exercise their skills as judging actors (Arendt, 1958). (3) It is recommended to develop an assessment tool that can adapt to different contexts and monitor the positive impact of these technologies, involving a wide range of stakeholders. These frameworks should also ensure seamless integration with educational standards as students progress through their educational journey and into their jobs.

2.3.1.2. Content Knowledge Dimension

The results from RQ4 indicate that while there are training programmes addressing ethical challenges in AIED for teachers, school administrators, and researchers, students usually receive limited exposure to the ethical use of AI. Their education primarily focuses on deontological and on-the-job ethics. This approach overlooks the valuable perspective that students can bring to the development of AIED environments, considering that they are the primary beneficiaries of these technologies.

To bridge this gap, curriculum infusion is proposed as a suitable strategy. It allows for the inclusion of emerging topics in a meaningful way, while accommodating the busy academic calendar. By incorporating AI concepts across various subjects and applying them to tasks utilising AIED resources, educators can facilitate discussions about AI's functions, impacts, and ethics from diverse subject perspectives. Ultimately, this approach enhances students'

acquisition of content knowledge, application to practical tasks, and cultivates a comprehensive understanding of AI systems and their effects on individuals and groups.

2.3.1.3. Learning Processes and Strategies Dimension

The results of RQ1, RQ2, RQ3, and RQ5 underscore the importance of preserving students' diversity and sense of agency in AI-based education. For personalisation in education to be effective, it needs to strike a balance between tailoring content to individual needs and ensuring that students have the opportunity to explore diverse perspectives and develop a holistic self-understanding. Exploring one's own interests, competencies, and values – three critical dimensions of identity – is essential for informed decision-making throughout life and fostering a strong attachment to the learning experience.

There is a need for epistemological reflection as an essential aspect of contemplating the significance of pedagogical innovations in the contemporary world (Trindade and Cosme, 2010). This goal can bring teachers and students together in project-based learning, encouraging them to explore diverse sources of knowledge and engage with them using various processes and literacies. This is even more important given the access to language models like ChatGPT. The challenges it entails present an opportunity to place questioning and critical thinking at the core of education. Approaches like the flipped classroom can help students engage meaningfully with AI technologies, fostering thoughtful exploration. Independent learning methods, followed by teacher-led discussions focused on problem-solving, offer a dual benefit: they allow students to deepen their understanding independently and encourage dialogic developmental processes when they share and explore their learning outcomes with their peers. Developing these skills empowers students as citizens, while assisting societies in navigating the evolving technology landscape.

Adapting evaluation criteria to accommodate diverse student expressions is critical when AI offers customised learning experiences. This process entails analysing AI-generated data and presenting it to students, teachers, and families for input and a systemic understanding. Effective use of AIED assessment tools can encourage self-assessment as a valuable learning tool, allowing students to critically evaluate technology, promoting self-reflection, self-regulation, and citizenship. In the era of dataism, educators face the challenge of exploring dataism's onto-epistemic grammar with their students, including its anthropocentric perspective, the drive for ontological security, and the thirst for absolute knowability (Andreotti et al., 2015; Lados et al., 2022; and Stein et al., 2017).

Only 4% of the papers in this field address the ethical aspects of AIED's emotional management (cf. RQ3), indicating a predominant focus on performance as the primary rationale for using AIED. However, this emphasis on performance neglects the fact that it results from a complex interplay of both cognitive and emotional factors. Emotional well-being and social skills are key for overall achievement. Currently, these AI technologies often align with an educational paradigm that prioritises performance and global rankings.

Adhering to established and conventional assumptions about knowledge, teacher and learner roles, educational goals, and learning methods when incorporating AI systems may exacerbate existing issues, making distances wider, instead of addressing broader educational and developmental needs. Such an approach also hinders the potential of these technologies to foster a decentralised and self-directed approach to education, as advocated by thinkers like Illich (1971). By solely concentrating on the short-term advantages, educators and policymakers might overlook or underestimate the broader impacts and ethical considerations associated with AI in education, hindering its beneficial potential.

2.3.1.4. Educational Actors and their Relationships Dimension

RQ1 and RQ2 shed light on the importance of recognising the significance of shared values, both explicit and implicit educational agreements, and the involvement of diverse stakeholders in decision-making concerning the use of AIED. MQ3 has also suggested that achieving geographical diversity in research on AIED ethics remains a persistent challenge. This recognition should be coupled with an awareness of disparities in access and variations in pedagogical quality. According to Christakopoulou et al. (2001), a school is a multifaceted entity encompassing aspects of a social, economic, and political community.

Also, it serves as a personal environment where attachments are formed, and memories are created. The introduction of AI-based education challenges all these dimensions, prompting a critical examination of the roles, relationships, and power dynamics of various stakeholders involved in the educational ecosystem. It calls for a reevaluation of how education is not only about imparting knowledge but also about nurturing a sense of belonging, empowerment, and active participation within a rapidly evolving educational landscape.

2.3.2. Uncovering Research Blind Spots on AIED Ethics

To tackle the challenge of rethinking education ethics in the age of AIED, it is crucial to carefully consider both the insights gained from this SLR and the aspects that remain concealed or overlooked. One of those missing aspects has to do with the lack of incorporation of AIED

ethics in philosophical or psychological paradigms of moral or ethical development. This vagueness in discussions may limit the definition of criteria to thoroughly evaluate the impacts of AIED use. Moreover, there is a notable oversight when it comes to discussions about the broader concept of transhumanism introduced by AI systems. There was only one paper that raised awareness of the question of whether there should be limits on using technology to extend or enhance cognitive abilities, as well as the dynamics of relationships in hybrid environments. How are issues of equity addressed when only a privileged few will have access to competitive AI features, granting them a significant advantage?

Discussions often lack critical perspectives on the ethical boundaries of the concept of intelligence as defined by AIED research and design. Overreliance on behaviourist and cognitive approaches may sideline aesthetics, emotions, morality, and social development, potentially reinforcing reductionist views of human intelligence. Prioritising optimisation may undermine the value of reflective and contemplative thinking, which is crucial to develop strategies for solving complex problems. Another perspective that can be gained from this study is that by adopting a participatory approach involving educators and students in research and inquiry-based learning, the full potential of AIED can be harnessed while encouraging a reflexive and critical attitude that supports comprehensive ethical growth.

2.3.3. Recommendations for Educators on AIED Ethics

It becomes clear that the ethical implementation of AIED requires a comprehensive approach. The results emphasise the importance of participatory methods and dialogic ethics within this context. The following recommendations on AIED ethics are based on the insights derived from this study and will be followed on the continuation of this project research: (1) engaging educators through focus groups – respecting the insights from this research, it is recommended arranging focus groups sessions with educators to provide them with a platform to engage in meaningful discussions about the ethical use of AIED in educational institutions; (2) developing a pilot training programme – data generated from these focus group sessions will form the basis for creating a pilot training programme aimed at promoting ethical considerations and practices in the integration of AIED within educational settings; (3) exploring content and delivery methods accordingly – during the programme's development, it is critical to align with research findings, adhering to a socio-constructivist framework that emphasises active participation and engagement when designing content, delivery methods, duration, and activities; (4) training in comprehensive topics – develop the training curriculum to cover a wide range of topics, including the ones that were overlooked in previous work

(AIED technologies, potentialities, and challenges; ethical considerations specific to education; interdisciplinary perspectives on AIED; stakeholders in AIED, power dynamics, interests, and needs; effects of AIED ethical challenges on student agency, self-determination, and emotional well-being; AIED's role in pedagogy and innovative learning processes; learning analytics, data collection, analysis, and interpretation in the context of AIED; evaluation of AIED effectiveness using informed criteria; promotion of communities of practice in AIED for knowledge sharing and addressing challenges); (5) role-playing with students – incorporate challenging and hands-on engaging tasks that encourage reflection on how AIED-specific applications could impact students' lives, which can then be further discussed in training; (6) evaluating through self and peer assessment – align with the training's defined criteria, facilitating an evaluation of the created resources' responsiveness and effectiveness; (7) promoting inclusive participation – invite a diverse group of teachers as trainees, including individuals from typically underrepresented countries and representing various fields; (8) assessing attitudes towards AIED ethics – include quantitative data collection moments at the beginning and end of the training programme to facilitate the measurement of its impact and inform future improvements; (9) gathering teacher perspectives – involve teachers in a final semi-structured interview to gain insights into their perspectives regarding key criteria that should be prioritised in a continuing professional development (CPD) programme on AIED ethics, with a specific focus on their training experiences.

2.3.4. Potential Limitations of This Research

Despite the inclusion of over 100 papers in this study, there are certain limitations to consider. Since our research covers papers published until the end of 2022, studies currently being published are not taken into account. The most recent research can provide updated insights as the field is growing both conceptually and in school practice. Another limitation of this SLR can be some articles left behind. In fact, the general search term “Artificial Intelligence” was chosen because it would be difficult to cover all its technologies, such as “robots”, “educational chatbots”, “machine learning”, “intelligent tutoring systems”, “exploratory learning environments”, “teachable agents”, “dual-teacher model”, “speech/image recognition”, “autonomous agents”, “affect detection”, and so on.

The idea was that papers covering a specific AIED technology would mention the word “Artificial Intelligence” at least once. Nevertheless, this may have excluded papers that used narrower terms without referencing AI directly. An additional 24 papers were included later as first or second references, but others with similar relevance may have been overlooked. To

address this, multiple screening rounds and comprehensive search strategies were applied, though some omissions may still have occurred.

2.4. Conclusion

Research interest in the ethics of designing, developing, and implementing AIED has been steadily increasing since 2018. This interest has experienced a surge after 2020, as a result of the COVID-19 pandemic and the rise of AIED technologies, leading to a significant increase in AIED ethics research. The majority of papers analysed in this SLR presented concerns about dimensions such as fairness, inclusion, autonomy, and agency. The publications analysed also showed efforts to implement AIED ethical programmes in K-12 and higher education. The main conclusions drawn from this SLR highlight the significance of participatory processes in researching and implementing AIED.

To achieve this, it is essential to engage a range of educational stakeholders, including students, teachers, school administrators, parents, and researchers. Although regulatory ethical frameworks for AIED have been introduced, they have come quite late and do not account for the specific needs of learners and pedagogy. Generic AI ethics principles do not adequately address the part of agency responsible for learning, which no technology with promises of better personalisation or guidance should struggle against. Therefore, ethical frameworks for the design and use of AIED should be developed through participatory processes that recognise the specific needs and tasks of its main actors (students and teachers), and respect each learning community's heritage, as well as the will and capacity for innovation.

This SLR uncovered the importance of providing teachers with support to effectively utilise AIED technologies while preserving students' sense of agency and promoting lifelong learning potential. Additionally, rethinking and aligning evaluation parameters is crucial to ensure that ethical concerns are taken into account and that data is incorporated from multiple feedback sources. The SLR also revealed a failure to incorporate AIED ethics into a philosophical or psychological paradigm of moral or ethical development. The lack of a defined conceptual corpus that fits the educational landscape makes it difficult to understand the extent to which AIED meaningfully addresses ethical aspects of individual, community, and organisational development.

Using AI technologies can facilitate self-reflection, including reflection on the actions of other agents (persons and machines), and balance personal standards and value systems while considering interpersonal connections and mutual obligations.

This systematic literature review sheds light on the key topics that should be included in teacher training on education ethics when using AIED. It also suggests that while AIED research has been addressing ethical considerations, there is still significant room for growth in terms of analysing its unique heuristics within the context of education ethics.

CHAPTER THREE

Empirical Research I: Futures Studies

3.1. Introduction

This section presents the results of a participatory futures research study exploring the ethical challenges posed by AIED. This study employed the Delphi Method to engage experts in a structured consultation, leading to the development of eight hypothetical future scenarios. These scenarios critically examine how AI integration in education may impact pedagogical goals, assessment practices, equity, student agency, privacy, accountability, informed participation, and ethical design. The study highlights the need for AI in education to be guided by agentic and meaningful pedagogical, social, and scientific evidence, ensuring that its implementation aligns with ethical principles and supports student development. The findings inform further consultation with educators to develop teacher training programmes that foster epistemic and collective agency when using AI in education. By mapping both opportunities and risks, this research contributes to the broader debate on responsible and agency-driven AI adoption in educational settings.

3.1.1. Arguments for a Participatory Approach to the Ethics of AIED Use

Recognising the significance of stakeholder involvement in the design of large societal projects (Bahadorestani et al., 2020), this study strives to give educational actors a voice in establishing a secure and valuable environment when using AIED technologies. The novelty of this research topic and the limited number of project impact assessments make it particularly relevant to involve educational actors in discussions. Zuboff (2019) highlights the exceptional nature of these technological advancements that cannot be captured by current frameworks, reinforcing the importance of using participatory methods and futures approaches. Stiegler (1998) characterises technology as the science that accompanies the creation of technical objects, and he believes that its greatest strength is precisely its unpredictability. Unlike what the “robot myth” suggests, technology’s dynamics are not controlled by automation, but rather by the object itself that is prone to unpredictability, making it difficult to predict its future development.

By using “technology maieutic”, however, experts can contribute to a constructive evaluation of the future of these systems and their applications. It is critical, therefore, to incorporate a variety of expertise, including local and traditional knowledge and practices, in policy design, implementation, and evaluation of AIED. This approach also increases transparency, accountability, and legitimacy of decision-making (ICAT, 2020). Regarding technologies, few research articles have examined how stakeholder engagement is considered

by research teams to evaluate key characteristics of the technologies to be developed (Nygaard et al., 2021). So, this is the proposal of the research presented throughout this paper: to merge ethics with participatory, deliberative and stakeholder approaches, based on the assumption that the public can make an ethically informed assessment of a new technology and that the moral insights of various individuals involved in the creation of these systems can enhance ethical evaluations (Brey, 2017). In fact, from a pragmatic point of view, a participatory study can lead to a more diverse and comprehensive analysis; and from an ethical perspective, this study recognises the human right to be part of public decision-making processes that affect people's lives.

This idea aligns with the democratic and emancipatory ideals of the modernist Enlightenment (Santos, 2012). So, the implementation of the ethics of discussion is justified because it is deemed more suitable than the Kantian dialectic in identifying practical solutions to real-world ethical dilemmas. The ethics of discussion is based on a dialogical concept of reason inspired by the “linguistic turn” of analytical philosophy. The tasks of deontological ethics must be carried out by communicative reason, embodied in an open discussion with the plurality of members of an ideal community of argumentation. This would allow for a closer connection between ethical argumentation, thought, and practical action, insofar as the social agents themselves, as ethical subjects, participate in argumentative activity and introduce various ethical content into the discussion in their materiality (Santos, 2012).

This study, therefore, intends to answer the following research questions: How do diverse stakeholders' ethical perspectives around the integration of AIED come together to inform and shape possible future scenarios? How might these imagined scenarios be crafted into a continuing professional development toolkit to support educators in engaging with the ethical dimensions of AIED? Slaughter (2020) suggests that concepts regarding the future should be incorporated into the curriculum, teacher training, and educational systems. This study will, thus, further enable conversations with educators concerning the primary themes of a training curriculum, which will aid in the facilitation of the toolkit's use. The aim of this training programme is to guarantee that AIED is used in a meaningful and ethical manner that prioritises educational and pedagogical objectives. This toolkit can also drive both risk and impact assessments that support educational policy design and implementation.

3.1.2. Futures Studies

As AIED theory is still in its early years of growth, this research importantly integrates futures studies methodologies to anticipate, monitor, and address the ethical challenges that these technologies may pose.

3.1.3. Historical Overview of Futures Studies: Theory and Method

After the Second World War, futures studies began to gain traction in the USA to assist military systems in the field of technological forecasting (Tolon, 2011). The US Department of Defense, through its research think tank – the RAND (Research and Development) Corporation – started publishing on these matters in the 1960s. By the end of that decade, both the World Future Society and the World Futures Studies Federation were created, and the 1st World Futures Conference was held in Oslo. In 1973, this Federation, founded in Paris, embraced pluralism through the diversity of its members and research areas, arguing that the singular form of the word “future” is conceptually and politically limiting. For over 60 years, futures studies have evolved from a method focused on making predictions to one that questions the possible, probable, or preferred transformations and impacts of an existing object in its transition from the present to the future(s) (Hines, 2019; Son, 2015; Tolon, 2011). Its research is usually linked to objects, subjects, and processes that involve a particular uncertainty regarding how they can be transformed and how they, in turn, can transform other layers and fields. It also relies on a predominantly Western linear concept of time that spread with modernity, excluding cyclical and spiral perspectives of time found in hunter-gatherer societies, particularly before the organisation of agriculture (Elias, 2007), or in cultures that believe the primordial moment can be renewed repeatedly (Eliade, 2016), based on the recurrence of cosmic and earthly processes (Fuchs, 2018).

According to Inayatullah (2013), futures studies refer to the systematic study of possible, probable, and preferable futures, taking into account both worldviews, myths. Bjerstedt (1982) identified a fourth orientation – the preparedness to act – which comprises self-reliance and solidarity, often referred to as prospective futures capacity. Its focus is on mapping alternative futures as a form of agency toward them.

This critical futures approach represents a post-structural turn in the field (Inayatullah, 2013), where the world is not conceived separately from its representation but rather in a loop of inputs, from the inner to the outer world. Within this theoretical landscape, four key approaches emerge (Inayatullah, 1990):

- A predictive approach, grounded in empirical studies.

- An interpretative approach, which seeks to make sense of different images of the future.
- A critical approach, in the Foucauldian tradition, which questions who benefits from certain methodological choices in futures studies.
- A participatory approach, based on stakeholder involvement, aimed at fostering more democratic assumptions and actions regarding the future.

By the end of the 20th century, Mannermaa (1992) proposed that futures studies could encompass three alternative paradigms: (1) a descriptive (mainly quantitative) approach, predicting the future based on past data; (2) a scenario-based approach, supporting decision-making through diverse perspectives on possible futures; and (3) an evolutionary paradigm, applying evolutionary laws and complexity research to understand and describe potential futures.

Gidley (2017), drawing on Inayatullah (1990), Masini (1993), Bell (2004), and Slaughter (1999, 2003), outlined five key traditions in futures studies. Despite their chronological emergence, all the following approaches remain relevant and are applied according to contextual needs: (i) empirical-positivist, which emerged with the World Future Society (1960s, USA), emphasising trend analysis and prediction; (ii) critical-normative, which gained prominence with the World Futures Studies Federation (1973), offering an alternative to the empirical approach; (iii) cultural-interpretive, which focuses on non-Western, civilisational, and planetary futures; (iv) empowerment-activist, which include action-research approaches that grew in Europe during the 1990s; (v) integral/transdisciplinary, which incorporates multiple perspectives and planetary studies (Gangadean, 2006; Montuori, 1999).

Strategies for futures studies include mapping, anticipating, timing, deepening, creating alternatives, and transforming. Traditional forecasting relies on quantitative methods such as (1) time series analysis, which uses statistical processes for predictions, and (2) explanatory analysis, often based on regression and correlation, both projecting linearly from past to future. Additional quantitative methods include (3) trend impact analysis, which adjusts time series based on expert insights on unprecedented events (Firminger, 2003), and (4) cross-impact analysis, which evaluates the conditional probability of events based on their interdependencies (Gordon, 2003).

Four key qualitative techniques include the Delphi method, futures wheel, simulation and games, and agent-based modelling. The Delphi method, due to its significance to this research, will be further explored. The futures wheel is a participatory brainstorming tool that maps cascading impacts of an event (Bengston, 2016). Simulation and games examine how variables

interact in controlled models (Rausch, 2003). Agent-based modelling uses computational agents that act upon their environment based on past experiences, simulating complex, interdependent systems (Salgado & Gilbert, 2013).

Some techniques combine qualitative and quantitative elements. The scenario method constructs narratives about potential futures, identifying opportunities and challenges. Scenarios encourage reflection, stimulate further questions, and integrate quantitative approaches while considering plausible cause-and-effect relationships. Unlike data-driven projections, their goal is not to predict but to present alternative perspectives for managing uncertainties (Dean, 2019; Martelli, 2014; Schwenker & Wulf, 2013).

3.1.4. Historical Overview of Futures Studies Application to Education

Three main approaches to futures studies in education can be found in literature. The first concerns youth participation, exploring young people's own views on what the future may or should look like. In the 1970s, this approach revealed a dissonance between the futures that young people envisioned for themselves (often conservative and conventional) and those they imagined for the world in general, which tended to be more pessimistic (Johnson, 1987; Toffler, 1974) or only moderately optimistic (Guilló, 2013). This section will not explore the reasons behind this discrepancy in depth, but researchers suggest that personal aspirations are shaped by modern influences, while perceptions of the world are influenced by postmodern factors, particularly the media (Rubin, 1996). Other factors that may contribute to this difference include cultural background (Guilló, 2013), age – with older students generally being less optimistic (Hicks & Holden, 1995) – and artistic education, which appears to enhance the ability to imagine preferred futures (Gidley, 1998). Another approach to futures studies in education focuses on research and practice related to specific concepts, subjects, processes, and resources for K-12 school contexts. Slaughter (1998, 2019) highlighted the importance of embedding futures concepts throughout curricula, teacher professional development, school systems, and strategic planning.

Various methodologies support this kind of reasoning in schools, including timelines, futures wheels, visioning, scenario-building through pictures, change cycles, brainstorming, drawing, jokes, cartoons, science fiction, social inventions, and back-casting – a process that links an envisioned future to the present and to specific actions, policies, or programmes (Gidley et al., 2004). While these methodologies are valuable for exploring futures, “futures fluency” is considered critical for engaging with them. This concept refers to the ability to critically, constructively, and imaginatively speculate about different future possibilities, a skill

that is becoming more complex as systems thinking evolves within contexts of chaos, decentralisation, massive distribution, and inclusiveness (Schultz, 2015). Recently, the concept of Futures Literacy has been proposed as a developing competence for anticipation, distinct from planning-oriented foresight. It focuses on learning how to use futures, with the implicit claim that anticipatory competence is the source of agency (Tuomi, 2024).

Finally, some studies focus on foresightful, transformative educational models and strategies. Slaughter (2002), however, reminds us that the application of futures studies in education still faces many challenges, particularly the resistance of predominantly industrial-era school systems to change. Nevertheless, the need for more integral, spiritual, and artistic educational models was recognised over a century ago (Aurobindo, 1933; Steiner, 1982, 1996) and has been revitalised in sustainable (Redman et al., 2018; Warner & Elser, 2015), active citizenship (Akin et al., 2017; Simonova et al., 2019), and artistic (Elyağutu, 2016) approaches to education.

The first wave of futures studies in education primarily explored young people's views on probable futures and spread across the US. Over time, it evolved to incorporate possible and preferred futures, shifting towards more imaginative, creative, ideological, and critical perspectives, aligning with the European critical futures tradition (Gidley et al., 2004). According to Slaughter (2003), the third wave of futures studies in education, prompted by Inayatullah's work, expanded into a multicultural and internationally diffuse collection of perspectives. Gidley et al. (2004) identified a fourth wave, emphasising empowerment-oriented educational futures, and proposed a fifth iteration, where the Integral Futures model is placed, following the works of some authors in the field of psychology (e.g., Jean Gebser, Kieran Egan, Robert Kegan).

3.1.5. The Delphi Method: Theory, Evolution, and Educational Use

Mapping alternative futures is also a way of agency looking into the future. So, the option of creating techno-ethical scenarios is justified by the fact that they have proven to be appropriate to study moral change. They allow for an ethical analysis based on the expected future moral values of the stakeholders involved (Brey, 2017). For this research, a participatory scenario planning approach was chosen (van Notten, 2006), and most of the content is based on the Delphi consultation (Bond et al., 2021; Dinges et al., 2020; Nuwan et al., 2021).

The Delphi method originated in the 1930s at Ohio State University and was initially used to develop criteria for faculty evaluation, a tradition that has continued in higher education. It is still considered important today for educators in designing curricula and career-relevant

learning experiences (Green, 2014). In the 1960s, the RAND memorandum (Dalkey & Helmer, 1963) described this research as an experiment aimed at reaching the most reliable consensus of expert opinion. This was achieved through a series of in-depth questionnaires interspersed with controlled feedback. The technique was developed to address areas where theoretical or quantitative approaches were insufficient. Initially, it was used in forecasting, later expanding into policymaking and patent participation identification in the early 1980s (Hilbert et al., 2009). Over time, it became a widely adopted research technique in fields such as medicine, nursing, and health services (Gibson, 1998), as well as in education (particularly in higher education), where studies have highlighted its applications in cost-effectiveness analysis, cost-benefit analysis, curriculum and campus planning, and university-wide educational goals (Green, 2014).

The method involves a structured sequence of questionnaires that collect data from experts through successive “rounds”. In each round, participants receive anonymised feedback from the previous responses, refining their views progressively. The goal is to reach a consensus within the group (Powell, 2003), at which point the process concludes. A key feature of this method is the anonymity of participants, which helps to prevent domination by certain individuals, reduces bias, and encourages open expression and attitude shifts throughout the process. According to the RAND memorandum (Dalkey & Helmer, 1963), this approach seeks to prevent the tendency to dismiss novel ideas, to rigidly defend one’s initial stance, or to be unduly persuaded by others’ opinions. Instead, it fosters the gradual formation of a well-reflected expert consensus.

Additional characteristics of the Delphi method include regular feedback, iterative rounds, and a facilitative role assigned to the researcher. The researcher is responsible for setting the task, structuring the flow of information, and integrating the Delphi consensus with institutional research, up-to-date literature, and contextual factors, such as the campus environment in higher education studies (Green, 2014). There are, however, no universal requirements for implementing this technique (Hasson et al., 2000), leading to variations in its application. For instance, the “argument Delphi” method is based on experts presenting distinct ideas, generating new arguments, and critically evaluating those of others, ultimately arriving at a decision through this dialectical process (Seker, 2015). This consensus-driven approach is particularly valuable for generating ideas on unexplored or contentious topics. It enables the participation of subject-matter specialists from diverse geographical locations, broadening perspectives on a given issue based on their respective areas of expertise (Green, 2014).

3.2. Methodology

The main objective of the current study is to investigate the ethical challenges related to the integration of AIED from the perspective of multiple educational stakeholders using the Delphi technique. As this theoretical field is still in its early stages of development, the Delphi method was deemed an appropriate choice to incorporate a broader range of perspectives. As Cabero Almenara and Infante Moro (2014) state, the Delphi method enables researchers to identify key topics for investigation, formulate research questions, establish a theoretical foundation, select relevant variables, determine causal relationships between factors, define and validate constructs, develop analytical or data-gathering instruments, and create a shared language for discussion and knowledge management.

In particular, the current study aims to develop an informed toolkit that can be utilised in continuing professional development for educators in different regions. The expertise, perspectives, and viewpoints of experts were sought to gather insights on various aspects, such as AIED technologies, applications, purposes, contexts, educational actors, subjective experience, impact on subjectification, socialisation, and qualification, as well as usage drivers, ethical concerns, and existing regulations. The goal is to provide educators with the necessary sensitivity, knowledge and resources, empowering them to participate in constructive discussions and make informed and meaningful decisions concerning the ethical integration of AIED in various educational environments.

3.2.1. Expert Group Constitution

The research coordination group consisted of three researchers and was intentionally formed through a criterion sampling method (Patton, 1990). This implied the selection of participants based on predefined criteria that focused on their substantive knowledge of the problem under study (Ogbeifun et al., 2016). In addition, there is controversy over the use of the term “expert” and how to appropriately identify a professional as such (Hasson et al., 2000). Therefore, in the context of this research and given the novelty of the debate on the ethics of AIED use (in research, policy and training), there was an urgent need to define consensual criteria for what an “expert” can be. Based on other Delphi studies (Arteaga-Martínez et al., 2021; García-Peñalvo et al., 2021), proven knowledge, extensive professional experience in the field of study, and sensitivity to scientific research (grounded on previous collaborations) were selected as preferred criteria. Additionally, the participants’ professional diversity was appreciated since it brings varying perspectives from individuals in distinct fields (Renzi & Freitas, 2015). To manage the impact of non-acceptances, a larger number of individuals were

initially invited, surpassing the preferred group size (Ogbeifun et al., 2016). So, during the initial stage, 30 AIED experts were chosen from different regions, and eventually, 18 of them consented to take part. Out of the selected participants, five individuals did not respond to the invitation, and seven faced difficulties in fully engaging with the process due to their professional commitments and ultimately declined the invitation.

The eligibility criteria selected were: 1. work experience in the field of technology for education (EdTech) as (a) government advocate, opinion-maker, or supplier; (b) researcher; (c) specialist in implementation and evaluation of technologies in education; and (d) specialist in EdTech development; 2. professional experience (PE) in the field of over 10 years; 3. previous collaboration with academic research (PCR); 4. perception of the self as a specialist in education with technologies or EdTech. If any of the four criteria were met, the potential participant was deemed ineligible. For the first criterion, participants were considered eligible if they had profiles (a), (b), (c), or (d), but it was not expected for all four profiles to be present in the same participant. Criteria three and four were implemented to ensure highly qualified panel members with a high level of expertise.

With this purpose, the coefficient of expert confidence ($K = \frac{1}{2} (K_c + K_a)$) was added (Cabero Almenara & Barroso Osuna, 2013; Sanromà-Giménez et al., 2021). K_c is understood as the self-assessment knowledge coefficient on the topic (on a scale of 0 to 10) multiplied by 0.1. The argumentation coefficient (K_a) was determined based on the participant's involvement in previous research (criterion #3) and their years of professional experience: 1 for more than 30 years; 0.8 for 20 to 30 years; 0.5 for 10 to 20 years. The coefficient of expert competence has been 0.73. To reduce observer bias, the data recorded in each round was analysed by multiple observers from the lead research team: the three researchers attempted to ensure interrater reliability of the collected data.

Responses to the first round came from 18 participants (100% participation rate) whose sociodemographic and occupational profiles are presented in Tables 10 and 11, below. The mean age of the participants was 44.5 years (SD 7.42), with 16.6% identified as women and 83.3% as men. They were employed on different continents and included countries as diverse as Portugal (PT), Timor, the United Arab Emirates (UAE), and United States of America (USA). All had professional experience in the field of education and held either a Master's or Ph.D. degree. The experts' occupational fields were corporate and business (61%), academic (22%), government (11%), nonprofit and community-based (6%).

Table 10*Experts' Sociodemographic and Professional Profiles*

DELPHI EXPERTS' PROFILES									
Professional Activity	No. of Experts	Age Range			Gender		Working Country		
1. Advocates, opinion-makers, buyers and suppliers of EdTech	5	31-40	41-50	51-60	M	W	PT	USA	
		0	2	3	5	0	3	2	
2. Researchers EdTech	5	31-40	41-50	51-60	M	W	PT	Timor	USA
		0	2	3	5	0	3	1	1
3. EdTech Specialists (implement & assess)	3	31-40	41-50	51-60	M	W	PT		
		2	1	0	1	2	3		
4. Specialists in AIED software development	5	31-40	41-50	51-60	M	W	PT	USA	
		4	1	0	4	1	5	0	

Table 11*Experts' Professional Profiles*

EXP No.	GENERIC PROFESSIONAL CATEGORY	QUALIFICATION	PROFESSIONAL CATEGORY	OCCUPATIONAL FIELD	PCR	PE	K
1	EdTech Advocate	Master in Management	EdTech broker	Corporate and Business	Yes	20 – 30	0,7
2	EdTech Specialist (implementation & evaluation)	Doctor in Education	Government EdTech advisor/ School Dean	Government	Yes	30 – 40	1
3	Researcher on EdTech	Doctor in Psychology	Researcher/ Education Lecturer	Academic	Yes	20 – 30	0,85
4	Specialist in EdTech development	Doctor in Psychology	CEO in Data Business Analytics	Corporate and Business	Yes	40 – 50	0,85
5	EdTech Specialist (implementation & evaluation)	Master in Computer Science/ Engineering	Joint venture analyst in EdTech	Corporate and Business	Yes	20 – 30	0,8
6	EdTech Specialist (implementation & evaluation)	Doctor in Sociology	Education programmes consultant and CPD trainer	Non-profit and community-based	Yes	10 – 20	0,65
7	EdTech Advocate	Master in Education	Government EdTech advisor	Government	Yes	20 – 30	0,8
8	Researcher on EdTech	Doctor in Anthropology	Researcher/ Anthropology Lecturer	Academic	Yes	20 – 30	0,75
9	EdTech Advocate	Master in Education	EdTech broker	Corporate and Business	Yes	20 – 30	0,75
10	Specialist in EdTech development	Master in Electrical Engineering	Software developer	Corporate and Business	Yes	10 – 20	0,6
11	Researcher on EdTech	Doctor in Chemistry	Researcher/ Computer Science Lecturer	Academic	Yes	30 – 40	1
12	Specialist in EdTech development	Master in Computer Science/ Engineering	Software developer	Corporate and Business	Yes	10 – 20	0,55
13	EdTech Advocate	Master in Management	EdTech broker	Corporate and Business	Yes	20 – 30	0,7
14	Specialist in EdTech development	Master in Civil Engineering	Software developer	Corporate and Business	Yes	20 – 30	0,8
15	Specialist in EdTech development	Master in Electrical Engineering	Systems engineer	Corporate and Business	Yes	20 – 30	0,75
16	Researcher on EdTech	Doctor in Philosophy	Researcher/ Media and Communication Lecturer	Academic	Yes	30 – 40	1
17	EdTech Advocate	Master in Computer Science/ Engineering	Chair of the Board at EdTech company	Corporate and Business	Yes	10 – 20	0,6
18	Researcher on EdTech	Doctor in Education	Pedagogy R&D in business and CPD trainer	Corporate and Business	Yes	10 – 20	0,7

3.2.2. Rounds Implementation

The implementation of the Delphi method involved three iteration loops, with a synthesis facilitated by the researcher's regular feedback and the comparison of the results with informed literature (Green, 2014). The various rounds included (1) answering a questionnaire; (2) reviewing first answers and select the most important critical points for each criterion; (3) vote on the new ideas to define a final list of criteria; this one would provide the inputs for constructing hypothetical scenarios that reflect the ethical challenges AIED poses; (4) discuss the plausibility of the scenarios, rewrite them and select those that better portray the ethical challenges of AIED.

3.2.2.1. Iteration 1

Participants were given an 8-item questionnaire (cf. <https://forms.gle/2CqBDsyy3p2n1jpE8>) to share their knowledge, vision, and opinion on the intersection of AIED and ethics. In the process of designing the questionnaire, the research team ensured its validity by drawing insights from various sources and grounding the questionnaire items on relevant literature, specifically codes for the responsible use of AI. The Artificial Intelligence's Ethics guidelines for trustworthy AI from the European Commission (2019) and Nesta's (2019) Map of the global AI governance landscape were instrumental in shaping the questionnaire, particularly questions 4, 5, and 6. These documents shared similar principles and recommendations across the stages of AI creation, function, and outcomes. However, while incorporating these recommendations, it became evident that some guidelines were broad and lacked specific direction for practical implementation in educational contexts.

To address this, questions 1, 2, 3, and 7 were derived from Holmes et al.'s (2019) seminal work on the promises and implications for teaching and learning of AIED. To ensure reliability, the questionnaire underwent a pilot test with a small group of individuals similar to the target participants, including an Edtech developer, an educational researcher on ICT, and an Edtech purchaser. Their thorough review of the questions helped identify any ambiguities or misunderstandings, leading to necessary adjustments to enhance clarity. Only one minor change related to language clarity was made in the last question. As well, the team employed the test-retest reliability approach to assess the stability of responses over time. Participants were asked to freely answer the questions and then answer them again after three months to evaluate the consistency of responses on separate occasions before the questionnaire was sent to the experts. After collecting the experts' 18 responses, the coordination team condensed each

meaning unit to identify broader categories, and descriptive statistics were used to determine the frequency of each category.

3.2.2.2. Iteration 2

In the second round, the 18 participants were presented with the results and asked to rank the importance of each category based on their personal views. Of the participants, 12 (67% participation rate) submitted their responses, which is still within the recommended range of 5 to 20 experts for qualitative research on a new topic (Landeta, 1999). While participants were aware of all study phases, the decrease in participation rate can be attributed to the demanding nature of this research phase, which occurred during the sudden pandemic-related restrictions and uncertainty in 2020. Content that was classified as medium-high and high was incorporated at this stage (representing challenging or very challenging issues), and categories with a sum of frequencies equal to or exceeding eight (more than half of the participants' votes) were retained for the third iteration.

3.2.2.3. Iteration 3

In line with the tradition of merging Delphi data with current literature, the third and final round employed a “Content Relational Analysis”, combining the collected data with the ethical principles outlined in “The Ethical Framework for AI in Education” (The Institute for Ethical AI in Education, 2021), which was created to guide the design, procurement and application of AI on behalf of learners. The goals of the ethical framework and the opportunities and challenges of AIED found by the experts were consistent and therefore interrelated by two elements of the coordination team and reviewed by two others. The cognitive mapping presented in Table 12, below, was used for the experts to construct hypothetical scenarios based on these ethical categories and the possible outcomes of AIED implementation in different scenarios and from the perspective of diverse educational actors.

As part of the process of hypothetical scenario construction, two types of scenarios were considered: normative and exploratory. Normative scenarios illustrate pathways toward desirable outcomes, while exploratory scenarios examine potential developments, regardless of their desirability (Kosow & Gaßner, 2008). In this study, we followed the basic steps of exploratory scenario planning proposed by Dean (2019). The first step was the (1) scoping phase, which involved defining the exercise's thematic coverage, stakeholders, and timeline. The (2) information-gathering phase analysed various data sources, including updated key reports like “The Ethical Framework for AI in Education”. The (3) trend and uncertainty

analysis involved analysing possible future situations in terms of their likely impact and level of uncertainty.

The principal investigators of this study performed this analysis, followed by the Delphi experts in the third iteration, as further described below. In the (4) scenario-building phase, the coordination team created eight hypothetical scenarios based on the experts' input. These scenarios were designed as short exploratory vignettes that presented a difficult-to-solve dilemma, following the orthogonal construction (Wright et al., 2014) and portraying one of four situations (the horizontal axis representing the degree of impact and the vertical axis representing the degree of uncertainty). They describe potential risks that may arise while striving to achieve eight out of the nine goals outlined in “The Ethical Framework for AI in Education”. The objective of managing administration and workload was excluded from this analysis as the experts' insights regarding opportunities and challenges did not align with this category. Table 12, below, provides an example of the approach taken during this phase.

The experts thoroughly examined all these scenarios. This exercise, which involves using key criteria to assess scenario quality, has a long tradition (Greeuw et al. 2000; Kreibich, 2007). Although scenarios are of a hypothetical nature, they are by no means arbitrary and must be evaluated according to criteria such as plausibility, consistency, comprehensibility and traceability, distinctness, transparency, degree of integration, and quality of reception (Kosow & Gaßner, 2008). In this study, participants were asked to give feedback on each scenario based on five criteria: 1. plausibility – whether it seems possible, 2. consistency – whether it makes sense logically, 3. comprehensibility – whether it is easy to understand, 4. relevance – whether it is relevant, and 5. distinctiveness – whether it is different from the others.

Table 12

Delphi Third Round: Tables with Hypothetical Scenarios Sent to Experts

SCENARIO 1. OBJECTIVE FROM “THE ETHICAL FRAMEWORK FOR AI IN EDUCATION”: ACHIEVING EDUCATIONAL GOALS				
<i>AI should be used to achieve well-defined educational goals based on strong societal, educational or scientific evidence that this is for the benefit of learner.</i>				
RISKS				
Future with AIED	AIED Targets	AIED Function	AIED Outcomes	Schools support
Broader Implications	School Administrators and Government	Content Production	Education Quality	Product Procurement
Information as commodity	Responsibility for frameworks and assessment	Accountability for content selection and curation	Accelerated Individuation	Framework definition for procurement
Mainstream thought and standardised behaviours			Reproduction of trainable skills	
Knowledge Management & Share			Human-human Interaction	
Instant data uploads on any topic			Challenge in terms of emotions and deep meaning	
Personalised/Smart Learning Platforms and MOOC's				
Skills enhancement				
Enhanced high-level processing				

HYPOTHETICAL SCENARIO: May 2033. Chloe is wakened by her robot friend Roy. Chloe checks her personalised learning chronogram for the week and starts her learning day in her garden. She activates her mindfulness chip and breathes in deeply. The programme suggested a Steve Jobs’ hologram on entrepreneurship for the morning, followed by a virtual meeting with John and Mitchell to discuss best practices to run a sustainable out-of-the-earth business ten years in the future. When they meet, they quickly benchmark what has already been done in the field, using their extended brains (what once was Internet is now connected to neural processes through implanted chips). Her backpack was prepared by Roy for the afternoon, with lots of sensors and some acrylic paints. Chloe creates a handmade painting that is able to provide multiple experiences through sensors: for example, if a viewer were to touch the sun, they might feel heat. All her academic performances are automatically graded on 10 parameters, including engagement, creativity, critical thinking, and empathy. She will only have one brick-and-mortar learning experience this week with ten students – five of them will meet on that day for the first time. The capacity to build learning opportunities with new colleagues is a highly appreciated skill set. She was given the names of the new colleagues and after receiving her teachers’ feedback along with the automated grading, she is ready for her favourite part of the day. She automatically screens all the new colleagues’ social media profiles through her favourite filters: dating history (longer and shorter relationships), trips to other countries, preferred clothing styles, and personal physical enhancements already completed. Each one immediately receives a score that matches Chloe’s already recorded preferences. She celebrates with Roy the fact that she found a 17-year-old boy who had a particularly high matching score. If nothing else, at least academic compatibility and empathy will be certain. She can hardly wait for that face-to-face learning day!

3.3. Results Per Round

3.3.1. Iteration 1

For each category created upon the participants’ responses, frequencies were determined by descriptive statistics (Table 13, below).

Table 13

Delphi First Round Results: Frequencies Per Category

1. If you were asked to forecast the future (academic/social/emotional data mapping for employability, online learning, impact and ROI...), of Education with Artificial Intelligence, what would you highlight?	FREQ
Broader Implications	
Education as business	1
Education dissolution	1
Information as commodity	1
Mainstream thought and standardised behaviours	2
Larger learning divide	1
School dissolution	1
Trade-off: privacy	1
Employment	
Better matching people-education-jobs	4
Knowledge Management & Share	
Instant data uploads on any topic	2
Personalised/Smart learning platforms and MOOC’s	3
Learning Processes	
Better teacher-student relationships	1
Learning personalisation	1
Learning practices impact & predictability	2
Real time engagement and performance assessment and feedback	2
Teacher routine tasks facilitation	2
Relational and Societal Factors	
Loss of interaction & detachment	2
More humanistic causes, leisure and culture	1
Skills Enhancement	
Enhanced high-level processing	4
Strengthened role of emotional intelligence and social skills	2
2. From the applications of AI in formal Education (automated assessment or feedback, curated content, learning personalisation, virtual teaching assistance...), which of them should encompass a particular ethically sound design?	FREQ
Automated Assessment/Feedback	
Integration of cognitive, social, emotional and psychological inputs into automated assessment	7
Curated Content	
Consider bias, limits, and manipulation	2
Learning Personalisation	
Constraint of personal freedom and capacity of thought	2
Misplacement of personal singularities	2
Privacy (publishers, vendors, schools, students)	2
Virtual Teaching Assistance	
Respect for students’ self-determination and autonomy in virtual teaching	2
3. Regarding the learner, the educator or the school institution, which AI target would be of more ethical concern to you?	FREQ
Learner	

Education systems not prepared to respond to kids' will to know more about AI	1
Performance-oriented solutions - avoidance of kids' "bad" emotions (frustration, boredom...) and refusal of its developmental insight	2
Students' exposure to AI conception shortcomings	2
Students' vulnerability to abusive power	6
Schools	
Data processing, encryption and storage	2
Educational actors training	2
Protection of school as a pre-technological space for thinking and imagining	1
School Administrators and Government	
Responsibility for frameworks and assessment	1
Teachers	
Preparation and training	3
Risk of replacement	1
4. When thinking of AI applied to Education, if you had accountability for AI creation which issues (accountability, biases, diversity representation, explainability, fairness...) would you be more ethically apprehensive about?	FREQ
Accountability	
Aligning industry's goals with education systems and individuals	5
Reinforcing accountability first to enable AI development	1
Diversity representation	
Inclusiveness, equity and fairness	8
Representation of local strengths and needs	1
Technological and ontological diversity	1
Explainability	
Fostering critical users	1
Enhancing transparent processes despite interests	2
5. When thinking of AI applied to Education, if you had accountability for AI function which issues (cyber security, decision making processes, hacking, privacy control...) would you be more ethically apprehensive about?	FREQ
Cyber Security	
Hacking and exploitation	5
Privacy control	6
Unauthorised use of personal information	3
Content Production	
Accountability for content selection and curation	4
6. When thinking of AI applied to Education, if you had accountability for AI outcome which issues (automation, citizenship, human-human interaction, labour changes...) would you be more ethically apprehensive about?	FREQ
Automation	
Blindfolded processes	4
Education Quality	
Accelerated individuation	1
Reproduction of trainable skills	1
Human-human Interaction	
Challenge in terms of emotions and deep meaning	3
Danger of virtual interaction replication	1
Distance between people	1
Interaction among community members	1
Political Citizenship	
Vulnerability for those who have lower self-determination	3
Training and Labor Changes	
Demanding labour market	2
Students' academic and career development	1
Vocational education and training (VET) threatened	1
7. In terms of schools' support (product procurement, students' privacy, teacher preparation...), which areas would you consider of priority care?	FREQ
Educational actors	

Communities' awareness and preparation	1
Involvement of all educational actors in AI rules definition	1
Product Procurement	
Framework definition for procurement	2
Sorted by internal school stakeholders	2
Teachers Preparation	
Capable of guiding students in tech usefulness	1
Greater AI literacy, awareness, and questioning	10
8. Considering the topics under study, please add any note or commentary, if you wish to.	FREQ
AI and its applications cannot be understood as all the same	1
AI cannot replace education as human relation	1
AI outcomes are political matters	1
Philosophical conundrum of ethics: utilitarian position ("the greatest good for the greatest number") VS. Kantian universal ethics	1
Schools should not be wired	1
Tech innovation should be an act of imagination exerted by all	1

3.3.2. Iteration 2

Regarding the future of AIED, some of the initially proposed categories have been excluded. This happened with the items "Employment" and "Relational and Societal Factors", which look at how education relates to impacts on other specific social layers. This appears to be true for both the negative perspectives – "Loss of interaction & detachment"– and the positive ones – "More humanistic causes, leisure and culture" (both from the category "Relational and Societal Factors") or "Better matching people-education-jobs" (category: "Employment"). The most extreme views concentrate a smaller number of votes: for example, "Education dissolution" (3 votes for medium-high and none for high) or "Schools dissolution" (5 votes combining medium-high and high).

The highest rated answers correspond to questions more directly related to academic achievement and the improvement of didactic resources: "Knowledge Management & Share", "Learning Processes", and "Skills enhancement". Looking at the critical positive/negative spectrum, four dimensions stand out positively, meaning they can have a noteworthy impact: "Instant data uploads on any topic" and "Personalised/Smart Learning Platforms and MOOC's" (both from the category "Knowledge Management & Share"), "Real time engagement and performance assessment and feedback" (category: "Learning Processes"), "Enhanced high-level processing" (category: "Skills enhancement"). All of them are related to improving academic performance. On the more challenging side, four dimensions stood out: "Education as a business", "Information as commodity", "Mainstream thinking and standardised behaviours", and "Larger learning divide". All of these points belong to the "Broader Implications" category, showing that these experts seem to agree on more global

negative impacts of using AIED, namely in terms of politics and asymmetries related to (the quality of) access to education, its instrumentalisation for profit and the emergence of dominant standardised attitudes.

3.3.3. Iteration 3

To assess and redesign the eight scenarios, the experts' feedback was analysed and broken down into meaningful parts. For instance, one expert raised a consistency question about the 2nd scenario related to evaluation: "[isn't there] more fear of correcting errors immediately or wanting to respond to the review? Do you think that could happen? The immediacy can trigger action and in some cases trigger fear" (Expert No.3). These inputs were considered and the scenarios updated accordingly. Five scenarios were evaluated based on how easy they were to understand. On the positive side, six scenarios were considered highly plausible, and five were deemed relevant.

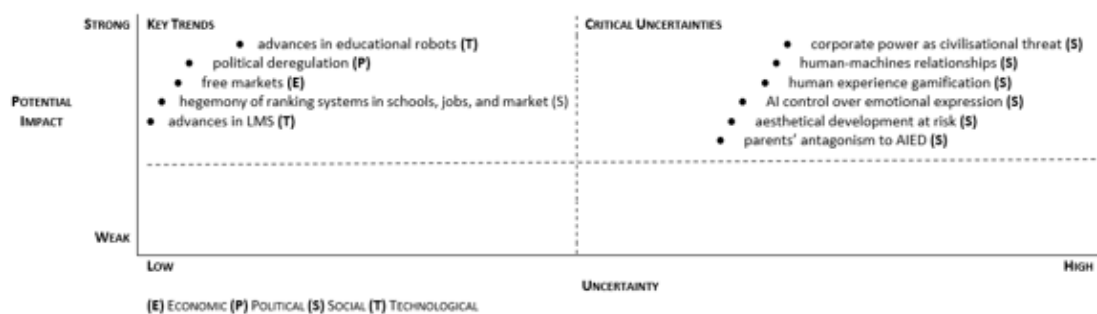
In terms of plausibility, however, five scenarios were less convincing (scenarios 1, 3, 4, 6, and 8). Parts of five out of the eight scenarios were neither believable (scenarios 1, 2, 4, 5, and 8) nor consistent (scenarios 1, 2, 3, 5, and 7). The fifth scenario, which concerns privacy, raised concerns about potential opposition during teacher training. The experts recommended clarifying the distinctions between scenarios 3 and 5 and proposed a dialogue between scenarios 1 and 8. Scenarios 1 and 5 underwent significant changes. The last three scenarios (6, 7, and 8) received fewer comments, either because they were more consistent or because the experts were less likely to provide feedback after initial involvement. In fact, the literature suggests that the optimal number of scenarios should be between two and four or five for ease of manageability (Dean, 2019).

In summary, the experts' recommendations were focused on several areas, including: 1. increasing the trade-offs between good and bad outcomes in the scenarios; 2. adding more biographical details to the characters; 3. setting some of the dates further in the future; and 4. creating clearer distinctions between scenarios that examine fairness and the preservation of a privileged status by private schools. Additionally, two participants predicted that some scenarios may be difficult for teachers to discuss due to their futuristic nature or because they portray teachers as passive. Regarding comprehensibility, the experts' understanding of future outcomes and trends resulting from AIED growth is depicted in Figure 6, below. Certain factors are emerging as trends (low uncertainty and high impact) in the economic (E), political (P), social (S), and technological (T) domains. Other aspects reflect critical uncertainties (high uncertainty and high impact), primarily in the social (S) domain. Examples of these critical

uncertainties include the dominance of classification and labelling systems, the gamification of human experience, AI's influence over emotional expression, the power of corporations as a civilisational threat, the importance of human relationships, the possibility of parental rejection of AI without proper ethical oversight, and the need to preserve students' aesthetic development in the highly technical global context.

Figure 6

Impact and Uncertainty Grid for AIED Futures Scenarios



3.4. Discussion

As Holmes et al. (2021) suggest, the ethics of AIED is expected to call our attention to the ethics of education in the first place. Each new technology entering the realm of education is, in fact, an opportunity to rethink education ethics and how particular technological features may hinder specific aspects of pedagogy. Such challenging systems present an opportunity for schools to discuss and define their ethical common ground and to develop strategies to overcome any obstacles that may arise. Collaborating in this way can be helpful to identify and solve unexpected problems that may emerge while using AIED, problems that were not taken into account during the design process. So, which aspects of pedagogy need to be safeguarded by ethical considerations? By using the ik-model (Mouta et al., 2015), which is designed to assess how technology is integrated into education while prioritising educational goals, it is possible to grasp the potential risks and benefits of having such an autonomous agent incorporated into pedagogy.

It can be argued that the majority of ethical preoccupations of using AI in education were found in the “relational dimension” (the social domain of the impact and uncertainty grid), highlighting the ethical importance of how people interact with one another in the advent of AIED. The designed scenarios acknowledge the importance of shared values, including both

explicit and implicit educational agreements. They also prioritise stakeholder involvement in decision-making processes, while being mindful of disparities in technological access and pedagogical quality. The first scenario centred around achieving educational goals and all participants agreed that the use of AI in education must be backed by robust evidence that demonstrate its beneficial effect on learners, what was previously highlighted by research (Richards & Dignum, 2019). The third scenario placed an emphasis on equity and highlighted the risk of a wider learning gap, where the disparities between public and private schools could become more prominent.

This disparity extends to significant gaps between developed and developing nations, socioeconomic groups within countries, and those who have AI-enhanced jobs versus those who are susceptible to being replaced by them (Miao et al., 2021). The seventh scenario addresses informed engagement and recommends that students and other education actors should possess an adequate understanding of AI and its implications. The experts suggest that individuals with AIED knowledge and the ability to question should participate in establishing AI policies at the school level. Levinas (1969) contended that ethics must primarily acknowledge the importance of the interpersonal dimension, when in search of its existential ground. And this comes before any consideration of concepts like utility, virtue, or duty. It is, thus, critical that the design, implementation (education with AIED), and evaluation (educational results) of activities with AIED be collaborative and shaped by people who have the capacity to consider the individual and societal benefits and drawbacks of its adoption and governance.

When designing these systems ("technological dimension"), it is important to be cautious about how the technologies are built, considering the type of connections people may form with machines. The sixth scenario emphasises the need for transparency and accountability in overseeing the operation of AI systems. In scenario eight, the importance of involving individuals who understand the potential consequences of AI on individuals and society in the design of these technologies is highlighted. This situation is exemplified by social robots that interact with humans. In fact, a review of the literature on the use of educational robots has evaluated their impact on four main dimensions, which are expected to be carefully scrutinised: 1. privacy; 2. human replacement; 3. impact on students; and 4. accountability (Serholt et al., 2017). To meet this need, it is encouraged to use a Trustworthy AI Ethics Guide in both creating and utilising AI technologies (European Commission, 2019) and also to promote a participatory

design of these technologies, informed by a variety of educational stakeholders and research fields.

Considering "content knowledge" and implementation through the lens of Levinas, education can be viewed as an ethical practice that aims to create spaces where individuals can engage with one another in a caring manner. The fourth scenario demonstrates how AI systems can undermine student autonomy, disregarding even the most capable and perceptive students. Dependence on automated decisions and AI-driven personalisation can limit opportunities for student interaction and focus on knowledge that is easier to automate, hindering their development of resourcefulness, self-efficacy, self-regulation (Miao et al., 2021), and the recognition of themselves as the citizens they already are. Moral deskilling can also affect educators, who increasingly rely on AI machines to make decisions and become less critical and morally engaged (Tundrea, 2020). The fifth scenario focuses on privacy and the use of personal data to achieve educational goals.

Experts warned of the possibility of education becoming a business with AIED being used, with many opportunities to enhance teacher training but at the expense of privacy and with the main goal of providing a specific service (Pammer-Schindler & Rosé, 2022).

In the dataism era, another ethical concern of AI-based education relates to the possibility of turning individuals into measurable and controllable entities through digital experiences. According to Han's (2014) argument that dataism could reduce self-tracking to mere self-surveillance, it's crucial to foster collaboration between teachers and students to envision and establish desirable futures with this unprecedented level of access to data. This is an invitation to reflect on what it means to be an individual in a group, and to foster mutual growth through reciprocal interactions. Educators also have the responsibility to unpack with their students the onto-epistemic grammar of dataism. This ethical undertaking involves exploring the anthropocentric perspective (Andreotti et al., 2015) underlying this desire, as well as the drive for ontological security (Lados et al., 2022) and the thirst for absolute knowability (Stein et al., 2017). This also provides a chance to use pedagogical strategies for a deeply purposeful and ethical learning experience.

Project-based learning and curriculum infusion can be powerful strategies for achieving integrated goals. By incorporating other ethical, societal, and political concerns from different fields' perspectives, these approaches can address the challenges posed by systems that can grow in agency through our own inputs, while still meeting curriculum standards. Educators can use a variety of subjects and make students apply them to the task, while also availing

themselves of AI resources. Engaging in discussions about AI functions from the perspectives of different subjects such as Mathematics, Science, History, and Languages (or any other) can serve as a means of strengthening newly acquired knowledge in these areas, applying it to practical and analytical tasks, and simultaneously building a more varied and intricate understanding of the AI systems in question. By doing so, education can move toward a more ethical exercise of freedom, even in the face of digital pressure.

Considering evaluation and “learning processes” dimension, AIED can provide just-in-time assessments, as well as new insights into how learning is progressing. But before recognising the potential benefits of incorporating AI-based assessment into learning environments, it is necessary to address ethical concerns related to educational assessment.

While it is true that obtaining high-quality knowledge is extremely valued and that AI can improve the processes of encoding, storage, and retrieval by offering personalised pathways, discussions with education experts indicate that this can only be achieved if there is mutual agreement and respect between individuals and what to expect from their interaction with autonomous systems. This is even more crucial now, as the pandemic caused by Covid-19 has given new impetus to technology (García-Peñalvo et al., 2021). As exam proctoring in some regions was a response to the problem of not being able to test students in physical situations, AI was identified as a possible solution to a large number of educational challenges. These ethical concerns were directed towards the second scenario that focused on forms of assessment. The experts believed that automated assessment and feedback on cognitive, social and emotional performance could become a reality in the near future and that this could present challenges and potential risks. Earlier research suggested that supervision is effective in decreasing deceitful actions. Students, however, may only behave honestly because they know they are being watched and not because of any intrinsic drive or self-reflection. This can lead to feelings of discomfort, such as a lack of privacy and anxiety, during the assessment process (Gudiño Paredes et al., 2021).

Drawing from Hannah Arendt's ideas, Coulter and Wiens (2002) suggest that to make sound educational evaluations one must establish a connection between the teacher (actor) and the researcher (spectator). It is critical to challenge teachers to become accurate judges and actors themselves, which involves creating opportunities for them to appear. In fact, this is the goal and major ethical responsibility of this research: to engage teachers in the development of the curriculum for a course on AIED as part of their continuing professional development, using this scenarios toolkit as a basis for discussion.

This represents an effort to urge teachers to become judging actors, which constitutes both a moral-political and an educational issue. These teachers are expected to engage with each individual child in complex communities, balancing guidance and agency and encouraging children to make informed judgments about the actions of others and reflecting on their own actions and choices. To expect teachers to foster these skills, they must receive training in these very principles. The process of using AI systems to evaluate performance presents a challenge: everyone involved should be asked to participate in understanding the feedback given. When personal perceptions of performance differ from the classifications provided by the AI system, it can be both a valuable opportunity for personal growth (through insight) and understanding of how AIED works. This approach will allow for constructive criticism and questioning, forming the foundation for critical engagement with the world.

It's important to recognise that AI is more than just a neutral tool; it's an agent that learns, interacts, and can impact outcomes, which can create conflicts between students, teachers, and the educational system in terms of agency. While AI-powered chatbots and virtual assistants can provide students with 24/7 support and resources, thereby increasing their autonomy, there are also risks associated with AIED that could undermine this. For instance, if AI is used to make decisions such as determining which courses learners should take or what career path they should pursue, it could limit their options and opportunities for self-determination. This, in turn, may restrict their ability to explore their identity (which is crucial for psychosocial development) and form a sense of self. Similarly, if AI is used to monitor student behaviour or performance, it could lead to a surveillance culture that restricts students' ability to take risks and make mistakes, which are deemed essential for growth.

The Vygotskian notion of 'scaffolding,' which involves a skilled mentor providing guidance and encouragement for action, may be interpreted differently in the context of AI. Since AI may not be able to offer the same level of support and encouragement as a human mentor, it could alter the perception of the teacher's role, who is expected to provide challenging emotional experiences that are crucial for confidently engaging with the world.

To conclude, this toolkit aims to bring intervention by providing educators with a comprehensive set of resources and guidelines, enabling them to effectively address pedagogical challenges, including the integration of AIED as supportive tools when appropriate, considering ethical aspects and potential challenges. Designed for use in continuing professional development of educators, the toolkit will consist of training modules and workshops covering fundamental AI concepts and their applications in education, how AI

can address current educational needs, the benefits and challenges of AI implementation in the classroom, and the contextual integration of AI in various settings. It will also foster participatory and collective agency and decision-making among educational stakeholders to define ethical and pedagogical aspects of AIED implementation that better suit their educational contexts and interests. Additionally, this toolkit can be effectively employed as a scenario-based learning tool for students in project or inquiry-based learning, encouraging exploration of real-life situations and challenges that arise when using AI in the classroom, thereby empowering students and enhancing agency within the school environment.

Finally, it can serve as a valuable resource for developers, providers, and educational decision-makers by offering guidance on ethical considerations related to AI usage in education. Within the scope of this study, the next stage involves the development of a continuing professional development course for educators. This phase will include close collaboration with teachers to explore the use of the toolkit to foster ethical reflection on education and AI. This methodological approach seeks to generate relevant insights, refine the toolkit's applicability, and systematically explore how it fosters educators' pedagogical reflection and supports students' ethical engagement with AIED.

This study does not rely on predictive analysis, but rather on plausible or possible futures (Brey, 2017). It uses short-term future narratives and the provision of information on the potential and dependencies of emerging technologies, which helps to bolster the decisions made throughout the work (Brey, 2017). The study, admittedly, faces certain constraints, including the difficulty of conducting research grounded in objective moral reasoning, ensuring fairness, mitigating unequal power dynamics, and fostering equal participation (Hagendijk & Irwin, 2006). To mitigate these challenges, the Delphi method was used.

Another potential limitation is that this study involved a relatively small number of expert participants (n=18), which may raise concerns about the representativeness of the insights gathered. There is some risk that some perspectives or expertise relevant to the topic might be underrepresented, leading to conclusions that might not fully capture the complexity and nuances of the ethical challenges related to the integration of AIED in education. To address the limitations of the sample size, the research coordination team made a deliberate attempt to include a diverse group of experts from different regions, including Europe, Southeast Asia, Middle East, and North America.

These experts possess varied backgrounds and expertise in the field of AIED. There was also a concern to ensure that the eligibility criteria encompassed a range of professional

profiles, such as government advocates, researchers, specialists in implementation, and specialists in EdTech development. Additionally, the Delphi method involves multiple iteration loops and expert feedback, meaning that data collection continues until a point of saturation is reached, where new insights or themes are no longer emerging from the panel. With 18 participants, it seemed possibly to efficiently reach this point, allowing for an in-depth exploration of the research questions. In the continuation of this research, qualitative research methods will be integrated, including focus groups with educators. These methods will complement the Delphi method, providing deeper insights into participants' perspectives and experiences.

3.5. Conclusion

This study reports on an expert consultation on AIED. The goal of the study was to foster debate on ethical AI integration in education and support teachers' continuing professional development through scenarios that will serve as a toolkit for discussing training syllabi. The scenarios created feature a combination of current AIED technologies and some dystopian elements. They highlight how these systems may significantly impact our daily lives, interactions, thoughts, and emotions, being reasonable to expect that there may be significant challenges that arise at a societal or even civilisational level. Therefore, it's important for educators to be mindful of the potential risks and benefits of using AIED, particularly with regard to emotion recognition and social choice, and collaboratively establish purpose for its use. This means knowing fundamental characteristics, potentialities, and challenges of AI, including its general functions, and being transparent about how AI is being used. It also means involving students in decisions about how these technologies are being implemented, how their inputs are incorporated into the learning experience, acknowledging that AI is expected to be used to support or expand student agency.

The scenario toolkit was designed to serve as the foundation for focus group discussions with educators, aiming to anticipate challenges and align educational objectives and practices with the context in which these AI technologies are intended to be used. During these discussions, experts advised exercising caution, as teachers may feel uncomfortable discussing unfamiliar topics, causing the conversation to steer towards familiar territory. In the medium term, the goal of this research is to equip educators with the appropriate resources to participate in such discussions, preventing resistance and fostering constructive dialogue that enhances the overall discourse.

The decision to employ a participatory method was taken to obtain a more comprehensive perspective on the ethical challenges of AIED implementation. Being the initial phase of a research project that will subsequently involve educators, this first step aimed to stimulate increased involvement from pertinent stakeholders who could influence policymaking. The experts' unique considerations may already enable them to contribute to unprecedented critical evaluation of AI technologies' impact in education as responsible actors in the field. In addition, the methodology adopted in this study aims to conform to ethical principles, active participation and agency, which are the same criteria proposed for the assessment of AIED. Not only does this methodology aim to ensure ethical research practices, but it also seeks to instil and promote the values it supports in the individuals involved, thereby guaranteeing their application in the way learners are encouraged to develop in the presence of AI-based systems.

CHAPTER FOUR

Empirical Research II: Participatory Collectives

4.1 Introduction

This section presents the findings from the third phase of the research. Conducted with teacher responsible for K-12 teacher training, the study investigated how ethical considerations and pedagogical agency shape professional development in the AI era. Using a participatory research approach, four focus groups analysed the scenarios from the previous phase and provided insights into AI integration in education, ethical frameworks, and strategies for fostering teacher agency. The findings inform the design of a training programme that addresses ethical concerns from an agentic perspective, ensuring that all participants have agency in shaping the ethical and meaningful integration of technologies, while critically examining the rationality of technique. This phase builds upon previous research stages, further reinforcing the need for teacher-led professional learning to challenge dominant narratives and promote ethical, context-aware AI practices in education.

4.1.1. Ethical and Meaningful Pedagogical Standards through CPD

The significance of continuing professional development for teachers in enriching pedagogical practices is well-established (Abakah et al., 2022; Darling-Hammond et al., 2017; UNESCO, 2015). CPD tends to foster heightened expectations that indirectly influence student achievement (Rubie-Davies, 2015) and holistic development by enhancing teachers' capacity for utilising evidence-based pedagogical methods, designing intentional and meaningful learning experiences, and implementing relevant assessment techniques. It can lead to experiences of rupture with methodological normativity and reaffirm the inevitability of pedagogical mediation (Trindade, 2012). Teachers involved in CPD can serve as role models for students and peers, promoting the value of learning, knowledge sharing, and critical thinking within learning communities. CPD may also strengthen teachers' influence in decision-making processes, enhancing their decisional capital (Hargreaves & Fullan, 2012). Additionally, CPD can help teachers adapt to changing contexts, respond to students' evolving needs, and meet broader societal expectations (Learning Forward, 2022). This is especially relevant given the impact of technological advancements on students' lives.

Following an Ellulian perspective, CPD can serve as a catalyst for questioning the underlying mindset that governs our society and for the critical examination of the purposes and dynamics of technique, assisting educators in comprehending and navigating the intricate interplay between technology, pedagogy, and the broader societal context. Only an approach responsive to the challenges posed not only by AI technologies, but by a world driven by

technique, can reveal how its dominance has permeated all layers of society and institutions, extending even into education. Such an approach can redirect discussions on educational ethics far beyond superficial debates centred solely on AI ethics.

Kowalczyk-Waledziak et al., (2019) also assert that, given the intersection of ethical considerations with pedagogical approaches, there is a need to reconsider teacher education. This could be done through a participatory approach involving educators, to make meaningful progress in pedagogical terms (Chichekian & Benteux, 2022). While UNESCO (2024a) recommends governmental agencies to build capacity for the proper use of GenAI in education and research, a systematic literature review exploring the ethical dimensions of AIED technology implementation in education from 2011 to 2022 revealed a noticeable absence of emphasis on ethical considerations, despite some training projects being introduced in schools (Mouta et al., 2023). However, more recently, some initiatives are currently underway to address these concerns and requirements. Set to launch in 2024, UNESCO's AI competency framework for educators takes a step further by outlining a three-stage progression that covers areas such as fostering a human-centric approach, addressing AI ethics, understanding AI fundamentals and applications, incorporating AI into pedagogy, and utilising AI for professional growth (UNESCO, 2023b). A training practical example comes from AI4T (2024a), an Erasmus+ K3 initiative designed by France, Slovenia, Italy, Ireland, and Luxembourg, which emerged to facilitate AI education specifically for teachers and school leaders. Understanding the ethical implications of employing AI in educational environments was included as a component of the shared learning objectives across all countries' professional development pathways.

4.1.2. Teacher Agency

The dominance of a world of technique (the pre-eminence of rationality and effectiveness), the increasing diversity of students, the necessity to prepare learners for environmental transitions and to critically examine what development might mean, and the attempt to enhance learning and assessment through digital means greatly influence the direction of education. Thus, it becomes central to incorporate the needs of students and society into an andragogic narrative that specifically caters to teachers, recognising them as active contributors in co-authoring the educational plot. Intelligent professionalism suggests educators to be consistently involved in the processes that educational systems claim are enhancing education (Thompson, 2021). This is expected to come alongside the commitment to change within schools and among teachers (Kukulka-Hulme et al., 2020) and the recognition of families and community

members as vital partners (EU, 2018). Being a teacher involves taking a public stance on important educational issues and public policy development, elevating professional status across all systems (Thompson, 2021; UNESCO, 2021).

Such an approach to professional development can strengthen teachers' shared epistemic agency (Damşa, 2014) and reinforce ethical practices related to AIED knowledge. Moreover, it can promote teachers' relational and collective agency, reassuring the structural context supports and facilitates the interdependence, while understanding how agency unfolds over time (Damşa, 2014). This perspective aligns more closely with an ecological approach advocated by Priestley et al. (2015), which suggests that teacher agency isn't an inherent attribute of individuals but rather a result of the interaction between personal capabilities and the environmental factors that influence their actions.

Ultimately, teacher agency is recognised as pivotal, as educators possess the ability to actively approve, adapt, or resist policies and programmes", (Severance et al., 2016). While it is true that numerous decisions regarding educational technologies are often made from the top-down, involving government bodies, big tech companies, national research, education networks, and appointed staff from learning institutions, the ultimate realisation of these technological initiatives heavily relies on teachers themselves. Their awareness, knowledge, and, most importantly, their attitudes and level of engagement play a key role. To inform these aspects, training may be significant. Once educators are active agents in designing, implementing, evaluating, advocating for, and experiencing quality professional learning and the systems that support it (Cochran-Smith et al., 2022; Learning Forward, 2022), training initiatives built upon their knowledge and deliberation may yield meaningful investments for schools. Teachers are also expected to undergo training centred on the same principles of agency they are to instil in their students, enabling learners to exercise it when interacting with autonomous AI agents.

In this context, this phase of the study is a participatory research effort aimed at strengthening teachers' agency through training on ethics and AIED. It is rooted in the theoretical underpinning of dialogic ethics, which acknowledges the multifaceted nature of ethical concerns and the need for a thorough examination involving various stakeholders (Santos, 2012). While existing research has explored the theoretical aspects of AIED ethics, it often overlooks the perspectives of educators themselves. This study distinguishes itself by focusing on educators' voices through the lens of Continuing Professional Development.

Specifically, the gap in participatory studies in this field is addressed by utilising focus groups with those responsible for shaping teacher education.

4.2. Methodology

This component is part of a comprehensive educational design research, which began with a systematic literature review (SLR) to pinpoint major gaps in AIED and ethics (Mouta et al., 2024a). The findings from the SLR guided subsequent research steps, including the recognition of the need for the involvement of stakeholders from diverse educational backgrounds in research and capacity building in AIED and education ethics, especially among teachers. As a result, the Delphi method was employed to investigate educational experts' ethical concerns regarding AI in education (Mouta et al., 2024b).

With the goal of encouraging comprehensive debates on the development of professional learning in AIED, education ethics, and agency, these concerns informed the design of futures scenarios to facilitate in-depth focus groups discussions. Focus groups are commonly employed in educational research because of their usefulness for programmes development (Nagle & Williams, 2013) and their ability to yield valuable data to shape policies and practices. This section delves into the initial inputs arising from those focus group discussions among professors, teachers, and trainers regarding the attributes of such training. While the development of the professional learning materials is planned for later stages, this phase provides insights and discussion around the course's objectives, implementation strategies, and design aspects. This study's third stage is the main focus of the upcoming sections and is structured according to the framework proposed by Silva et al. (2014), which includes planning, preparation, moderation (referred to here as "implementation"), data analysis, and results communication.

4.2.1. Planning

The planning phase of focus group research is key to ensuring that clear objectives guide the moderation process.

4.2.1.1. Research Guiding Questions

To guide the design of meaningful training for K–12 educators on AIED, ethics, and agency, this stage addresses the following research questions (RQ):

- RQ 1: Which structuring dimensions do educators consider important for training programmes on AIED, ethics, and agency targeting K-12 pre-service and in-service teachers?

- RQ 2: What objectives do educators prioritise for such training programmes?
- RQ3: What specific aspects should be taken into account when developing a programme training programme in this field?
- RQ 4: What strategies can be used to update teachers' knowledge and skills on AIED while also fostering agency and ethical thinking about its potential impact on education?

4.2.1.2. The Moderation Scripts

To facilitate participants' reflections and deepen their responses, it is recommended to use a moderation script with a "funnel" strategy, where questions become increasingly specific throughout the session (Borges & Santos, 2005). While the main objective of these focus groups was to identify the goals, topics, and delivery methods for training programmes on ethical AIED use for teachers, some initial questions on AIED ethics were included to familiarise participants with the topic under discussion. In advance of the session, each participant was provided with two hypothetical scenarios that had been generated during the Delphi phase of this research (Mouta et al., 2024b). The aim was to motivate the participants for the discussions by providing these narratives. It is worth noting that these scenarios resulted from a three-stage iterative process, leading to their creation in 2021 for discussions held in 2022. AI applications to the field of education have progressed significantly since then, as emphasised in the introduction. Despite this, the scenarios remain relevant for the purpose of this research. Their dystopian aspects, however, and the potential shift towards vignettes illustrating current experimental cases during the training will be investigated in the discussion section.

The scenarios presented to the participants were the same for each element within a group, but they differed between groups. The script utilised during the session featured open-ended questions on participant attitudes towards AIED, ethical concerns, student agency, and the design of an educator training programmes within the context of AIED.

4.2.1.3. Participant Recruitment

To prevent dispersion, a relatively uniform group was established, selecting participants according to criteria that corresponded to the research objectives (Silva et al., 2014). For this study, participants with professional certification for teaching and experience in the field of education at any level were eligible for selection. Additionally, they were required to hold responsibilities in K-12 teacher pre-service or in-service education.

Participants were required to be native Spanish speakers or highly proficient in the language, as Spanish was the primary language of the research centre where the study was conducted. Initial recruitment used convenience sampling due to time constraints at the start of the research, followed by snowball sampling to ensure cultural diversity. To balance effective communication with cultural diversity, invitations were sent via email, with a larger pool (40) to offset potential refusals (Ogbeifun et al., 2016). Declines were mainly due to professional commitments or discomfort with the topic. Nineteen participants agreed to join and were divided into four groups. All provided informed consent, and the moderator explained the research objectives and confirmed their willingness to participate (Silva & Fortunato, 2021).

The mean age of the participants was 44 years old (SD 7.46) and 47.4% identified as women and 52.6% as men. The group included individuals from five countries, all of whom held responsibilities in preparing pre-service or in-service teachers in the K-12 education sector. They possessed pedagogical experience in either K-12 or higher education. Participants general sociodemographic and occupational profiles are presented in Table 14. below.

Each group engaged in an in-depth discussion of two scenarios selected from the hypothetical set developed in the previous Delphi phase, totalling eight vignettes analysed across four groups. According to Hennink et al. (2019), appropriate sample size for focus groups depends on factors such as study purpose, code characteristics, group composition, and desired saturation. Focus group studies typically involve three to six sessions, with more sessions used for diverse populations (Coenen et al., 2012; Guest et al., 2016). In this study, code and meaning saturation were reached after four discussions, meeting recommended saturation levels for non-stratified focus groups.

Table 14

Teachers' Sociodemographic Profile

FOCUS GROUP PARTICIPANTS' PROFILES						
GROUP no. 1						
Professional Profile	Age range			Gender		Working Country
	25-35	36-45	46-55	F	M	
Director of Digital Transformation in Higher Education (Tecnológico de Monterrey)			X		X	Mexico
Associate Professor in Research Methods and Diagnosis in Education (University Salamanca)	X				X	Spain
Former K-12 teacher. Director of the Ministry of Education's Master Plan			X		X	El Salvador
Associate Professor of Research Methods in Education (University León)			X		X	Spain

Former K-12 teacher. Manager of Academic Development for Higher Education at the Ministry of Education		X	X	El Salvador
Former K-12 teacher. Researcher in Teaching of Mathematics in K-12 education	X		X	El Salvador
GROUP no. 2				
Associate Professor (University Salamanca). Coordinator of Primary Education Degree	X		X	Spain
K-12 teacher. Director of the Fontán School Network		X	X	Colombia
K-12 teacher. Pedagogical Coordinator at Fontán School	X		X	Colombia
Teacher Trainer. Associate Professor in Information Technology (University El Salvador)	X		X	El Salvador
GROUP no. 3				
Full Professor in Computer Systems (University Salamanca). Academic Director of Quality Assessment		X	X	Spain
Full Professor in Research and Diagnostic Methods in Education (University Salamanca)	X		X	Spain
Associate Professor in Didactics of Experimental and Social Sciences (University Valencia)	X		X	Spain
Former K-12 teacher. Assistant Professor in Primary Education (Fray Luis de León University School)	X		X	Spain
Teacher trainer and project Manager (University Salamanca)		X	X	Spain
GROUP no. 4				
Assistant Professor in Educational Theory (University Salamanca)	X		X	Spain
Associate Professor in Didactics and School Organisation (University Santiago Compostela)	X		X	Spain
Full Professor in Research and Diagnostic Methods in Education (University Oviedo)		X	X	Spain
K-12 teacher trainer	X		X	Portugal

4.2.1.4. Preparation

Participants were notified through emails, which contained details about the research objectives, session duration, group specifics, focus group dynamics, session recording, and informed consent. The two hypothetical scenarios were dispatched 48 hours prior to the session. Since the invitees originated from diverse regions and the world still faced pandemic restrictions, it was deemed more appropriate to conduct virtual meetings, through the Zoom platform. Despite some limitations, research has shown that virtual focus groups can be just as productive as in-person focus groups in generating rich and meaningful data (Krueger & Casey, 2015).

4.2.2. Implementation

Four focus groups were implemented in the first semester of 2022, with varying numbers of participants (6, 4, 5, and 4) to accommodate their availability.

Participants were mostly unfamiliar with one another, except for two cases in the first and second groups, where professional roles were unrelated. No dominant speakers emerged, and participants ensured inclusive discussions. Session durations varied: 80 minutes (first), 64 minutes (second and fourth), and 62 minutes (third).

4.2.3. Data Analysis Procedures

Following Roberts et al. (2019), the deductive phase included a preliminary codebook to ensure credibility through three elements: content (sampling and instrument design), criterion-relatedness (tool consistency), and construct validity (alignment with research questions) (Long & Johnson, 2000). The following sections describe the methodology, including iterative coding, category refinement, and validity and reliability testing.

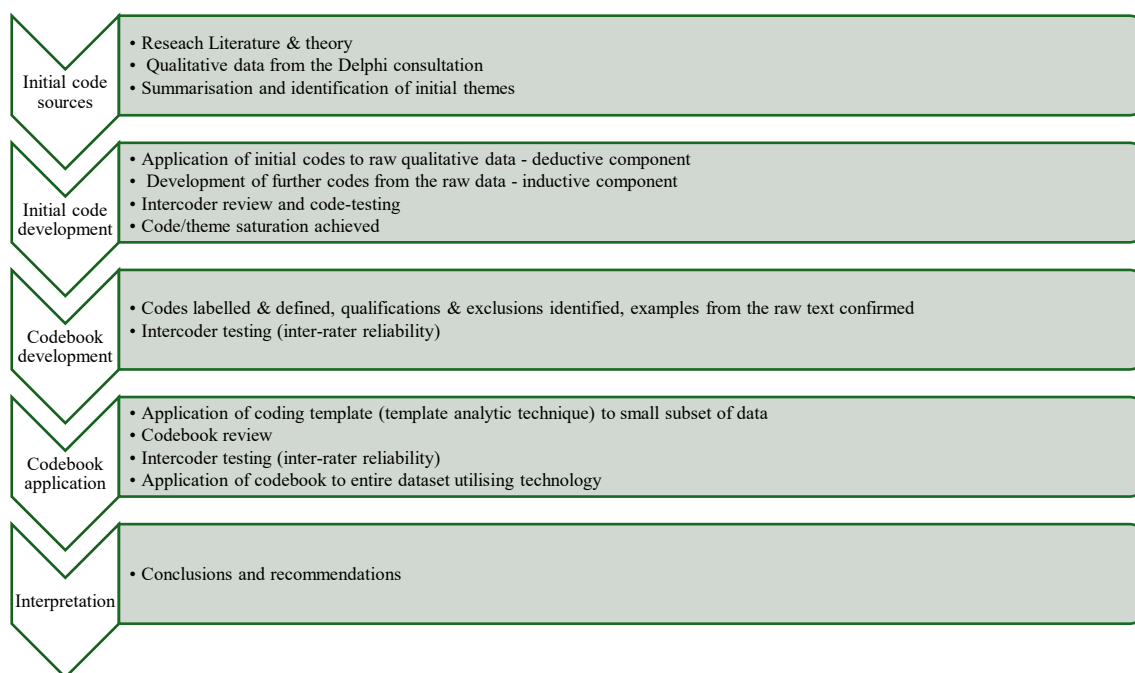
All sessions were recorded using the recording option in Zoom, after obtaining email and verbal consent from all participants. The recordings facilitated the verbatim transcription of the session's content using AI-powered software (Trint). All transcriptions were then thoroughly reviewed and cross-checked against the video recordings by the principal researcher. The analysis was based on the research question being asked ("How should training programmes targeting ethics in AIED for K-12 teachers or aspiring teachers be structured and organised?"), the initial analysis of the literature, the Delphi method approach undertaken as part of the project, and a preliminary scan of the raw data. Additionally, an inductive approach enabled the identification of unexpected themes during coding. Preliminary analysis involved a close review of the transcribed data, with selected segments marked for coding and analysed using NVivo software.

Moderator contributions were transcribed but not included in the analysis. The research team agreed on unit separation criteria upon importing the data into the software. A thematic approach was used, with a structural coding matrix developed from existing literature and initial readings of the raw data. New codes were added as needed when data did not fit the existing framework. This iterative process continued until no new codes emerged, confirming the matrix's validity. This process was repeated multiple times until no new codes were identified, confirming the matrix's validity. The identified codes were added to NVivo as nodes, and the coded text was matched to the nodes in a systematic way.

As shown in Figure 7, below, the combination of deductive and inductive approaches enabled the study to identify patterns both within and outside of the predetermined codes, resulting in a comprehensive and nuanced analysis of the research question.

Figure 7

Process of Code Creation and Testing



Note. Adapted from “Attempting rigor and replicability in the thematic analysis of qualitative research data: A case study of codebook development”, by K. Roberts, A. Dowell, and J. B. Nie, 2019, *BMC Medical Research Methodology*, 19, Article 66 (<https://doi.org/10.1186/s12874-019-0707-y>). © 2019 by The Author(s), distributed under the terms of the Creative Commons Attribution 4.0 International License.

The final selection of nodes resulted from two reliability tests: a test-retest reliability and an inter-rater/coder analysis (cf. https://drive.google.com/drive/folders/13fF4h6fGbrwf9woDKApJ7NulfoeNzLEr?usp=drive_link), conducted by two researchers. The Pearson correlation between initial and follow-up coding was 0.86, showing high consistency. Intercoder agreement was 81%, and Cohen’s Kappa was 0.83, indicating substantial agreement beyond chance. Analysis used a framework matrix with four main categories (attitudes toward AIED, benefits and opportunities, ethical

challenges, and teacher initial and CPD) plus an "other" category for unrelated content (Table 15). Descriptions and examples for each NVivo node are available in Mouta et al. (2025a).

Table 15

NVivo Framework Matrix

NODE NAME	
A. Attitudes towards AIED	C6. Education as a system
A1. Affective attitudes	C7. Examples
A1.1. Negative	C8. Families
A1.2. Positive	C9. Future of Society and Civilisation
A2. Cognitive Attitudes	C10. Investing in projects, research, and theories
A2.1. AI tech as neutral tools	C11. Lifelong Learning
A2.2. People's Agency	C12. Problems with Data sets
A2.3. Technological determinism	C13. School Administration
B. Benefits and Opportunities	C14. Students
B1. Citizenship and Equity	C15. Teachers
B2. Education as a system	C16. Work and Employability
B3. Examples	D. Teacher initial and Continuing Professional Development
B4. Families	D1. Broadened ethical thinking
B5. Future of Society and Civilisation	D2. Collaboration between teachers and with families
B6. Lifelong Learning	D3. Compliance with a policy framework
B7. People in general	D4. Dealing with resistance and transition
B8. School Administration	D5. Hands-on and skill-based training
B9. Students	D6. Informal self-directed CPD
B10. Teachers	D7. Innovative and purposeful pedagogical practices
B11. Work and Employability	D8. Learning by example
C. Ethical Challenges	D9. New role requirements
C1. AIED tech design and developers	D10. Pointless without AI tech
C2. Citizenship and Equity	D11. Pre-Service Training
C3. Communities and cultural context	D12. Questioning Educational Paradigms
C4. Developing Frameworks and Policies	D13. Setting priorities
C5. Different interests between users and corporations	D14. Update of knowledge and skills
	E. Other

Instances arose where information exhibited overlapping characteristics, being assigned to multiple categories simultaneously, or left uncoded if it did not align with a specific category. For this analysis, categories encompassing all the gathered information were established, although only specifically studying the data directly relevant to the main research question. In fact, the initial open-ended questions were designed to prepare participants for the discussion. Since the topic under consideration is relatively new or often overlooked in current educational discourse, these questions aimed to provide the necessary context for discussing training methods. The need to code such information stems from the fact that coding ensures that no data is overlooked and that all related aspects are taken into account during the analysis process.

However, for the purpose of this analysis, a deliberate decision was made to analyse and present the data that directly addresses teacher training in ethics in AIED.

4.3. Results

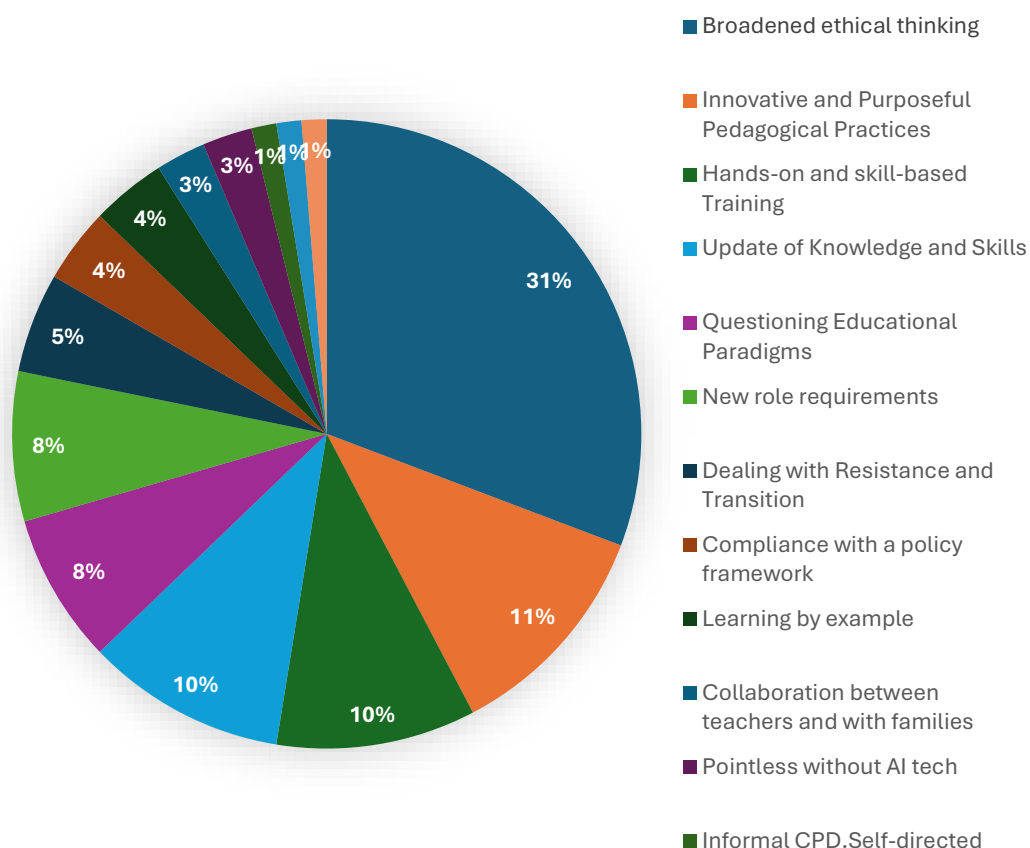
This section presents the key findings from the focus group discussions in relation to the research questions guiding this stage of the project.

4.3.1. RQ1: Structuring Dimensions for the Training Programmes

Concerning the findings on teacher initial and CPD, the relationships between the coded text excerpts and their corresponding nodes are examined. This analysis uses the data shown in Figure 8, below, which identifies the topics participants discussed the most. This figure provides an overview of the percentage frequency of each category mentioned across the entire dataset, highlighting its prominence in the overall discussions.

Figure 8

Number of References Per Item



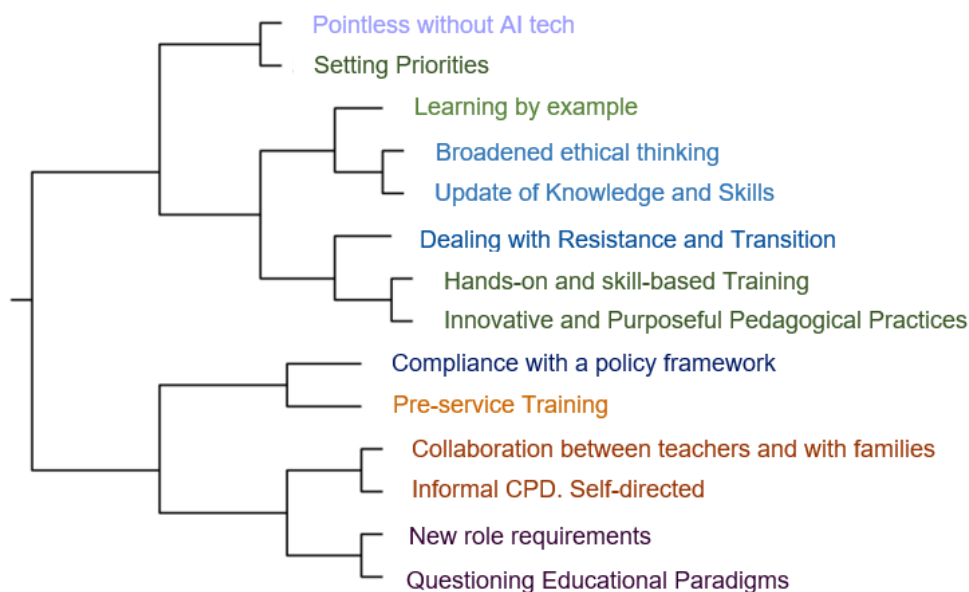
These discussions were further categorised based on different structuring dimensions of the training course. Analysis of node distribution revealed three key structuring dimensions: training aims and objectives, design considerations, and implementation strategies. Table 16, below, presents the percentage of discussion devoted to each category, highlighting their relative prominence in planning. For instance, while “broadened ethical thinking” constitutes 31% of the overall discussion dataset (Figure 8, above), it comprises 51.06% of the references associated with the “aims/objectives” category in Table 16. This “aims/objectives” category was indeed the primary focus of the participants' discussions, centring on the purpose of the training. With 60% of the discussions dedicated to appraising and explicitly articulating its objectives, it is clear that participants were committed to ensuring the course aligned with teachers' needs.

Table 16

Optimising Teacher Training: Planning Dimensions

AIMS/OBJECTIVES CATEGORY	NO.	DESIGN CATEGORY	NO.	IMPLEMENTATION CATEGORY	NO.
Broadened ethical thinking	51.06%	Dealing w/ Resistance and Transition	36.36 %	Innovative and Purposeful Pedagogical Practices	45%
Update of Knowledge and Skills	17.02%	Collaboration between teachers and with families	18.18 %	Hands-on and skill-based Training	40%
Questioning Educational Paradigms	12.77%	Pointless without AI tech	18.18 %	Learning by example	15%
New role requirements	12.77%	Setting priorities	9.09%		
Compliance with a policy framework	6.38%	Pre-Service Training	9.09%		
		Informal CPD. Self-directed	9.09%		

To further analyse the relationships between these discussion topics, NVivo's cluster similarity function was employed. This function measures similarity between nodes, considering the presence or absence of shared words between each pair of nodes. The results of this cluster analysis are displayed as a horizontal dendrogram (Figure 9, below). The following paragraphs will discuss categories and subcategories, as well as the connections between them, highlighting those revealed by the analysis.

Figure 9*Items Clustered by Coding Similarity*

4.3.2. RQ2: Aims and Objectives Prioritised by Educators

4.3.2.1. Broadened Ethical Thinking

The aims and objectives category accounted for a total of 47 references across five items. The emphasis on this category reveals a significant concern among the participants regarding the purpose and direction of an AIED ethics training course. Their main concern revolved around the necessity to foster broad ethical thinking; a topic extensively discussed in approximately 51% of the discourse segments. In this context, one of the participants expressed, "when I think about teacher training, I don't see it from a technicist point of view.... [it] is the space for reflection and construction, joint construction about what the challenges, the difficulties, the opportunities might be". This highlights the understanding that dealing with the ethical aspects of technological progress necessitates a holistic strategy that encompasses all disciplines. By cultivating a climate of ethical consciousness and accountability regarding technological advancements, individuals are given the opportunity to question the very ethical nature of the educational project.

4.3.2.2. Update of Knowledge and Skills

By recognising the significance of keeping knowledge and skills up to date, as evidenced by the subcategory that garnered 17% of the references, participants suggested that training is expected to question what is technically known, ethically debatable, and required to create a theoretical and conceptual rationale. The following excerpt illustrates this perspective: “we must perhaps have a significant ability to develop competencies for knowledge, to know the advances of science and their relationship with the benefits and problems that ethical aspects in sciences and technologies can generate for us”. Participants also underscored the ever-changing nature of the field and the imperative for trainees to remain abreast of the latest advancements.

In the context of CPD, offering training that not only responds to teachers’ current needs but also challenges them to anticipate and explore emerging challenges supports the development of professional agency. This dimension may also support educators to respond to any emotional discomfort that may arise from global shifts affecting their learning environments, whether due to emerging resources, technologies, or pedagogical trends. It may also encourage them to explore alternative approaches, experiment with emerging resources and help students engage meaningfully with the learning experience, keeping teachers motivated and satisfied in the long run.

The cluster similarity analysis in Figure 9, above, highlights a connection between these two subcategories: broadened ethical thinking and knowledge/skills update. This association underscores that extensive ethical reasoning in AIED leads to the recognition of the importance of updating knowledge and skills in this field. Conversely, a strong knowledge base is crucial for informing ethical deliberations accurately.

4.3.2.3. Questioning Educational Paradigms

As a result of this mindset and skills update, teachers will be instigated to question current educational paradigms, a subcategory that garnered 13% of the references. In this case, participants acknowledged the significance of motivating trainees to challenge the established norms and embrace a more comprehensive and ethical educational approach, facilitated by the dilemmas posed by this recent technological endeavour. As one of the participants stated, “It has to come accompanied (...) by a significant innovation (...) within the educational system. For example, it’s about ending the concept of exams, the concept of a teacher, as we see it today”.

4.3.2.4. New Role Requirements

Once teachers begin to challenge current educational paradigms, their role in the classroom will also be challenged, as stated 13% of the time and as the cluster similarity analysis underscores (Figure 9, above). This is exemplified by an excerpt from one participant: “becoming guides means positioning ourselves horizontally in the processes of development, in learning, in the construction process, in existence”. So, training is expected to encourage teachers to re-evaluate the adequacy of their current roles and practices, critically question what it means to be a “guide” when integrating AI technologies and understand data-driven insights in a holistic way.

4.3.2.5. Compliance with a Policy Framework

Participants recommended preparing teachers to comply with a policy framework, which accounted for approximately 6% of the recommendations, as expressed in this discourse: “teachers should indeed receive training (...) on the current legislation and how to apply it in the context of using applications that collect personal information, such as (...) Artificial Intelligence systems”. This highlights their concern for achieving conformity once legal and regulatory guidelines are established. Furthermore, it underscores the significance of enhancing teachers' professional credibility and safeguarding students' rights by ensuring consistent practices across contexts.

4.3.3. RQ3: Design Considerations for the Training Programmes

The design category was the least mentioned, receiving a total of 11 mentions. However, the existence of six subcategories related to this topic suggests that the participants differentiated their insights into the strategic and organisational aspects of the training programmes.

4.3.3.1. Dealing with Resistance and Transition

In 36% of the discourses, the subcategory dealing with resistance and transition was mentioned. It suggests a concern about how teachers might respond to the integration of these technologies in learning environments, due to their cognitive and affective attitudes or lack of familiarity with the theme of AIED. It encompasses resistance to change and managing transitions comprehensively. As one participant emphasised: “I see what I call the transition risk, meaning, in the initial stages, the first teachers may not adopt these new technologies. There will be a transitional phase”.

It is proposed that the training tackles concerns related to a smooth and contextually meaningful transition. The dendrogram's analysis offers insights on this matter, indicating the design category's proximity to two implementation categories (Figure 9, above). It suggests that addressing concerns within the design category may necessitate practical interventions to gain teachers' support. Hands-on strategies and innovative pedagogical methods aligned with a major purpose could prove instrumental in building teachers' confidence.

4.3.3.2. Collaboration Between Teachers and With Families

The subcategory that pertains to the collaboration among teachers and between teachers and families was covered by 18% of the data, indicating a strong concern in engaging with key stakeholders who play a proximal role in shaping students' ethical understanding and behaviour, as highlighted by this participant: “But it is here where we, especially teachers, need to take a stand, collaborating closely with families to enhance the quality of students' learning”. This also recognises that these actors hold the power to either support or undermine an educational project.

4.3.3.3. Pointless without AI Technologies

The subcategory that questions the relevance of training without AI technologies available at a school level was covered by 18% of the references, shedding light on the belief that ethics training in AIED holds significance when it is integrated with and informed by AI technology. Participants acknowledged that educational endeavours should be purposeful and aligned with the actual needs and practices of stakeholders, grounding ethics training in the available current technological landscape, as demonstrated by this participant: “It wouldn't make sense to do it too far ahead of the time when teachers will have the technologies to work with”. The cluster similarity analysis reveals an interconnection between this subcategory and the one that raises a thought-provoking question on setting priorities (Figure 9, above).

4.3.3.4. Setting Priorities

In 9% of the discourses, the importance of adding more training specifically focused on AI was questioned. Other shortcomings in teacher education related to inclusion and gender perspective were mentioned as priorities. The participant expressing this concern highlighted the following: “Are we prepared to take on more training, specifically in the field of AI, when we still have so much pending training related to inclusion, gender perspective, and many other areas in which we are...failing in teacher education?” Once again, the dendrogram provides insights on this matter, particularly highlighting the farthest neighbour clustering between the

categories of training's relevance without AI technology and the necessity for debates regarding new role requirements or educational paradigms (Figure 9, above). According to this pragmatic approach, without setting priorities for training focuses and ensuring teachers have access to AIED, this discussion is spurious.

4.3.3.5. Pre-Service Training

Nine percent of the references had also been made for equipping educators with the necessary knowledge and skills related to AIED ethics during their pre-service training, as demonstrated by this quote from one of the participants: “However, in many instances, the discussion of ethics needs to start from a more fundamental standpoint. This particularly applies to the generations now entering universities”. This integration has the potential to cultivate a culture of ethical consciousness and responsibility among future educators. Additionally, when teachers with different levels of experience and diverse knowledge exchange their perspectives in educational settings, positive outcomes can result.

4.3.3.6. Informal and Self-directed CPD

Nine percent of the discourses have emphasised the importance of maintaining an informal and self-directed continuing professional development (e.g., “And the positive awareness that comes with this type of tool, without the need for other types of courses”), revealing the recognition that professional growth in ethical practices should extend beyond formal training programmes. This can be fostered through reflective practice, learning communities, diverse resources, and collaborative learning. Such subcategory and the one that pertains collaboration among educational actors appear connected in the cluster similarity analysis (Figure 9, above). This suggests that teachers and families who actively engage in collaboration are more likely to be the ones that take ownership of their professional development and seek out opportunities for learning and growth independently. This proactive approach to professional learning and alliance-building reflects collective agency and a shared commitment to improving educational outcomes.

The cluster similarity analysis reveals that the nodes "compliance with a policy framework" and "pre-service training" share some characteristics, despite being represented by different colours, indicating a less close association between them (Figure 9, above). This near neighbour clustering highlights the importance that some participants place on establishing a strong foundation for the integration of AIED, suggesting it be formalised through policy and incorporated into the curriculum for pre-service educators. This concern calls attention to the

efforts that are expected to accompany teacher professional learning and that relate to policy development and curriculum design.

4.3.4. RQ4: Implementation Strategies

The implementation category received 20 references, indicating that the participants placed importance on the delivery methodologies of the training programmes. These references likely stem from the participants' experiences as trainers, suggesting the value of hands-on, skill-based training and innovative pedagogical practices in the context of AIED ethics.

4.3.4.1. Innovative and Purposeful Pedagogical Practices

The subcategory centred around innovative and purposeful pedagogical practices has received 45% of the references, indicating a strong desire for intentional and creative practices. In this regard, one of the participants stated:

“There, regarding the assessment aspect, there needs to be a modification. I still don't understand why teachers are reluctant to (...) all kinds of exploration, advancement, and the construction of students' knowledge with open-book assessments. For instance, presenting a case like the one you've just raised and, around it, initiating a discussion or debate. If knowledge and experiences are meant to be shared among everyone, to discuss common points, identify areas of differing thoughts, and establish consensus – which is the most challenging part... I mean, the focus should be on the knowledge-building processes, in utilising what is being generated through Artificial Intelligence”.

This passage illustrates the assumptions participants hold regarding the impact of traditional practices, particularly assessment, in education. To address this concern, they advocate for creativity coupled with purpose. Aligning intentionality with innovative practices reflects a deliberate desire to challenge existing paradigms of meaning-making in education and stimulate fresh approaches to thinking and action. From this viewpoint, participants propose that trainers not only consider how to assess students but also question whether the objectives and nature of what is being assessed remain static. These practices may initially disrupt conventional norms, but they hold the potential to spark meaningful shifts in the educational context, shifting from learning better to an overall experience of better learning.

4.3.4.2. Hands-on and Skill-based Training

The subcategory concerning hands-on and skill-based training received 40% of the references, indicating a strong emphasis on experiential training and the desire to provide opportunities for applying knowledge in real-world scenarios. This perspective is exemplified

by a participant who stated: “I believe that courses for professionals who are currently working should be very practical. They should not be awareness courses”. As highlighted earlier, insights gleaned from the cluster similarity analysis, illustrated in Figure 9, above, suggest a connection between these two subcategories and indicate that both approaches can address potential resistance and ease transitions by instilling confidence in participants regarding their existing level of awareness in their field.

4.3.4.3. Learning by Example

The topic of learning through examples arose in approximately 15% of discussions, highlighting the significance of practical illustrations and real-life scenarios, through which participants can gain insights into ethical dilemmas and develop skills to navigate them proficiently, as demonstrated by this quote: “The only mechanisms that come to mind for training in that area would be to show examples of what not to do so that they can learn how to use them....Teaching small cases tailored to the user and having them identify if they share that use and if they see potential risks in using it. And then, they are presented with half of the case, and the consequences are observed”. This subcategory, as indicated in Figure 9, above, is linked with the promotion of broadened ethical thinking and the update of knowledge and skills. It implies a recognition of the importance of examples, role-playing, and role-taking in stimulating perspective taking and the affective-cognitive dissonance that leads to the emergence of ethical dilemmas and to a deeper understanding of ethical decision-making processes. These examples serve as valuable learning tools by offering direction on which areas to prioritise for additional learning. This is especially important when considering the informed allocation of resources needed for continuing professional development.

4.4. Discussion

The findings of this research contribute to shaping the next steps in developing an AIED and education ethics training course for K-12 in-service and pre-service teacher programmes. This evolving framework aims to foster dialogic practices that allow teachers to reflect on the ethics of education in the presence of AI technologies. The discussion examines the findings concerning the research questions to guide the next steps in developing the training course. It integrates theoretical and conceptual elements from the current draft of the AI Competency Framework for Teachers (UNESCO, 2023c), emphasising how the findings align with ongoing work by experts in this field. Additionally, it interprets the results through the lens of the ik-model domains (Mouta et al., 2015). The ik-model builds upon the TPACK framework by

introducing diachronic dimensions to enhance the understanding of technological integration in learning environments. These dimensions focus on relationships and processes, adding depth to the technological and content domains. They provide a perspective on how relationships between different educational stakeholders are shaped and change, and how pedagogical processes are initiated, evolve, and adapt contextually to foster meaningful learning experiences. Additionally, participants' responses naturally encompassed the dimensions of the ik-model in their narratives, thereby demonstrating the model's relevance and applicability for discussing the results. The following discussion of each research question will draw on the findings to explore their practical implications for teacher training in AIED, ethics, and agency.

Regarding the first question, which concerns the structuring dimensions of training programmes on AIED, ethics, and agency, participants expressed a nuanced perspective on what should be implemented, moving away from a positivist approach to training design. In fact, teachers spontaneously covered all the ik-model's domains as relevant parts to be considered in such a training: (a) technological domain – AIED techniques and technologies; (b) content domain – ethical thinking, educational paradigms, teachers and students' roles, policies frameworks, best practices; (c) relational domain – roles and collaboration between key educational stakeholders; and (d) processes domain – dealing with resistance and transition, experiential practices, learning by example, self-directed learning.

Relational and processes dimensions are often sidelined in training programmes on AIED for education, which predominantly focus on content knowledge and pedagogical content knowledge. The categories addressed within the processes domain also demonstrate participants' proficiency in andragogical principles, possibly stemming from their roles as teacher trainers, indicating a deep understanding of the dynamics involved in designing and facilitating adult learning.

Regarding the second question, which focused on objectives, participants addressed the course's technological and content domains. The first implication for defining the course's aims is as follows: (1) consider it an ethical decision in itself to discuss AIED frameworks as either supporting tools or constraints to meaningful practices and pertinent innovation.

This concern and recommendation regarding teacher competency align with the UNESCO draft, which proposes the inclusion of “ethics of AI” as an asset. This asset encompasses an understanding of fundamental ethical principles related to AI, as well as participation in communities for the iterative development of institutional and societal regulatory environments, bridging the gap from understanding to active engagement. In these focus

groups, participants put emphasis on the importance of discussing AIED-oriented policy frameworks during training. The adoption of such frameworks is deemed suitable solely when enabling a comprehensive approach tailored to local interests and school microsystems.

As these frameworks are typically developed through extensive research, consultations, and expert input, with the aim of reflecting societal values, legal requirements, and best practices in the field, they can provide reassurance to teachers during transition stages. For instance, the “Beijing Consensus on Artificial Intelligence and Education” (UNESCO, 2019a), “The Ethical Framework for AI in Education” (The Institute for Ethical AI in Education, 2021), and the draft “AI Competency Framework for Teachers” (UNESCO, 2023c) can serve as valuable guidelines in this regard. These frameworks are scalable and allow for some contextualisation and progression, serving as an initial step in addressing the ethical challenges posed by AIED and ensuring its alignment with pedagogical principles.

On the one hand, the decision to follow these guidelines in the context of AIED can be seen as an ethical decision in itself. In fact, it can help reduce excessive reliance on singular experiences and personal biases. On the other hand, it may prompt considerations as to whether these frameworks might unintentionally impede meaningful innovation, creativity, or more dynamic and fully context-driven approaches. Nevertheless, exploring these frameworks may provide a basis for discussing a set of principles, considering the perspectives and interests of various stakeholders.

Regarding the third question about particular considerations for designing a training programme in this field, participants mainly reflected on the course processes domain. The second implication for preparing this training is as follows: (2) provide opportunities for addressing teachers' concerns, needs, and uncertainty during transition processes throughout course's design and implementation phases.

The utopian and dystopian narratives surrounding AI, the current knowledge, and conditions for its integration, combined with past experiences of integrating digital technologies into learning environments, make it key to address the cognitive, affective, and behavioural aspects related to AI integration during training. In fact, participants underscored the necessity of assisting teachers in handling resistance and facilitating transitions to foster positive training effects. Optimising attention to change, uncertainty and transition processes can also be strategically used to measure the course's impact on educators' cognitive, affective, and behavioural dimensions of attitude.

Cukurova et al. (2023) devised a reliable instrument for assessing the comprehensive factors that impact teachers' adoption of adaptive learning platforms in educational settings. According to their study, it is imperative that these technologies do not impose an additional workload on teachers, requiring them to switch between different tools, or navigate various pedagogical practices during implementation, while ensuring the minimisation of ethical concerns. Certain frameworks concentrate on the emotional and cognitive phases that teachers undergo when faced with change, including endings, neutral zones, and new beginnings (Bridges, 1991; Hall & Hord, 1987). Frameworks like The Diffusion of Innovations Theory (Rogers, 2003) and Change Management Models highlight the importance of communication, social networks, stakeholder involvement, leadership support, and the perceived advantages of the innovation in influencing teachers' acceptance and positive implementation. The selection and combination of frameworks may vary depending on the context, nature of the change, and the specific requirements of teachers and educational institutions, recognising and exploring transition pathways throughout the implementation phase.

Understanding uncertainty as the movement through which agency can be restored is crucial for enabling educational actors to lead implementation. The process of integrating new technologies is inherently non-linear and open-ended. Within this space, imagination, experimentation, and the recombination of old and new elements can thrive, fostering meaningful decision-making and agency. Becoming is not only about moving towards new possibilities but also about letting go of what will no longer be or other alternative modes of being. Understanding and incorporating this dynamic movement into the strategic training axis is indeed particularly rich for engaging educational actors in practices of meaning and collective agency.

In relation to the fourth research question, focused on implementation strategies, participants covered the course technological, relational and processes domains. With respect to the technological domain, they proposed the following: (3) use AI not only to support course design, delivery, and assessment but also to redefine and concretise AI tools within dialogical practices.

Participants highlighted the need for direct access to AI tools in order to comprehend their techniques and subsequently assess their potential and pitfalls. They suggest complementing theoretical discussions of ethical concerns with practical training using various AI resources to support course delivery and assessment. Methods for implementation include offering guidance in course forums through AI teaching assistants, fostering collaboration with AI

facilitators, and evaluating course activities using learning analytics. Additionally, GenAI can be trained to provide models for teacher assistance, incorporating school and national educational guidelines. This includes addressing several topics which can be chosen by the trainees, such as managing challenging classroom behaviours or offering pedagogical tips to improve student engagement with a given subject. The suggestion aligns with the competency outlined in the UNESCO draft, which advocates for teachers to progress from understanding fundamental AI concepts and functions to comprehensively mastering the adaptation of AI tools to develop solutions tailored to their own educational contexts.

This strategy should not be perceived solely through an instrumental lens. It goes beyond mere tool usage to understanding its functionality, enhancing learning experiences, and addressing the ethical challenges these resources may pose. A perspective aligned with Bardone and colleagues' (2023) concept of tinkering would be more dynamic, exploratory, and diachronic for all educational actors. Through this approach, they may participate in the process of concretisation of these tools, as defined by Simondon (2008). A new iteration of the technological object would facilitate the concretisation of its overabundant functions, potentially leading to new technological possibilities within a technogeographical space situated between nature and technics. This process involves the integration of the technological object with its surroundings, demonstrating a level of internal compatibility that produces its external adaptability. Importantly, such an approach holds dual significance: it pertains to the concretisation of pedagogical practices themselves, reshaping evolving roles among all actors; and it also transforms power dynamics in shaping narratives surrounding AI for education.

The relational domain was also covered with respect to the fourth research question on implementation strategies, with participants advocating for: (4) fostering individual and collective agency on AIED ethical issues through informal and self-directed learning channels.

In addition to engaging in critical reflection, prioritising training initiatives that enhance collective agency is crucial. In fact, teachers are expected to translate insights into meaningful actions aligned with the specific demands of their educational environments. So, another insight derived from this study is that any training fostering a reflective and collective agency journey in AIED and education ethics must persist through various channels, including professional learning communities, communities of practice, online professional learning networks, teacher associations and unions, faculty meetings and committees, school leadership teams, and action research projects. The UNESCO draft on AI for Professional Development also encompasses this perspective, viewing AI as a facilitator of learning for teachers. It

emphasises the importance of teachers critically adjusting or altering AI tools to better meet the evolving needs of professional communities and contexts. This makes it clear that in the context of AI, professional learning is key, namely through informal communities and self-directed endeavours.

Collective agency can also be fostered in educational cultures that allow for contextualisation and autonomy, through curriculum flexibility, adapting to the diverse needs of student populations and broader community requirements in feasible timeframes (OECD, 2018). This work can support the building of alliances that advocate for the pertinence and sustainability of bottom-up decisions. According to Crary (2022), these bottom-up strategies are critical to collectively reshape the landscape of technology in education, contributing to a social movement that brings about collective benefits and shared risks.

This assertion gains heightened significance in light of the prospect of AIED automating pedagogical practices that lack contextual meaningfulness, coupled with the risk of potentially disempowering teachers in the long run. In a more positive approach, by incorporating AI techniques thoughtfully and ethically, educators can amplify collective agency. This can be achieved by leveraging data in a contextualised and insightful manner to inform decision-making and collaborative planning processes. Moreover, collective agency can be fostered when educators are not only motivated to ethically integrate AI into educational practices (educating with AI) but also to develop students' understanding and critical thinking about AI itself (education for AI). Initiatives such as the European Commission's Code Week (European Schoolnet, 2021) and the MIT RAISE programmes (RAISE, 2024) exemplify this approach by supporting or providing K-12 curricula to expose students to AI at an early age. As Gertz (2016) argues, technology not only mediates our nihilism but also mediates our responsibility, serving as a means to prompt reflection on the human values that shape and nurture AI today.

A fifth implication for the implementation dimension regarding the process's domain has to do with the need to: (5) employ narratives to contextualise these AIED technologies, considering ecosystemic factors.

The subcategories that emerged from participants' discussions, focusing on dealing with resistance and facilitating change, are associated with the implementation of training through innovative and purposeful andragogic strategies. In line with this objective, the ethical scenarios developed in the initial stages of this educational design research were seen by participants as valuable for supporting knowledge acquisition and fostering consensus. As a result, another recommendation is to use narratives to contextualise AIED technologies by

considering ecosystemic factors, implementation strategies, and societal dynamics. In the design of the training course, these dilemmas will be revisited to integrate updated examples that reflect recent advancements in AIED while addressing the same ethical concerns. The resulting scenarios, blending both dystopian and utopian elements, will serve as heuristic tools for reimagining education through ethical and agentic perspectives. This approach can support the development of the UNESCO-recommended competency centred on a human-centred mindset, which includes recognising opportunities and risks, understanding societal impacts, and engaging in transformative action to address emerging challenges.

Another implication in terms of implementation strategies related to the process's domain concerns the following aspect: (6) incorporate meaningful experiential learning as a means to stimulate dialogic ethics in AIED-related discussions.

The broad acceptance of GenAI has sparked concerns among educators across secondary to tertiary education levels. Despite LLMs not being crafted for educational settings, it has raised worries, particularly in scenarios where it can complete homework for students. Nevertheless, these platforms offer an opportunity to reassess and redefine the learning processes and corresponding pedagogical methods, potentially leading to more meaningful outcomes. This can entail students interacting with these platforms and then assuming the role of teachers to correct the AI responses. They may also position themselves towards the information, taking on roles such as researcher or opponent. Additionally, involving other community actors in the learning process to diversify information sources and fostering group and project-based learning are viable approaches. Indeed, teachers are the ones who understand the microgenesis of learning, the progression that unfolds throughout a learning session.

Using Large Language Models can be an opportunity to challenge tasks of assimilation, accommodation, and evocation of knowledge, while creating agentic opportunities to deal with this technological novelty within school. This can be purposefully achieved by engaging students in diverse learning processes, such as perspective-taking and distributed agency (from both humans and machines), introspection (focusing on the uniquely human process of self-reflectiveness), imagination (inspired by the vast possibilities enabled by socio-technical entanglements), affective-cognitive dissonance (introducing dilemmas brought by the interaction with AI that stimulate ethical reasoning), and exploration of engagement with action through collective agency (teachers share decisions with their students about how to learn and be assessed with GenAI).

Another suggestion arising from this study is that training is expected to integrate meaningful experiential learning as a means to fully embrace dialogic ethics in learning settings. This can be maximised through activities such as role-playing and role-taking, enabling educators to assess and refine their strategies directly in their classrooms during the training period. Subsequently, they can return to the secure training space to evaluate their practices alongside peers and trainers. This also enables creating a platform for debate on ethics among educators, between educators and trainers, among students, and between teachers and students, ensuring that inter-generational perspectives on AI challenges are not only debated but also infused into subjects, potentially leading to valuable outcomes.

This strategy directly aligns with the UNESCO suggested competency of guiding teachers to progress from identifying the potential benefits of using AI systems to critically evaluating AI in pedagogical practice, ultimately moving towards transformative pedagogy.

4.5. Conclusion

Until recently, the ethical foundations of education in the context of AIED have not been thoroughly explored, and the corporate narrative continues to dominate public discourse on AI for education. Debates on this topic seem to be especially pertinent as we live in a world of technique, as Ellul (1964) defined it: a world embedded in the pursuit of methods derived through rational means and achieving optimal effectiveness across all areas of human endeavour, where the humanities have lost their pace. AI technologies have become more pervasive, and major multinational corporations, in cooperation with at least one local government, are attempting to influence educational practices. Non-governmental curricula are already incorporating AI learning outcomes at middle and high school levels (UNESCO, 2022).

Various research findings indicate that AI curricula only marginally address AI ethics, and there is a shortage of opportunities for teacher preparation to tackle contemporary challenges in this area. To bridge this gap and capture diverse stakeholders' perspectives, this study focused on insights on education ethics in the AI era stemming from teachers and teacher trainers focus group discussions. Envisioning the subsequent step as the development of a training course in this area, this research phase contributed significant insights on course objectives, implementation strategies, and design stages.

By prioritising transversal assets beyond AI techniques and technologies, participants highlighted the significance of continuing professional development as a platform where teachers actively contribute to shaping their professional growth and standing, safeguarding their agency against potential encroachments by tools or policies. They were also recognised

as key figures for challenging established knowledge in their field, while also questioning the very nature of epistemology and the conditions under which it is nurtured. Participants articulated their perspectives regarding the revitalisation of educational paradigms and teachers' roles, not only as gatekeepers of high-quality content, but as vital agents of thought and action in the face of the polysemic landscape of AIED technologies.

While education remains subject to state regulation, opportunities exist for elevating the status of teachers across diverse educational systems and impacting public policy development, in the interplay between technology, pedagogy, and society. Additionally, increased investment in professional learning and participatory policy reforms, leading to policies that are more comprehensive, has the potential to significantly enrich the outcomes of national investments in AI for the educational sector. In this context, the present research has sought to provide educators with an avenue to explore relevant educational possibilities and to assume proactive roles as agents of meaningful transitions in their learning environments.

Participants' insights shed light on the importance of a broad ethical exercise and strategic collaboration among educational actors. They also revealed a nuanced understanding of the cognitive and emotional challenges teachers face during the integration of AI technologies, where uncertainty plays a key role. This calls for a comprehensive consideration of meaning indeterminacy and underscores the importance of carefully considering signification processes as they unfold in pedagogical practices. While the recommendations arising from the study underline the critical role of teacher autonomy in professional development, especially in the era of AIED, realising this objective can be a complex undertaking due to constraints related to limited resources, hierarchical organisational structures, and cultural factors that may hinder teachers' capacity to exercise professional autonomy.

This clearly exemplifies the complex interplay between structure and agency to which the current research and further training aim to offer actionable insights. The participatory methodology used in designing research and teacher training situates this study within the broader discourse of dialogic ethics. This approach is sought to be responsive to the challenges posed not only by AI technologies, but by a world driven by technique. It's only when uncovering the technique layer that is overarching education too, that a consistent dialogue on ethics for education can be established, going beyond mere debates on the ethics of AI.

Finally, while this research advances past studies that merely explore the ethical challenges of AIED by offering concrete recommendations for training programmes", targeting teachers in this domain, it is necessary to acknowledge its challenges and limitations. One challenge

arises from the disconnect between the primary examples of the scenarios toolkit and the rapid advancements in AI technologies. This underscores the imperative of integrating more relevant examples that align with the current state of the art, thereby enhancing the applicability of the toolkit for the upcoming training. The scenarios will be used with these updated examples, as they depict ethical dilemmas that remain relevant today. Other constraints in the study emerge from concerns about generalizability, given that the insights derive from a particular participant group, even though efforts were made to encompass participants from various geographical backgrounds. Moreover, the study primarily focused on providing recommendations for teacher training in AIED, agency, and dialogic education ethics, without exploring the actual implementation and effects of such a training programme.

This research calls for the next step, which pertains to an examination of its impact on agentic decision-making regarding the use of AIED in educational settings. Additionally, cultural and societal factors can significantly influence the level of autonomy granted to teachers, complicating efforts to introduce alternatives within existing structures, reform those structures, or challenge and resist them, as advocated by Wright (2010). In certain educational contexts, traditional norms and hierarchical power dynamics impose strict adherence to prescribed curricula and pedagogical methods, leaving limited room for individual autonomy or innovation.

The lack of professional autonomy restricts teachers' ability to shape their own professional growth and may hamper their capacity to read their learning ecosystems in light of evolving educational demands, practices, and pressures. As a result, these contexts are less likely to play a critical role in challenging the predominant dominance of technique, reaffirming educational ethics and leveraging the potential of AIED technologies when integrated meaningfully. Consequently, they are more prone to encounter ethical challenges associated with their use.

CHAPTER FIVE

Empirical Research III: The Shifting Nature of Agency in Education with AI

The findings from the previous research, which involved focus groups with teacher trainers on the ethical considerations and pedagogical strategies for AI in education, helped identify key areas to address in teacher training programmes. This phase primarily focused on gathering general insights into the ethical implications of AI, as well as the elements to consider when designing professional development for teachers in the context of AI integration. As the research evolved, however, it became clear that a more targeted exploration of the concept of agency, especially in relation to AI's impact on education, was necessary. To refine this focus, a secondary analysis of the focus group data was conducted, specifically examining how agency was being understood and discussed by educators. This analysis revealed important gaps in how agency, both at the individual and collective levels, was being framed in the context of AI-integrated education.

The refined focus on agency led to a more comprehensive understanding of how teacher training could not only incorporate ethical dimensions of AI but also support educators in maintaining their sense of agency amidst technological advancements. These findings are now crucial in informing the design of the teacher training programme that is the central goal of this dissertation research project.

To further explore these insights, this chapter delves into how AI technologies in education influence human agency at subjective, intersubjective, and collective levels. Through qualitative analysis of discussions with teachers and professors, it examines how educators perceive and negotiate agency in relation to AI, shedding light on both its potential and its risks. While teachers demonstrate a strong awareness of students' individual agency, they may overlook the broader intersubjective implications of AI, potentially reinforcing narratives that diminish the role of educators. The findings contribute to the development of professional training frameworks and highlight the need for further research on sustaining agency within AI-integrated educational environments.

5.1. Introduction

In an age of increasing automation and technological reliance, the dynamics of agency are evolving. Interactions with socio-technical systems, especially Artificial Intelligence (AI), present new challenges and opportunities for individual and collective agency. In response, there is a pressing need to foster forms of agency that are not only personal but also relational and collectively distributed, grounded in individuals' capacity for internal conversation, which enables them to reflect on and shape their engagements (Archer, 2014), and in their

interdependence with others and with the natural world in which we live and learn (UNESCO, 2021).

More recently, the interplay between human and non-human entities has challenged the idea of agency as an individual property. These non-human entities can include machines, algorithms, technologies, animals, and even objects or environmental factors. Thus, debates have been exploring the extent to which agency is distributed across a network of actors, both human and non-human (Deleuze & Guattari, 1980; Haraway, 1985, 2016; Latour, 2005; 2018). Latour (2018), in particular, explores how the alignment (or misalignment) between macro-level planetary concerns, meso-level institutions, and micro-level everyday actions is entangled, calling for multiple levels of agency. These discussions hold particular significance in helping remember and actively engage with our shared agency within this trading zone. As Latour (2014, 2018) suggests, there are two main values to this approach: first, it allows us to identify potential actants before they become actors; second, it fosters a connection between citizens and their concerns, linking them to their surroundings and shared environment, gradually shaping a common world to be collectively built.

Building on these broader debates on distributed and relational agency, this chapter turns to the field of education as a critical space for exploring how agency is experienced, challenged, and shaped in the age of AI. Education is not only where individual autonomy is cultivated, but also where collective responsibilities are negotiated.

The present study primarily focuses on the findings from the focus group study on the theme of agency, it is situated within this wider EDR framework, which also engages with the regulatory and policy implications of AI in education. In particular, the findings highlight the absence of clear agency frameworks to guide the ethical adoption of AI technologies, which is an issue that will be further explored through the subsequent design, implementation, and evaluation stages of the training project.

To explore these questions in greater depth, the next subsections of the introduction elaborate on two core foundations: first, by tracing how the concept of human agency has evolved across philosophical, psychological, and sociological traditions, and second, by exploring how agency has been framed and fostered within educational contexts. These discussions then lead into a framework for examining how AI affects agency on three interconnected levels (subjective, intersubjective, and collective), which structure both the theoretical framing and the empirical analysis presented in the paper.

5.1.1. The Concept of Human Agency

The concept of human agency, which can be largely defined by the capacity to act independently and make choices that influence one's environment, holds profound significance for human and societal development (Bandura, 2001). As a temporally embedded process, shaped by the past, oriented towards future possibilities, and responsive to present contingencies, it allows individuals to explore and reconfigure their relationship to social structures, which are themselves temporal and relational. Social actors continuously shift their orientations within overlapping temporal contexts, enabling them to engage in flexible, situated action (Emirbayer & Mische, 1998). This notion extends beyond mere individual decision-making, encapsulating social, cultural, and technological dimensions that shape collective experiences and behaviours. In a world characterised by rapid technological advancements and evolving societal norms, understanding agency has become increasingly critical, particularly within the realms of education. The ability to assert agency is not only foundational for personal development but also key for collectives, fostering responsible citizenship, engaged practices, and contextual-relevant frameworks for action within societies (UNESCO, 2021).

Debates surrounding agency have long focused on the nature of free will, moral responsibility, and the tension between individual autonomy and social structures. During the Enlightenment, thinkers framed human freedom as a response to deterministic views rooted in religious morality. People were seen as wilful beings, capable of independent judgment and autonomous action, embodying the Enlightenment ideal of rational individuals with moral agency. Philosophical perspectives on agency evolved to include both normative, non-rational action (influenced by Kantian ethics) and rational, instrumental action, later developed by American pragmatism and Continental phenomenology. The former reflects a moral will, where actors (individually or collectively) engage in moral action (Kant, 2005), while the latter frames agency as rational, goal-oriented behaviour to achieve personal interests or fulfil material needs (Marx & Engels, 1976).

Throughout the early modern period, questions of personal identity and consciousness began intersecting with concepts of agency, with existentialists like Sartre highlighting the importance of individual responsibility in one's choices (Sartre, 1943). By the 20th century, sociological thought had expanded to explore how agency is deeply influenced by historical and social contexts, investigating its embeddedness within socio-economic structures and power dynamics (Foucault, 1975). This exploration led to the development of theories aimed at bridging the structure/agency dualism. Giddens (1991) proposed that social structures

emerge from repeated patterns of human action and are simultaneously shaped by, and constraining of, that activity, thus placing structure as both a product and a constraint of agency. Bourdieu (1990) similarly argued that habitual patterns, or practices, become structure over time, guiding future actions. However, reactions to Giddens's theory emphasised that agency is not merely the repetition of patterned action; it also encompasses the ability to break with established rules, reflecting thought, reflexivity, and creativity (Emirbayer & Mische, 1998).

The rise of phenomenology introduced discussions about self-ownership and the pre-reflective experiences of causing actions, deepening the understanding of how individuals experience agency. Within the framework of social cognitive theory, Bandura (2006) offers one of the most comprehensive accounts of human agency and remains a central reference in educational and developmental psychology. His theory acknowledges the contribution of genetic heritage in providing the neural structures and mechanisms necessary for distinctively human traits like generative thinking, symbolic communication, intentionality, foresight, self-regulation, and reflective self-awareness. It rejects the dualism between individual agency and disembodied social structures, positing that humans intentionally influence their functioning and circumstances. Bandura (2001, 2006) identifies three forms of agency: individual, proxy, and collective, each playing a distinct role in how individuals and groups interact with the world and shape their existence. Individual agency is characterised by a person's ability to exercise control over their own actions. For Bandura, self-reflectiveness, the metacognitive ability to observe and evaluate one's thoughts and actions, adjusting strategies as needed to accomplish goals, is the most distinctive feature of human agency. This capacity for reflection enables individuals to align their actions with moral standards, making agency not just a matter of execution but also of ethical and moral self-regulation.

Beyond individual agency, this author is also one of the few major theorists to develop the concept of proxy agency, which occurs when a person depends on others who have the necessary resources, skills, or authority to act on their behalf. While this form of agency can be crucial for achieving certain outcomes, it also creates a reliance on external agents, potentially leading to a diffusion of personal responsibility and affecting moral accountability. This is particularly pertinent in the context of AI and education. Contemporary uses of AI systems, whether through adaptive learning platforms, generative tools, or automated feedback mechanisms, frequently involve students and teachers delegating cognitive or pedagogical tasks to technological agents. Bandura's conceptualisation helps unpack the psychological implications of such delegation, including how it affects motivation, ownership, and

accountability. This dimension of proxy agency is still underdeveloped in many contemporary theories of agency, which tend to focus either on individual autonomy or distributed, collective enactments without addressing the intermediary or surrogate roles played by technologies.

On the other hand, collective agency arises when individuals engage together to achieve shared goals. This form of agency is distinguished by not merely being the sum of individual contributions but emerging from the synergistic interactions among group members, where each plays a key role. It is especially significant in social and educational settings, where a sense of belonging and shared responsibility fosters meaningful endeavours. Bandura argues that these three forms of agency often intermingle.

5.1.2. Human Agency and Education

The 1960s and 1970s saw a resurgence of the New School movement, originating at the start of the century, which advocated for greater democratic involvement in education. This shift was, in part, a reaction to the impacts of industrialisation, which had turned traditional schooling into a standardised and conformity-driven system. Influenced by socialist and anarchist ideals, New Schools aimed to create spaces where students could develop freely and critically engage with societal norms and structures. It wasn't until the 1990s that education fully pivoted toward fostering learner autonomy. This shift focused on students' responsibility for their own learning (Jones, 2007) and was influenced by Carl Rogers's humanistic approach (Rogers, 1983). In terms of learning, agency is seen as both a process and an outcome.

According to Bandura (2001), learners with agency can bring about things intentionally and be responsible for their self-development, adaptation, and self-renewal. At the individual level, a robust sense of agency enhances motivation, enabling people to respond to challenges with resilience. Research indicates that high self-efficacy correlates with sustained engagement and achievement across various contexts (Vera et al., 2014). Initiatives prioritising the cultivation of agency encourage students to embrace autonomy, build confidence, and engage in self-directed learning (Davis-Kean, 2005).

While crucial for individual achievement, the field of education extends far beyond the outcomes of personal agency. Educational environments are privileged spaces for cultivating the political foundation of democracies (Dewey, 1916) through collective agency, enabling students not only to take ownership of their learning but also to engage in dialogic practices that reflect the interests, dynamics, and unique experiences of their institutions. This interplay between individual and collective agency is also demonstrated by Zumeta et al. (2016), who found that shared emotional experiences and group identification in collective activities

enhance team achievement, highlighting the importance of community in nurturing agency. Collaborative learning experiences in educational settings foster not only academic achievement but also social connectedness and shared responsibility – critical components for building cohesive societies (OECD, 2012). Experiencing opportunities for agency within educational contexts enables students to recognise their role within a larger community, fostering a sense of belonging and collective participation. In this regard, Dewey's (1916) advocacy for an education based on principles of experimental democracy gains renewed significance.

5.1.3. Education and AI

Contemporary research highlights the potential of Artificial Intelligence in Education (AIED) to personalise learning experiences, streamline administrative processes, and offer tailored pedagogical support (Mousavinasab et al., 2018; Wang et al., 2023). Intelligent tutoring systems and adaptive learning platforms can now analyse large datasets on students' strengths, weaknesses, preferences, and learning styles, creating customised learning pathways that adjust dynamically to individual needs (Luckin et al., 2016; Miao et al., 2021). More recently, greater attention has been given to the fact that students are neurodiverse, bringing with them a range of assets derived from their experiences at home, within their communities, and across their cultural backgrounds. In response to this diversity, researchers have recognised that AI systems, with their capacity to detect patterns in complex datasets and to generate customised content, have the potential to extend beyond the most common learning trajectories and to offer personalised educational experiences that would be difficult for individuals to produce unaided (U.S. Department of Education, Office of Educational Technology [USDOE-OET], 2023). Moreover, embedded adaptive assessments facilitate real-time feedback and allow for continuous monitoring of learners' progress, engagement, and emotional states, thereby supporting more responsive teaching practices (Luckin et al., 2016; USDOE-OET, 2023). AI-driven tutoring support, through interactive simulations and virtual dialogues, can further enhance comprehension by delivering targeted feedback and gradually refining instructional strategies through ongoing learner interaction (Khosravi et al., 2022; Wang et al., 2023).

As AIED systems continue to collect and analyse performance data, they increasingly enable early identification of students at risk of falling behind, thus allowing for timely pedagogical interventions. Beyond supporting individual students, the use of AI in school administration is often seen as a way to strengthen collective agency, enabling institutions to

make more informed decisions, better allocate resources, and respond more thoughtfully to the needs of their communities.

Although these technological advancements are often celebrated for their efficiency and their potential to enhance learning and education quality, they also carry the risk of diminishing individuals' sense of control and autonomy (Moore, 2016). Studies indicate that the design of human-machine interfaces significantly impacts users' perceived agency, with greater agency associated with human-like interactions (Berberian et al., 2012; Obhi & Hall, 2011). This necessitates a critical examination of how we perceive and integrate these socio-technical systems into educational practices, ensuring that they serve as tools that provoke our collective engagement towards agency. Furthermore, the ethical implications of agency extend beyond personal development to influence societal norms and values. A diminished sense of agency can lead to moral detachment and detached behaviour, affecting how individuals attribute responsibility for their actions (Haggard & Tsakiris, 2009). In educational contexts, sense of agency has also a role in equipping students with the ethical reasoning skills necessary to question complex societal challenges, promoting accountability and moral engagement with life (Bandura et al., 1999).

As educators (and society more broadly) continue to engage with systems dominated by techniques that prioritise rationality and efficiency (Ellul, 1964), which are deeply embedded in the secular liturgies of school practices, their stance is particularly necessary for reviving the spirit that animates schools and preserving their organic capacity to respond to their own experiences. Practically, this means that teachers are expected to play a meaningful role in shaping the policies they are tasked with enacting, alongside families and community members, who are critical partners in educational development (European Union, 2018). This need for agency is especially relevant in the context of technological advances and narratives surrounding AI's autonomy. Pasquinelli (2023) argues that labour is at the root of the historical development of AI, which is shaped by a composite of dynamics reflecting social relations and power structures, where the current control is exerted by high-tech companies. This understanding highlights why agency cannot be separated from collective dynamics and historical contexts. Teacher agency, therefore, is also the result of interactions between personal capabilities and the environmental factors influencing their actions (Priestley et al., 2015). In an era of high automation, educators have a role in reclaiming the significance of agency and challenging the contested narratives surrounding automation and work. As AI evolves into a regime that replaces an episteme of causation with one of automated correlations

(Pasquinelli, 2023), educators are called to cultivate spaces where collective knowledge and agency can flourish. This perspective has direct implications for the way teacher professional development is conceived and designed, as well as for a comprehensive understanding of human and societal development processes, positioning collective agency as one of the ultimate goals of education.

So, in the context of education, three levels of concern shape our understanding of how AI systems could impact agency: the subjective, intersubjective, and collective levels. On the subjective level, agency is explored through dimensions that examine how AI may affect human experiences in areas like decision-making, moral judgment, deliberate action, achievement, individuation, and project ownership. Intersubjectively, agency is negotiated among peers, between teachers and students, and among other stakeholders, including families. Here, power dynamics, teachers' status, sense of belonging, and scaffolding experiences come into play. Collectively, agency is examined through situational freedoms within institutional policies, collective deliberation and achievement, political development, and democratic practices.

5.2. Methodology

5.2.1. Research Question

This study seeks to understand how teachers conceptualise the transformation of human agency in relation to AI in education, drawing on the theoretical and methodological framework of dialogic ethics. The research is based on a qualitative analysis of data collected from the focus groups presented in the previous chapter (see Chapter 4). While the initial purpose of the study was to explore collective perspectives on what a training programme on such themes should include, a significant portion of the generated text – though not directly focused on that theme – was closely related to socio-technical AI systems and education.

This prompted a new approach to data analysis, leading to a revised goal deemed highly significant for understanding the role of agency at the intersection of these technologies and pedagogical practices, as perceived by one of the key educational actors: the teachers. The research question for this phase was defined as follows: How do teachers conceptualise the transformation of human agency in relation to AI in education on subjective, intersubjective, and collective levels?

5.2.2. Participants, Implementation, and Data Analysis

As described in the previous chapter, participants were selected based on their experience in K-12 teacher education and their proficiency in Spanish, the primary language of the research centre. Initially, convenience sampling was used, followed by snowball sampling to ensure cultural diversity. Out of 40 invitations, 19 educators from five countries (Colombia, El Salvador, Portugal, Mexico, Spain) participated, all of whom held roles in pre-service or in-service teacher education. The study involved a balanced representation of males (N=10) and females (N=9), with a mean age of approximately 45 years. Among them were professionals in various roles, including a Director of Digital Transformation in Higher Education, several associate professors, and former K-12 teachers who held significant positions in the Ministry of Education's Master Plans. The collective expertise of these participants spanned various aspects of education, including research methods, teacher training, and the integration of information technology in educational settings, reflecting a strong emphasis on the intersection of education and technology.

In 2022, they engaged in discussions of hypothetical scenarios – dystopian perspectives on AI use in education derived from a Delphi method expert group – based on earlier research (Mouta et al., 2023) across four sessions to ensure data saturation. Conducted virtually via Zoom due to pandemic constraints, these sessions gathered a range of perspectives that contributed to the programme's development. Sessions were recorded, transcribed using AI software, and verified by two researchers. This phase of the analysis combined deductive and inductive methods, identifying themes based on existing literature, earlier project phases, and insights from raw data. As the purpose for this phase of the study is to understand teachers' perspectives on how the sense of agency can be influenced or transformed by the use of AI systems in education, the entirety of the data was once again explored through different search strings, namely: "agency", "autonomy", "author", "authorship", "capacity for action", "control", "decision-making", "independence", "initiative", "personal sovereignty", "self-determination", and "self-governance".

This approach emerged from the observation that participants often discussed the topic of human agency using various related terms, which, although not explicitly labelled as "agency", resonated within the same conceptual universe. This practice of broadening the scope to include associated concepts ensured a comprehensive understanding of how teachers perceive agency. The final selection of categories was determined through two reliability tests conducted with a

two-month interval between coding stages, as well as an inter-rater analysis performed by two researchers who compared the data directly (in the second stage) and reached a consensus.

5.3. Results

This section outlines the main findings from the focus group discussions, addressing the research question that guided this phase of the project: How do teachers conceptualise the transformation of the sense of agency in relation to AI in education on subjective, intersubjective, and collective levels?

The results are presented in a structured table format that delineates the findings across different levels of agency (Table 17, below). Each row of the table is categorised by (1) “Level” (subjective, intersubjective, collective), highlighting the context in which teachers discussed agency. Within each level, specific (2) “Dimensions” are outlined, reflecting the key aspects of agency that emerged during discussions. Accompanying each dimension is a (3) “Description”, which contextualises the findings and provides a concise overview of the issues identified. Furthermore, (4) “Examples” from the focus group discussions are included to illustrate the points made within the descriptions.

Table 17

Framework for Agency in Education: Teachers' Perspectives on AI and Human Agency

LEVEL	DIMENSION	DESCRIPTION	EXAMPLES
SUBJECTIVE	DECISION-MAKING PROCESSES	<p>Critical Reasoning AI-driven decisions may undermine students' ability to reason about personally significant variables.</p>	<ul style="list-style-type: none"> • “When AI decides what is best for the student... we’re taking away the student’s ability to decide what they want to learn, or how they want to approach a topic”. • “it makes the student unaware of why a certain personalisation process is beneficial. There is an entire external system that is managing their personalisation, which means that instead of fostering autonomy, what is actually being encouraged is heteronomy”.
		<p>Power of Structure Influence of social, cultural, and environmental factors that shape and may constrain individuals' choices based on available resources and visible opportunities, underscoring the need to critically assess whether technology and AI can expand or restrict these options.</p>	<p>"the reality is that the decisions we make are constrained by our surroundings, by what we can see. We choose to study from the options available in our community. (...) we shouldn't assume from the outset that technology or AI will restrict that freedom. Instead, we need to assess whether we truly have those freedoms today”.</p>
		<p>Reduction of Diversity and Circumscription of Choices AI systems that rely on pre-defined learning paths may restrict students' ability to explore a diverse range of possibilities.</p>	<p>“I think there's a loss of autonomy because if the student is only able to choose from a list of options... they're just choosing what was already decided”.</p>
INDIVIDUATION		<p>Differentiation and Authorship AI systems are expected to support the educational community exploring the diverse developmental trajectories and unique personal expressions of each student (e.g., learning pace, behavioural repertoire, specific manifestations) rather than enforcing standardised measures.</p>	<ul style="list-style-type: none"> • “This individual vision of each student and their needs contrasts with the generic view that ultimately states we don't want everyone to be the same; we need to delve deeper into each person being distinct and unique. This type of system that attempts to homogenise and make everyone equal ends up breaking down the personal characteristics of each individual”. • "what we do is help the student, year after year and day by day, develop their capacity to take ownership of the condition of being the author of their own life. They must create their life, but that is not easy. It is definitely not easy”.
FORESIGHT, SELF-REGULATION AND SELF-REFLECTIVENESS		<p>Agentic Properties Qualities that enable individuals to take ownership of their lives and decisions, enabling individuals to actively shape their experiences and achieve their goals.</p>	<p>“The ability to take ownership of one’s own life (...) the day my time runs out, my life runs out. Autonomy is about development (...) about beginning to give the student time so that they can take ownership of their time, envision their future, plan ahead, and be capable of committing their ability to act over time”</p>
MORAL DEVELOPMENT		<p>Moral Disengagement</p>	<p>“efforts are being made to incorporate Artificial Intelligence into these autonomous cars (...) Artificial Intelligence can reduce our capacity for</p>

			Progressive detachment from responsibility and erosion of ethical standards occurs when individuals justify or rationalise actions, often by attributing decision-making to external forces, such as technology or authority figures. This can diminish personal accountability and moral reasoning.	agency. For some people, this may be satisfying because making those decisions is certainly not easy, and there are individuals who may feel relieved not having to make those choices”.
INTER-SUBJECTIVE	PURPOSE AND VALUES	Teachers’ roles	Erosion of teachers' roles and purpose, as AI begins handling core pedagogical tasks, potentially turning their role into one of mere oversight.	“If AI starts taking over the planning and grading, then where does that leave us as teachers? We don’t want to just be there to supervise what the machine is doing”.
	AUTHORITY AND PROFESSIONAL STATUS	Proxy Agency	AI systems are viewed as supportive tools, with the expectation that teachers will maintain their role as proxy agents, preserving their professional autonomy and authority.	“AI can make suggestions, but it shouldn't make decisions. It should be the teacher who decides... because we know the students better than any algorithm”.
	MODELLING AND CONFIDENCE	Scaffolding	Supportive process through which teachers leverage their own experiences and socio-emotional skills to guide students in developing critical competencies.	“the experience and socio-emotional skills they have developed, along with the failures they have encountered (...) allow the teacher to guide students in discerning what information is good (...) not good, how to work in teams, how to communicate better, how to collaborate, and how to solve problems (...) help their students develop these socio-emotional skills”
COLLECTIVE	POLITICAL SPACE	Participation	Importance of a participatory approach where all stakeholders (e.g., students, teachers, parents) are involved in shaping AI's role in education.	“There needs to be a way for students, teachers, and even parents to give input on how AI is used... so we should all have a voice”.
		Distributed Decision-making	Call for collaborative decision-making in AI integration, with teachers advocating for a role in shaping its implementation, highlighting the need for democratic practices at an institutional level.	“We need to have a say in how AI is being used in our schools. It shouldn’t be a top-down decision”.
		Democracy and Rule of Law	Importance of participation, agency, and responsibility, rooted in the needs of the community. It promotes a society guided by established institutional frameworks that protect individual freedoms and enhance social welfare.	“In my view, I believe that two beautiful aspirations in a person's life can be striving to live in a democratic system and being guided by the factors established within the framework of institutions. This means aspiring to these two ideal goals that exist in most countries”.
	DISTRIBUTED AGENCY	Socio-technical Assemblages	Dynamics and shared agency between human and technological actants create a unique context for learning and reinforce the idea that the educational experience is co-constructed, shaped by the interplay of diverse influences.	"We could engage in dialogue with parents and teachers about viewing Artificial Intelligence and other aspects of information and communication technologies as added values to the learning processes and the construction of knowledge that have been developed. The very systems of life and the evolution of societies drive us to recognise that there must be different approaches to build learning”.

5.4. Discussion

This discussion section is organised into two main parts, drawing on the qualitative data presented in Table 17 above. The first explores how agency is perceived by teachers across individual, relational, and systemic dimensions of school life, and how AI socio-technical systems may interfere at each of these levels. Particular attention is given to how teachers interpret the evolving role of AI in relation to students' developmental processes, the professional identity and responsibilities of educators, and the democratic ethos of educational institutions. The second part presents an integrative analysis that identifies key trends across the three levels, highlighting patterns and interdependencies between forms of agency. These include tensions between automation and authorship, the evolving nature of proxy agency, and the need for moral scaffolding and shared responsibility.

5.4.1. Agency Levels

5.4.1.1. Subjective Level

Teachers expressed concerns about fundamental developmental processes, such as individuation and moral development, acknowledging their role as committed keepers of the educational mission, which extends far beyond the mere goal of content learning (Biesta, 2020). They also highlighted specific meta-cognitive and affective processes, which serve as qualitative differentiators for holistic development – a responsibility shared by all educational actors and formal institutions. These processes include the architecture of decision-making, encompassing critical reasoning, recognition of power structures, and the consideration of diverse options.

In terms of critical reasoning, teachers recognise that schools have a responsibility to create conditions that help students understand the personal variables influencing their choices throughout life; factors which, as many vocational psychologists have suggested (e.g., Guichard, 2009), directly impact vocational exploration, career self-determination, and work well-being. This task, which can be pedagogically designed, is significant for exploring personal critical development dimensions, such as interests, competencies, values, influences (both proximal and distal), and both positive and less positive personal bonds. This can be achieved through curriculum infusion activities, project-based learning, and role-taking experiences, where intentionality and systematicity are key. In this context, the role of AI systems is expected to be critically examined. AI systems that provide overly rapid feedback

and excessive automation may progressively be interpreted as authoritative figures that “hold the truth”, negatively impacting students’ sense of control. As research suggests, this perceived loss of control can be explained by high levels of automation and hierarchical structures within collaborative tasks. According to Gozli’s (2019) hierarchical model of agency, such a loss is not merely cognitive but begins at the foundational level of sensorimotor fluency, the capacity to skilfully and confidently engage with one’s environment. When this fluency is disrupted, students may experience agency as distant or inaccessible, leading to increased passivity and conformity. Consequently, such systems can reinforce passivity and conformity while distancing students from engaging with diverse significant others, potentially limiting the freshness of perspectives and contributing to emotional blunting.

Teachers also demonstrated sensitivity in recognising that social, cultural, and environmental factors shape personal agency (Larreamendy, 2011) at various instances and levels throughout different moments in life. This expanded awareness opens up the possibility of understanding agency as emerging from the interplay of heterogeneous elements, where objects and technologies do not merely assist human action but actively participate in shaping outcomes (Latour, 2005). In schools, this calls for a shift in how learning experiences are understood: rather than viewing agency as a purely human trait expressed through choice or autonomy, it becomes something co-constructed with the tools, spaces, and systems that surround learners. As the human-machine boundary becomes increasingly blurred (Haraway, 1985), and as material environments exert their own quiet pressures and affordances (Bennett, 2010), schools are called to move beyond simplistic binaries, such as active/passive or autonomous/dependent, and instead design for agency as something distributed, situated, and always in-the-making.

The power of structures, and their tendency to push the margins towards the centre, is reflected in education through the dominant discourse of efficacy that permeates the Big EdTech world and underpins AI applications in learning environments. This often places the burden of adoption and exploration of associated challenges on individuals, irrespective of the structural inequalities they may face, reinforcing a model shaped by minimal governance and free-market dynamics (Selwyn, 2016). Rather than fostering individual agency, such a distanced and individualised approach to what Zuboff (2019) describes as an unprecedented transformation in the exercise of power may lead to a diminished, if not detached, experience of agency, ultimately weakening the capacity to connect with others and build agency

collectively. Furthermore, as Watters (2021), Williamson (2022), and Saltman (2020) highlight, educational platforms and AI applications mirror Big Tech strategies by embedding practices of categorisation, optimisation, and impact measurement, promoting models of learning that reinforce privatisation, outsourcing, and labour exploitation. Such logics, while marketed as efficient, risk narrowing the space for variation, error, and alternative pedagogical approaches. This structural emphasis on predictability also affects how uncertainty is valued in emotional and cognitive development. Yet uncertainty can bolster creative and resilient approaches to dealing with circumstances and chance, and it may even contribute to the concretisation of technological tools in the process of their appropriation.

The role of education in supporting the development of students' individuation through differentiation authorship was also acknowledged, recognising a process in which one's voice, the diversity of experiences and actors, and the ability to articulate one's experiences through timely discursive connectors set the stage. How can a speculative argument on personalisation account for processes of meaning-making and signification? Adaptive learning systems may weaken or disconnect knowledge and learning from the cultural backgrounds, unique identities, experiences, and subconscious processes that are foundational to subjectivities. While educational sciences and psychology have moved beyond strict behavioural and cognitive models, the continued reliance of adaptive learning systems on these frameworks, as both inputs and outputs, raises questions about their adequacy in addressing the complexity of human learning.

The properties of agency, as described by Bandura (2001, 2006), were referred to. Once again, teachers demonstrated their awareness of the conditions that enable their students to become not only knowledgeable but also the architects of their own lives. They recognise that intentionality and planning are only possible when opportunities for self-regulation and self-reflectiveness are experienced. These factors are, in fact, critical ingredients for learning and meaningfully exploring the world – enabling individuals to act and integrate through processes of signification within meaningful relationships. As AI systems grow more powerful and sophisticated, their explainability becomes increasingly difficult. Moreover, a greater sense of agency arises from outcomes generated by humans rather than machines, with emotional feedback playing a crucial role in behavioural adjustment. This leads to speculation that automated feedback may be less relevant for fostering lasting behavioural changes. Moreover, AI evaluations used for grading, particularly when lacking critical discussions from peers and

teachers, can negatively impact self-regulation and self-reflectiveness processes. This may result in students struggling to connect curricula with personal development, leading to disengagement from the learning process and transforming the experience of achievement into a dualistic and impoverished experience of success versus failure.

Finally, teachers recognise that values and norms are context-dependent and can vary within a classroom, from classroom to classroom, from school to school, and from region to region, as well as among different stakeholders. With the advent of Generative AI, it may be challenging to harmonise perspectives on what should be accepted from this tool. For instance, can a student's work that utilised Generative AI be considered original? If students and researchers are honest about their use of the tool, can their work genuinely be regarded as their own? Are honesty and transparency equivalent? Furthermore, can Generative AI be seen as a proxy agent? The answer to these questions, as well as the plurality of these questions and their capacity to provoke further inquiry will be critical for moral reasoning and development moving forward.

5.4.1.2. Intersubjective Level

In terms of the intersubjective dimension, teachers primarily emphasised their role, which can be explained by narratives that view AI as a potential replacer or, at the very least, as a tool for fulfilling several critical functions traditionally held by teachers, while stripping away their core purposes and values. This raises the question: when teachers are relieved of administrative tasks, what other responsibilities are they freed from, and where are they displaced to? Some insights into possible answers may help explain why teachers highlighted their importance as role models, as well as their unique capacities for scaffolding. They also stressed their significance as proxy agents, possessing an authority that must be preserved to maintain their professional status and consistency.

The teachers involved in this research were notably silent regarding transformations among peers and the relevance of a sense of belonging in the learning experience. They did not address parents' concerns, nor did they discuss how distributed agency – where parental engagement is key – could serve as a touchstone for a meaningful accomplishment of educational initiatives. Such an approach to parents could help ensure that learning is transferable and generalizable to other contexts while preserving both the conditions and content relevant to learning. Furthermore, such a commitment may support the rise of AI literacy levels within educational communities.

Considering how detailed and prolific teachers were regarding the potential effects of AI systems on the subjective dimension of agency, this focus on their roles on the intersubjective level may express not only how hegemonic narratives from Big Tech can affect their professional autonomy and deliberation but also their capacity to critically examine the impacts on various relationships and actors within schools. This narrowing of thought can be a consequence of these procedures, as well as the processes through which these socio-technical systems reinforce their own entrenched status and solidify their influence, pushing aside the relevance of, or obscuring, the various levels of relationships that education encompasses.

5.4.1.3. Collective Level

From a collective perspective, teachers continue to uphold the vision of schools as democratic spaces, echoing Dewey's principles. In today's society, where political spaces are increasingly under threat – dialogues being replaced by (short and efficient) talks, plurality giving way to homogeneity, and politics overshadowed by policies – this vision holds significant importance. Creative writing and expression (often diminished by Large Language Models), for instance, do not merely produce novel outputs; they provide intellectual nourishment, prompting meaningful, context-driven innovation. Furthermore, achieving shared flow and emotional synchrony can foster a strong sense of belonging, enhancing participation, accountability, and distributed agency with expanded roles and synergies.

Situational freedoms have long been threatened by psycho-pedagogical techniques that presented themselves as a pharmakon in education, obscuring the vocation of education as a collective foundation for experiencing societal challenges and exploring personal genealogies of feelings and bonds. Technologies like automated essay scoring, predictive learning analytics, and AI-enhanced classroom management tools may offer immediate or short-term benefits but pose potential threats in the medium to long term. Critical pedagogy, which has traditionally integrated technology to support self-directed and heutagogic learning processes, is at odds with behaviour conditioning practices, such as gamified engagement. For instance, AI-enhanced peer tutoring systems can provide guidelines for constructive feedback, and gamified learning platforms powered by AI can facilitate role assignments tailored to students' unique characteristics.

Teachers and students are less likely to view this technological support with suspicion if there are opportunities for organic interactions and spaces for personal and collective recognition of each individual's contributions. Such an environment can provoke affective-

cognitive dissonance, which is crucial for embracing contradiction and divergence in social interactions and for building genuine common ground, far removed from illusions of full mastery.

Distributed agency is considered here as a subset of collective agency, reflecting the increasingly entangled relationships between human and non-human actants (e.g., AI systems, algorithms, data infrastructures) in educational settings. Within this framing, distributed agency does not imply a dilution of human responsibility but rather offers a conceptual stance for maintaining the primacy of the human domain, especially in contexts shaped by socio-technical systems. By acknowledging the roles of both human and non-human agents in shaping outcomes, this perspective promotes an ethically grounded awareness of how agency is negotiated, shared, and enacted across networks. In this sense, distributed agency, as incorporated into collective agency, supports a political and ethical commitment to preserving human values, judgment, and participation in decision-making processes, particularly in contexts where human agency is extended or mediated by technological elements.

5.4.2. Cross-Level Patterns and Interdependencies

A cross-dimensional analysis of the qualitative (see Figure 10, below) data reveals thematic patterns that cut across the subjective, intersubjective, and collective levels. These themes illustrate the intricate nature of agency in educational contexts and demonstrate how AI socio-technical systems may influence or reconfigure agency across these domains. Particularly, teachers' perspectives emphasise that agency is not a fixed attribute but a dynamic, relational process, continually shaped by institutional structures, technological environments, and interpersonal interactions.

5.4.2.1. Tension between Automation and Authorship

Across all three levels, teachers voiced concerns about how automation might displace human authorship and control. At the subjective level, this concern manifests in the perceived loss of students' autonomy, especially when AI-driven decisions bypass personal meaning-making. At the intersubjective level, the fear of being reduced to mere supervisors threatens teachers' roles as intentional and relational figures; roles that are critical to education's dual desideratum of subjectification (which is always made possible through the alterity of significant others) and socialisation (Biesta, 2020). Collectively, the emphasis on distributed

decision-making reflects a broader anxiety over top-down AI implementations that may further erode stakeholder agency.

This analysis of the recurrence of the themes of authorship, meaning-making, and control across the three levels offers a comprehensive vision of agency as a composite phenomenon that intrinsically nurtures itself: the more it grows subjectively, the more it can be gained collectively; and the more it grows collectively, the more it can be regained subjectively.

5.4.2.2. Foreclosed Agency as a Substitute for Proxy Agency

Another thematic recurrence concerns the complex balance between proxy agency (teachers and technologies extending the agency of students) and a kind of foreclosed agency, in which both significant others and tools acquire the power to resonate or act on behalf of students. While teachers accept a mediating role, they caution against the complete outsourcing of educational decisions to AI. At the intersubjective level, teachers assert that professional autonomy must be preserved; this very assertion constituting a foundational act of agency. At the collective level, they advocate for participatory processes in AI implementation, ensuring that decisions are not dictated solely by external agents or commercial interests.

Once again, this pattern illustrates that agency is deeply relational. The development of autonomy in students is conditioned by the autonomy and reflexivity of educators, which, in turn, depends on democratic and transparent policies and institutional frameworks.

5.4.2.3. Moral and Developmental Scaffolding

Teachers repeatedly highlight their role in guiding moral development, fostering foresight, and supporting emotional and meta-cognitive growth, especially at the subjective and intersubjective levels. These processes, including the development of critical reasoning and self-reflection, are seen as critical for meaningful and context-sensitive learning. At the collective level, teachers view democratic participation and dialogue as key not only for institutional functioning but for cultivating students' sense of social responsibility.

Teachers conceive themselves not merely implementers of curriculum or AI facilitators working solely toward qualification (Biesta, 2020). They not only recognise, but seem to embody, a professional and expressive dimension: to be moral and affective agents in the lives of their students.

5.4.2.4. Subjectification and Diversity

The theme of diversity, both in learning experiences and in the exploration of possible developmental paths, is central. At the subjective level, AI's potential to homogenise education poses a threat to processes of subjectification (Biesta, 2020), where learners construct their identities through engagement and choice. Within the intersubjective space, teachers draw on their socio-emotional skills to support diverse student trajectories and to cultivate bonds and context-sensitive socio-learning experiences. At the collective level, an emphasis on inclusive, participatory decision-making reflects the value of diverse perspectives, both in fostering political and democratic development and in reinforcing a sense of belonging, shared flow and emotional synchrony, which significantly contributes to a sense of collective achievement (Zumeta et al., 2016). Therefore, as individuals become more attuned to their environment and to one another, they are also better positioned to exercise agency in transformative ways. This aligns with Archer's (2015) concept of double morphogenesis, which describes how structural changes create new conditions that elicit responses from social actors, who in turn reshape those very structures. In schools, evolving educational policies, social expectations, and technological shifts do not simply impose change from above; rather, they invite reflective engagement, collaborative tinkering, and new practices within communities. Through this recursive process, teachers and students move from merely responding to projects and initiatives to actively shaping them, making collective agency both a consequence and a catalyst of meaning-making and transformation.

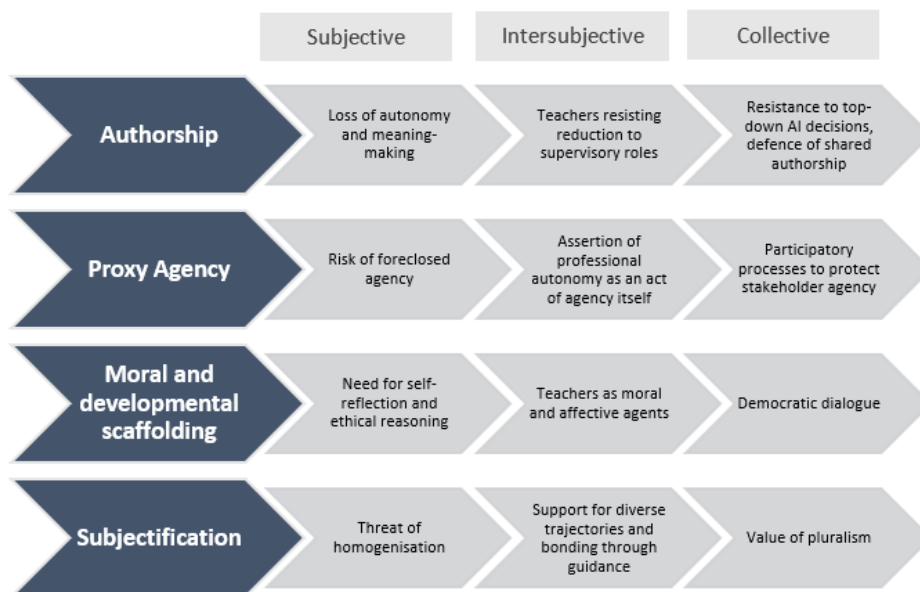
These perspectives point to a vision of agency that is developmental, dialogical, and contextually situated (Archer, 2014; Biesta & Tedder, 2006; Emirbayer & Mische, 1998). Moreover, the analysis of these aspects is valuable not only as a safeguard for the mechanisms that sustain human authorship at each specific level, but also as a recognition of agency as a compositional phenomenon, where the diminishing of agency at one level appears to reinforce its erosion at others, ultimately cutting off the roots of schools as spaces for the joint elaboration of possible subjective and collective futures. And this goes beyond the idea of merely creating futures from which the present condition deprives us. It raises a more fundamental question: what kinds of futures can we imagine without a present that nourishes desire? In digital capitalism, the collapse of desire manifests as a frantic urgency to appear alive, set against an ever-expanding void of meaning. A truly collective future is to be conceived as a project both dictated by and dedicated to attentional forms of care, at once psychic and social (Stiegler,

2015). This is a project that understands desire as the intensification of individuation, rooted in the dynamics of attention, memory, and becoming. In this sense, agency can be understood as protention (the anticipation of the *about-to-come*), in Stiegler’s (2015) terms: the capacity to anticipate and project oneself into the future through the practice of imagination and desire. It is a temporal orientation that links subjectivity with the not-yet, with what can still be created, decided, or hoped for. Within this framework, technologies, particularly AI, could occupy a meaningful role as forms of tertiary memory, that is, as externalised supports of human thought and collective experience. Rather than replacing memory, judgment, or intentionality, they could reconfigure traces of past knowledge and imaginative projections, helping individuals and communities to sustain or question their orientation toward meaningful futures. However, for this potential to be realised, we must break the cycle of closure, standardisation, and overdetermination that shapes these technologies; and this means the obvious: they must be reinvented.

The following diagram visualises the articulation discussed above, mapping the four key thematic categories onto the three analytical levels; each representing a distinct dimension of agency: subjective, intersubjective, and collective.

Figure 10

Thematic Flowchart of Subjective, Intersubjective, and Collective Levels



5.5. Conclusion

This study highlights how human agency may be transformed through the use of socio-technical AI systems in education. It explores the potential impact of these systems at subjective, intersubjective, and collective levels, drawing on the perspectives of teachers and professors responsible for teacher training. Concrete AI applications are presented in relation to their potential dimensions of impact, particularly concerning metacognitive and affective processes that contribute to long-lasting developmental processes, such as vocational, moral, and political growth.

The insights gathered reveal that teachers demonstrate high levels of differentiation when considering their students' subjective agentic dimensions. However, their focus appears more constrained when exploring the impact of AI systems on the intersubjective substratum. This narrowing may indicate an epistemic encroachment into narratives of dystopian futures devoid of teachers, paving the way for their own alienation from meaningful references within the framework of school bonds and strategic connections necessary for bringing projects to life and achieving impact.

This study is not without limitations. While the qualitative approach provides rich insights, it may not comprehensively capture the breadth of experiences and perspectives across various educational contexts and regions. Additionally, the focus on a specific demographic of educators may not represent the wider educational landscape. Futures studies could include a broader range of voices, particularly from students, parents, educational administrators, and even vendors, to gain a more comprehensive understanding of the potential impact of AI on agency and institutional powers. Moreover, although the focus groups included educators from diverse backgrounds across K-12 and higher education, the findings were not differentiated systematically according to educational sector, subject area, or level of teaching experience. Future research could meaningfully explore these differences. For instance, teachers working in vocational education, operating closer to industry demands, might conceptualise agency differently from those in more academically oriented settings. Similarly, younger or less experienced teachers, potentially more familiar with digital tools and AI applications, may engage with these technologies in ways that differ from more experienced educators, who may be more sensitive to relational disruptions and the complexities of situated educational contexts.

Furthermore, while this study discussed AI in broad terms, it did not distinguish between the implications of different types of AI systems. Future work could usefully examine whether co-creative AI tools, which invite students to generate, reflect, and revise, better foster agency than prescriptive systems that may automate decisions and narrow learner autonomy.

Regulatory frameworks to support agency-centred integration of AI into education were also beyond the primary focus of this study. Nonetheless, the findings highlight the absence of such frameworks as a critical gap, which future research should address, particularly in designing participatory and ethical approaches to AI governance in education.

Future research could also explore longitudinal studies examining the long-term effects of AI on agency dynamics across diverse educational settings.

CHAPTER SIX

Course Design and Evaluation

6.1. Introduction: Situating the Course in the Research Trajectory

Rather than treating AI as a neutral innovation to be adopted at scale, this doctoral research positions it as a complex and contested socio-technical system that demands critical, situated, and ethical engagement. Amidst rapid technological shifts, ethics can be preserved, first and foremost, through the assertion of human agency, as a foundation for deliberate thinking, responsible action, and the safeguarding of an educational ethos in which educators are recognised as gatekeepers of educational purposes.

In alignment with the generic model of educational design research, which unfolds through three iterative and interconnected phases, namely analysis and exploration, design and construction, and evaluation and reflection, each stage of this project was designed to foster meaningful interaction with educational practice while also contributing to theoretical advancement and the maturation of a pedagogical intervention (McKenney & Reeves, 2018).

From the outset, the project was designed to move iteratively from theoretical exploration to grounded application. It began with a Systematic Literature Review using the PRISMA framework, (Mouta et al., 2024a), which identified critical blind spots in how ethics is addressed in AI and education scholarship. This review laid the foundation for the next phase, a Delphi-based Futures Study (Mouta et al., 2024b), which engaged cross-sector stakeholders to co-develop future-facing ethical dilemmas involving AI in education. These dilemmas became entry points for a series of Focus Groups with teacher trainers (Mouta et al., 2025b), enabling contextual reflections on the risks, potentials, and pedagogical demands of working with AI systems in formal learning environments, and informing the design of a teacher training course on these very issues.

The fourth phase was dedicated to constructing a conceptual framework on AI and agency (Mouta et al., 2025a), which explored how educators perceive the shifting contours of human agency in relation to socio-technical systems. Attending to the subjective, intersubjective, and collective dimensions of agency, this framework, together with insights from the previous phase, helped inform the design of a training programme that addresses ethical concerns and fosters distributed human agency. It provided the theoretical foundation for developing a participatory, educator-centred response to the challenges identified throughout the research process, embodying the iterative and cumulative framework of EDR.

The course introduced in this chapter, “From Automation to Autonomy: Educational Agency with AI”, is the practical and pedagogical outcome of this research arc. It represents

not merely an instructional tool, but a deliberate intervention into dominant narratives around AI, automation, and the future of teaching and learning, aimed at preserving the educational purposes of qualification, socialisation, and subjectification. The course is designed as a professional development experience for K–12 educators, grounded in the principles of relational pedagogy, critical reflection, and human agency. It was constructed using an educational design research methodology, ensuring that every element, from content to delivery platform, reflects both empirical findings and previous conceptual commitments.

This chapter serves a dual purpose: it documents the design rationale, implementation architecture of the course, and initial evaluation of its potential relevance to the designated targets, while also offering a pedagogical reflection on what it means to design professional development initiatives for teachers with and about AI, resisting technocratic simplifications and promoting agentic practices that instil agency throughout the learning design of each task. Through this, the chapter contributes to ongoing conversations on ethical AI integration, teacher professional learning, and the defence of education as a fundamentally human-relational practice.

6.2. Methodological Positioning for a Research-Based Pedagogical Intervention

The development and evaluation of the professional development course presented in this chapter are consistently situated within the methodological paradigm of EDR, as it offers a robust framework for addressing complex, context-dependent educational challenges through iterative cycles of design, implementation, and reflection (Barab & Squire, 2004; McKenney & Reeves, 2018). As outlined in the introduction to this thesis, EDR is particularly well-suited to emergent domains such as AI in education, especially Generative AI, which introduces a new form of agency into the educational landscape and demands renewed negotiations of meaning. It is most appropriate when conventional pedagogical templates or validated instructional strategies are absent, underdeveloped, or themselves part of the problem. Such "open" problems require an iterative posture and the gathering of a trail of evidence that informs projects redesign and theory revision (Kelly, 2006). Therefore, the chapter's methodological positioning embraces this view by recognising that the educational problem it confronts is “wicked” in several interlocking ways: it is epistemically unstable, as the AI-and-education-related content to be taught remains emergent and not yet consolidated, even among experts; pedagogically uncertain, given the limited established practice on how to meaningfully engage with it; institutionally fragmented, since teachers often lack adequate resources and

support systems, while researchers themselves remain only partially equipped; and socio-politically charged, as broader policy, ethical, and cultural concerns further complicate the integration of AI into schools.

Central to this approach is the recognition that teacher training serves not only as a tool for practice but also as a vehicle for generating contextually situated knowledge and advancing theoretical understanding. Drawing on socio-constructivist epistemologies, this study embraces knowledge as co-constructed through social interaction, experience, and reflection, principles that directly informed the participatory and dialogic design and initial evaluation of the course. This methodological alignment is reflected in several interrelated aspects of the course's creation and structure. First, the course content and activities were not developed in isolation but were co-constructed with practicing educators through a series of focus groups and iterative co-design cycles. This participatory approach ensured that the professional learning experience was not only theoretically informed but also contextually grounded and directly responsive to teachers' expressed needs, contexts, and ethical concerns regarding AI in their classrooms and schools. Second, the design process integrated empirical insights, such as teacher narratives, ethical tensions, and real-world classroom scenarios, which were analysed through established conceptual frameworks, including Biesta's educational purposes and the three layers of agency.

This integration ensured that the learning remained closely aligned with teachers' professional realities, making the course both relevant and applicable. Finally, the course as a whole, and each of its sessions individually, were subject to iterative testing and refinement. This process of ongoing adaptation, which is a hallmark of EDR, allowed the content to evolve in step with the ethical reflections and pedagogical feedback of the participants. As a result, the course stands as a methodologically sound response to the complex and context-specific challenges faced by educators in learning environments where AI technologies are in use, exemplifying EDR's capacity to produce both theoretically robust and practically meaningful professional learning interventions.

Through these participatory processes, the course supports teachers not as passive recipients of knowledge but as emancipated researchers and active agents shaping their professional learning and ethical understanding: a stance aligned with critical theory and post-structuralist perspectives on agency in education (Arendt, 1958; Ball, 2013; Biesta, 2013; Foucault, 1980; Freire, 1996; Giroux, 1988). Positioning this work within the EDR tradition thus enables a reflexive, adaptive, and collaborative research process that respects the

complexity of AI’s educational integration. It balances the dual goals of producing a practical, meaningful teacher development programme and contributing to the theoretical discourse on agency and ethics in AI-mediated educational environments.

6.2.1. Assessment in an Age of Automated Scaling: A Pedagogical Defence of Relational Learning

This course deliberately positions itself against the prevailing trend of extensive automation in education, particularly as it addresses the intersection of human agency and AI-driven learning technologies. In an era increasingly shaped by platform logics of scalability, efficiency, and data-driven optimisation, educational solutions often promote the so-called “personalised learning pathways” as a hallmark of innovation. Yet, from the standpoint of this course’s pedagogical intentionality and participatory design, priority is given to relational, dialogic, and agency-centred practices, which are fundamental to deep, contextual learning, whether delivered face-to-face or in blended formats.

This is not to dismiss the potential benefits of fully digital education, which can play a valuable role in broadening access and supporting flexible learning when aligned with pedagogical purpose. Nationally supported online platforms, including Massive Open Online Courses (MOOCs), provide broad access to topics ranging from digital skills to subject-specific content, especially for older learners and adult education (OECD, 2023). In the case of teacher professional development, many countries offer curated digital resources, some within closed-access systems, designed to enhance competencies such as digital literacy and pedagogy. Some platforms also support peer learning and professional exchange, which facilitate teacher collaboration and material sharing (OECD, 2023). However, despite the growing availability of online resources, the actual adoption of advanced, adaptive digital learning tools remains limited. As of 2024, the majority of digital learning materials provided on national platforms remain static in nature, digital textbooks, video lectures, one-way simulations, and tests. Intelligent tutoring systems and smart classroom analytics are still rarely implemented at scale (OECD, 2021, 2023).

Even the more advanced interactive digital textbooks used today are often non-adaptive and linear in design. Governments and institutions have begun exploring the applications of Generative AI, but systematic guidance, strategic investment, and pedagogical vision are still lacking when it comes to designing and using digital learning resources for agency development, not just content delivery. Moreover, studies consistently show a disconnect

between the availability of digital resources and their actual pedagogical use, with underutilisation or superficial application as common outcomes. This underlines the need for ongoing research, data collection, and monitoring frameworks to explore how to make the best of digital learning resources (OECD, 2023).

A relevant example of a fully online teacher training initiative is the AI4T (Artificial Intelligence for Teachers) project, a European MOOC developed to help secondary school educators critically engage with AI in education. While the project successfully delivered content across five countries and reached over 1,400 teachers, its evaluations reveal the limitations of fully online models (AI4T, 2024b; 2024c). One of the reports shows that although teachers accessed the materials, many engaged only at a surface level, with limited evidence of deep pedagogical transformation (AI4T, 2024b). The comparative evaluation further highlighted that meaningful impact occurred primarily where the MOOC was complemented by peer dialogue, contextual support, or school-based facilitation (AI4T, 2024c). These findings confirm that awareness alone is insufficient; sustained professional learning requires relational, reflective, and contextually grounded engagement. This is further supported by research demonstrating the potential of blended learning to enhance both learner outcomes and engagement when designed with intentional interaction and support (Dixit & Pathak, 2023; Zagouras et al., 2022).

This course draws from that same potential to propose a purpose-driven integration of technology, one that situates AI not as the end goal, but as a tool embedded within reflective, dialogical, and agency-nurturing learning environments. Nowadays, educators are called not simply to use AI tools but to support learners in developing a sense of agency: the discernment to decide when and how to engage with AI in ways that are critical, ethical, and creative (Gulya, 2023). Education, particularly in the context of AI, is expected to engage with symbolic, ethical, and intergenerational dimensions: questions of agency, imagination, intention, memory, legacy, tinkering, and the possibility of radical new beginnings. These questions cannot be addressed through automation alone; they require spaces for collective inquiry and institutional capacity-building. As Mollenhauer (2014) reminds us, education is ultimately an ethical act: it concerns the question of why we are with learners and what it means to represent the world to them.

Building on sociological perspectives of pedagogical discourse (Morais et al., 2005), this course acknowledges that meaningful teacher learning depends not only on access to content,

but on the social and discursive conditions under which that content is appropriated. Many automated training models promote solitary, consumption-oriented engagement, framing learners as users exploring pre-sequenced content with minimal social interaction. Such models fragment the learning experience and reduce opportunities for negotiated meaning-making, and collective knowledge-building, all of which are central to learning and developing institutional and pedagogical agency (Vygotsky, 1978; Dewey, 1933). It is through sustained interaction with peers, communities, technologies, systems, and ideas that professional identity and educational responsibility are formed and reinforced. Therefore, the architecture of this course is intentionally free from adaptive algorithms or fully automated sequencing. Instead, it establishes a shared pedagogical space where educators engage collaboratively with ethical dilemmas, conceptual challenges, and socio-technical complexities of AI. Such an approach reflects sound design principles in online education: alignment of tasks and tools with clear purposes, scaffolded learning experiences, and a strong emphasis on collaborative construction, ethical co-reflection, and feedback-rich environments.

By resisting the lure of scalable automation, the course preserves the possibility for locally contextualised implementation. It is designed with the flexibility to be delivered entirely face-to-face, in blended formats, or even fully online, provided it remains situated within a supervised, collaborative setting rather than in isolated, self-paced consumption. Regardless of format, the course offers a consistent structure, a robust resource base, and a triangulated assessment framework aligned with its pedagogical values. This includes multiple modes of evaluation (self, peer, facilitator), all reinforcing an ecosystem where assessment is relational, formative, and agency-enhancing, not driven by data extraction. Ultimately, this course positions educators not as recipients of innovation, but as critical co-designers of their professional practice, and as gatekeepers of education who ensure it remains both ethically grounded and pedagogically meaningful.

6.2.2. Reframing Beyond SCORM and Personalisation: Rethinking the Role of Technology in Course Delivery

Rather than defaulting to fully digitalised or prescriptive technological systems, this course adopts a contextual and situated pedagogical approach that responds to the diverse realities, values, and professional contexts of its participants. The focus of the current deployment and assessment is not on tracking learner behaviour through data analytics or using it to drive

decisions. In this way, the course maintains pedagogical coherence: form and content are aligned in support of a learning model that is responsive, relational, and ethically attuned.

This course is deliberately structured to foster both subjective agency through reflective activities such as learning diaries and personal mapping, and intersubjective and collective agency, via collaborative discussions, conceptual mapping exercises, and ethically complex role-play scenarios. The digital platform functions not as a rigid Learning Management System (LMS), but as a flexible pedagogical toolkit, adaptable for use, preferably in face-to-face training, but also in blended learning environments or as a reflective companion for self-directed exploration, always prioritising pedagogical intentionality over technological prescription. It is not that SCORM-compliant or eLearning platforms are inherently incapable of supporting dynamic, well-designed courses. On the contrary, many such systems offer features that can enhance interactivity, adaptability, and learner engagement.

In the context of this course's core educational philosophy, centred on fostering collective agency in institutional contexts such as schools, the decision to avoid a fully individualised, algorithm-driven learning path is intentional and principled. In this context, the adoption of a Canva-based platform is not incidental, but deliberate and principled: it is a pedagogical commitment. It consciously avoids fragmenting knowledge into atomised, linear sequences. Instead, it embraces a design grounded in dialogue, relational learning, and critical co-construction. Unlike conventional LMS that often foreground individual tracking and automated analytics, Canva, by design, does not offer built-in features for automated tracking, user analytics, or algorithmically driven content sequencing. This absence can be seen as a strength, as it may help reinforce human agency. By not defaulting to learning analytics or user monitoring, it reinforces a learning culture that resists technocratic evaluation models in favour of formative, dialogic, and situated forms of assessment.

It reinforces its refusal to adopt assessment mechanisms that reduce learning to quantified behavioural proxies, such as time-on-task, click patterns, or passive completion metrics, which are not only insufficient, but epistemologically misaligned with this course's objectives regarding teacher agency, where the teachers are conceived as active participants who engage critically with the content. The use of self-assessment, peer-assessment, and facilitator feedback within the Canva-based platform creates an environment of collaborative learning, where reflection and discussion are prioritised. These activities are central to the goal of

encouraging ethical deliberation and critical thinking about AI in education, rather than just technical proficiency with AI tools.

Canva's visual, collaborative, and open-ended design space makes it far more appropriate for context-sensitive appropriation, dialogic exploration, creative synthesis, and participatory meaning-making than conventional LMS environments optimised for individual progression and automated scoring. Therefore, what distinguishes this choice is the intentional prioritisation of pedagogical integrity over technological determinism, which enables the nuances of individual experiences and the collective wisdom of a community of practice. It supports a pedagogical ecology that views assessment as a shared, continuous process, one for which each individual is responsible and consciously engaged, while intentionally limiting the temptation to rely on learning analytics as the sole or dominant form of assessment. It maintains fidelity to the course's core values and pedagogical architecture, ensuring that the very tools used mirror the type of agency, responsibility, and collective authorship the course seeks to promote within institutional educational settings.

Another important factor in the decision to avoid a traditional eLearning platform is the accessibility and usability for educators who may not have strong technological backgrounds or who work in settings where internet access or advanced eLearning systems are limited. The Canva platform is user-friendly and easily accessible across different devices, and the course content can be exported as an e-book for offline use, further enhancing accessibility for diverse participants. Additionally, the face-to-face training component is strengthened because facilitators can utilise the platform's rich materials in a way that is more immediate and direct.

The following sections elaborate on how these methodological commitments manifest in the course's design, delivery, and evaluation.

6.3. Pedagogical Framework and Course Design

This section details the design and implementation of the professional learning course, "From Automation to Autonomy: Educational Agency with AI" (<https://anamouta.my.canva.site/from-automation-to-autonomy-educational-agency-with-ai>).

6.3.1. Aligning Course Design with Key Research Findings

This course, publicly available at <https://anamouta.my.canva.site/from-automation-to-autonomy-educational-agency-with-ai>, is a direct response to the findings from earlier phases of this study, particularly those published in the papers "Comprehensive Professional Learning

for Teacher Agency in Addressing Ethical Challenges of AIED: Insights from Educational Design Research" (Mouta et al., 2025b) and "'Where is Agency Moving to?': Exploring the Interplay between AI Technologies in Education and Human Agency" (Mouta et al., 2025a).

6.3.2. Design and Implementation Aspects

In response to Research Questions 3 and 4 ("What specific aspects should be taken into account when developing a training programme in this field?"; "What strategies can be used to update teachers' knowledge and skills on AIED while also fostering agency and ethical thinking about its potential impact on education?"), raised during the focus group stage of this research, the course design has been aligned with the main findings.

The course intentionally articulates the strategic and organisational aspects of professional learning. Each session includes a detailed session plan (cf. Appendix E) with clearly defined goals, thematic stages, specific objectives, methodologies, targeted learning processes, and required resources, reflecting a deliberate and well-structured design that directly responds to the need for systematic programme development. It also explicitly tackles the "transition challenge" by prioritising a smooth and contextually appropriate integration of AI technologies into meaningful activities, all driven by a clear pedagogical purpose. Sessions are structured to build familiarity and confidence gradually, starting with broad historical narratives on intelligence and technology, providing a less threatening context before diving into AI specifics. This approach addresses potential cognitive and affective resistance from educators. Furthermore, the course fosters confidence and facilitates transitions by encouraging teachers to engage with AI through a process of tinkering in a supportive environment. This is complemented by deliberate moments of shared and individualised exploration of teachers' own practices, using Task Diaries, Reflective Prompts, Group Discussions, and Scenario-Based Activities.

A key design principle is that the training would be "pointless without AI technologies". Therefore, the course explicitly integrates AI tools into both its design and delivery, ensuring that learning is rooted in hands-on engagement with tools such as Synthesia for critical reflection, Character.AI for simulated dialogues, Mentimeter for collaborative reasoning, Wayground for creating flexible assessments, or ChatGPT/Gemini for ethical reasoning and exploring distributed agency. The course focuses on agency and ethics in AIED, while acknowledging its deep interconnection with broader educational priorities such as qualification, socialisation, and subjectification, in a contextually assessed manner.

Although designed for in-service teachers, the foundational ethical awareness and critical perspectives fostered by this course could serve as a robust model for pre-service training, providing future educators with a strong conceptual foundation in agency, ethical awareness, and socio-technical literacy. The experiential and dialogic approach, such as role-playing in Session 7, could be adapted for university-level pedagogical courses. These critical perspectives are highly transferable and provide a solid groundwork for pre-service training.

The course also promotes informal and self-directed professional development, encouraging educators to take responsibility for their learning beyond formal sessions. Reflective practices, such as Task Diary prompts, foster this self-directed growth, while collaborative learning opportunities build a sense of community that extends beyond the course structure. The inclusion of "Suggested Resources for Further Exploration" further supports continued professional enrichment, enabling teachers to take ownership of their professional development and pursue ongoing learning independently.

The course is designed with a strong emphasis on innovative and purposeful pedagogical practices that challenge traditional norms, shifting the focus from conventional assessment to "knowledge-building processes". This is exemplified by activities where teachers redesign lesson plans using digital scientific tools with AI to explore how automation influences student engagement and meaning-making. It also encourages creative synthesis through activities like the "Hashtag Reflection" in Session 3, which challenges conventional forms of participation and is attuned to current social media trends in a critical way, making it both timely and easily transferable to schools.

Hands-on, skill-based training is reinforced through "learning by example", using real-life scenarios to foster ethical reasoning and perspective-taking. In Session 3, for instance, participants examine the hidden labour behind AI via a reading on Kenyan data labellers. Role-playing (Sessions 4 and 7) and role-taking (Session 8) further support exploration of the affective-cognitive tensions involved in engaging with technology and dialogic ethics.

Finally, this training aims at fostering collaboration primarily among teachers, recognising their key role in addressing complex ethical issues as they are understood in their own contexts. While the provided session plans emphasise collaboration among educators, family engagement is encouraged in the final phase through the preparation of a lesson plan based on key course takeaways. This dialogic approach to ethics, encompassing educational

relationships and peer dynamics, lays the basis for teachers to comprehensively facilitate broader stakeholder engagement.

6.3.2.1. Modules and Sessions Overview: Integrated Research Insights

The subsequent sections outline the course’s progression across modules (Figure 11, below), each accompanied by a session-by-session lesson plan (Figure 12, below) that integrates research findings and theoretical concepts to ensure a cohesive and practically grounded learning experience.

Figure 11

Modular Course Layout on Canva: Opening Slide View

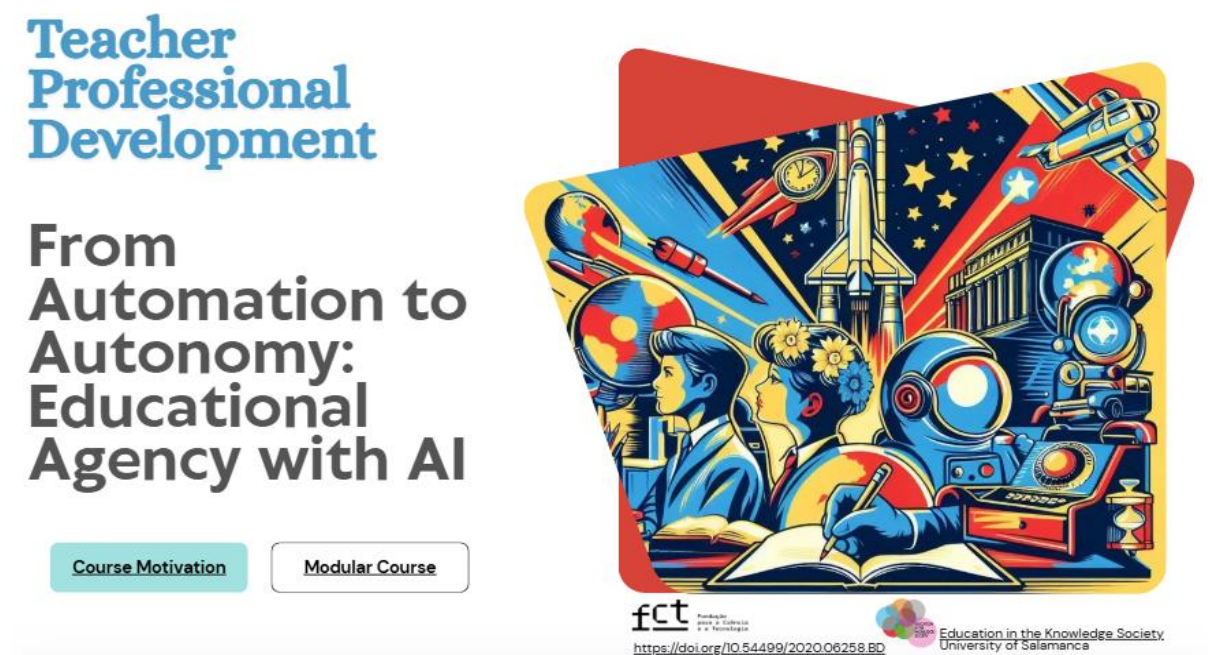


Figure 12

Sample Session Plan Demonstrating Applied Pedagogical Framework

MODULE 2, SUBJECTIVE AGENCY, SESSION 6, RECLAIMING AGENCY AND MEANING IN AI-CLASSROOM INTERACTIONS							
MAIN GOALS							
<ul style="list-style-type: none"> To foster connection and ownership in AI-mediated learning environments by exploring the educational implications of how students may feel when interacting with systems they do not fully understand or cannot influence. To support teachers in designing classroom experiences that enhance subjective agency through choice, goal alignment, and feedback fluency. <p><i>* Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.</i></p>							
SUGGESTED RESOURCES FOR FURTHER EXPLORATION: Schreiner, M. R., Neszemlyi, B., Schwarz, K. A., & Kunde, W. (2025). Goals rather than predictions determine the sense of agency. <i>iScience</i> , 112583. https://doi.org/10.1016/j.isci.2025.112583 ; Frith, C. D. (2014). Action, agency and responsibility. <i>Neuropsychologia</i> , 55, 137–142. https://doi.org/10.1016/j.neuropsychologia.2013.09.007 (Papers).							
THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from session 5	1. RECONNECT	To help participants integrate and transfer key ideas from the last session by reflecting on their meaning and restating them in their own terms and experiences.	Pair dialogue, whole-group share-out	Recap, exploration, differentiation, insight, synthesis, transfer	NA	NA	10'
	<p>Description: In this opening activity, participants revisit the previous session. First, they reflect in pairs on the following question: "Did the previous session surface any tensions between your values, available tools, and practices?". Then, in a process of "recapitulation", they restate a key concept, takeaway, or metaphor from the session using their own words or images, which they can share with the whole group.</p>						
Development: Feeling Control in Data-Driven Classrooms	2. EMBODIED EXPERIENCE: "BALL VS. BALLOON"	To allow participants to experience the difference between fluent, intentional action and disrupted, unpredictable feedback, forming the basis for analysing agency in AI-enhanced learning.	Embodied paired simulation, guided reflection	Sensorimotor awareness, introspection, perspective-taking, conceptual abstraction, emotional-cognitive integration, transfer	One foam ball (medium) per pair; one balloon per pair (extra balloons and balls on hand in case of pops or drops); Course's Canva platform	Group-discussion-based reflection	15'
	<p>Description: In this activity, participants work in pairs, with one taking the role of Initiator and the other as Responder. In the first round, using a foam ball, the Initiator throws the ball to the Responder, who catches and returns it in a smooth, rhythmic pattern for 6 to 8 exchanges. This establishes a sense of predictable, controlled interaction. In the second round, the ball is replaced with a balloon. The Initiator throws the balloon in the same way, but due to its unpredictable floating and wobbling, the outcome may feel delayed or misaligned with their intention. After both rounds, pairs reflect on the experience, considering the following questions: "When did you feel more in control?"; "Did the object's behaviour match your intention?"; "What did it feel like to act when outcomes were unpredictable?". These questions help explore the relationship between feedback, control, and the subjective sense of agency. These reflections are then expanded in a short plenary debrief. The facilitator links the experience to the research of Schreiner and Frith on the sense of agency (cf. "Suggested resources for further exploration"), which argue that our feeling of control over actions is not a direct perception, but rather an inference based on cues such as fluency, timing, and outcome consistency. When feedback is unpredictable or feels misaligned, even when the action is technically</p>						

6.3.2.1.1. Module 1. Collective Agency: Where Do We Stand in the Education and AIED Landscape?

This module introduces the course by grounding the abstract discourse on AI in collective narratives and shared educational experiences, drawing from socio-constructivist and phenomenological traditions. It sets the foundation for participants to critically examine dominant narratives about AI in education, while affirming their roles as epistemic agents, active shapers of knowledge, meaning and pedagogical practice rather than passive consumers of technological innovation. At its core, the module explores the concept of collective agency, the capacity for shared action, meaning-making and responsibility among teachers, students, technologies and institutional actors. Rather than viewing agency as the domain of isolated individuals, this perspective emphasises interdependence, co-construction and the ways educational futures are shaped through dialogue, participation and shared ethical commitments. Participants will develop foundational literacy in socio-technical AI, exploring core concepts and applications as they relate to educational contexts. They will also engage in critical self-reflection, exploring their personal positioning in relation to AI technologies in education, both emotionally and ethically, as well as from a professional standpoint.

By surfacing affective and ethical dimensions of automation, teachers begin to understand how their own assumptions, fears and aspirations intersect with broader socio-technical imaginaries. Through structured dialogue, reflective tasks, and participatory exploration, this four-session module fosters collective critical inquiry as the first step in building a community of practice grounded in agency, dialogical ethics, negotiated reflection, and a shared educational purpose.

Session 1: EdTech Imaginaries and Socio-technical Assemblages. This session addresses the research finding that teachers often lack structured opportunities to critically analyse dominant narratives surrounding AI in education. It aims to promote critical awareness of how educational technologies are historically imagined, socially constructed, and ideologically shaped. Conceptually, it introduces "educational imaginaries" as socio-technical constructs, linking to the findings that highlight the need for understanding AI's social shaping. An activity example involves participants exploring a visual timeline that highlights claims and promises of educational advances through technological integration. It emphasises how similar underlying imaginaries are often revisited and reintroduced through newer, more sophisticated technologies. Teachers then reflect in their Task Diary and engage in a discussion on what "innovation" truly means in this context, engaging with the political dynamics of imagination, time, and legitimacy that influence the dominant AI narratives and seek to shape the future of education.

The assessment strategies include self-assessment through a Task Diary reflective entry on personal positioning within AI discourse. Peer assessment via layer matching and quote interpretation dialogue during a Mentimeter activity. Facilitator observation of narrative depth and conceptual clarity in "Snapshots" presentations.

Session 2: From Users to Shapers: Agency, AI, and Shared Responsibility. This session operationalises the typology of agency (subjective, intersubjective, collective, distributed) developed through qualitative coding of teacher interviews in the research. It deepens teachers' understanding of the distinct educational purposes (qualification, socialisation, subjectification) and their relationship with different forms of agency. Key activities include Small-Group Dialogue, where teachers explore common AI tools (e.g., data dashboards) through Biesta's educational purposes, reflecting on how these tools serve qualification. Another activity, Role-Based Ethical Mapping, involves teachers role-playing scenarios to explore AI's impact on students' agency within an ethical framework. The final core activity,

character.ai prompting, allows teachers to deepen their reflection on AI's role in human agency through a pre-set Generative AI figure. This activity promotes reflective transfer between abstract concepts and classroom practice, fostering dialogic sense-making and reinforcing intersubjective and collective layers of agency.

The assessment strategies include self-reflection on personal agency in AI usage through a Task Diary. Peer feedback on prompt development in small groups. Facilitator evaluation of discussions on the impact of AI across different layers of agency, ensuring conceptual depth.

Session 3: Beyond the Algorithm: Human Stories in AI Education. This session aligns with the research's emphasis on developing ethical consciousness as a form of agency. It focuses on the invisible labour, bias, social, and environmental costs of AI systems, highlighting the lack of formal spaces to explore these tensions. The session strengthens teachers' capacity to critically, creatively, and contextually analyse education's role in the AI landscape, exploring the socio-technical and human dimensions of algorithmic systems. Key activities include a "Safety vs. Surveillance Debate", where teachers engage in structured debates on ethical dilemmas such as the use of facial recognition or AI grading. Another activity involves participants crafting a hashtag (e.g., #BehindTheAlgorithm) that captures their concerns and insights from the session, fostering a collective dialogue on the ethical, social, and environmental implications of AI as they relate to their own living and learning contexts.

The assessment strategies include self-reflection through writing on ethical discomforts with AI in practice. Peer constructive feedback on ethical dilemmas shared by peers. Facilitator observation of critical engagement during debates.

Session 4: Collective Agency in Action: AI and Participatory Practices in Schools. This session enables teachers to engage in transdisciplinary collective agency by examining how algorithmic decisions are made and audited in educational settings through real-world cases. Key activities involve Group Redesign, where teachers apply AI tools to reframe lesson plans for their discipline, integrating both AI-generated and student-generated content. Role-play activities allow teachers to explore how collective agency can be strengthened in environments where AI is integrated in ways that either facilitate or challenge practices of authorship and ethical development within groups, highlighting different levels of participation in lesson planning.

The assessment strategies include self-Assessment through insights from teachers' experience in role-play activities on how AI integration for subject matter can impact collective agency, authorship, and ethical development within groups. Peer Assessment through the evaluation of the lesson redesigns within small group settings. Facilitator Assessment through the observation of the depth of the role-play dynamics.

6.3.2.1.2. Module 2. Subjective Agency: Relational Pedagogy in an AI Age

This two-session module focuses on nurturing teachers' capacity to recognise and support subjective agency, the internal, personal dimension of agency through which students feel, perceive and enact their ability to make choices, influence outcomes and act with intention in learning contexts. Subjective agency is closely tied to engagement, self-awareness, self-regulation and a sense of ownership over one's learning. The module approaches this through the lens of ethical dilemmas as threshold concepts, moments of cognitive and emotional dissonance that prompt critical reflection and open up diverse ways of seeing teaching and learning. Rather than presenting ethics as an abstract theory, the module situates ethical reasoning within pedagogical decision-making in context, asking what kinds of choices students really have and how AI systems shape their experience of autonomy, authorship or marginalisation. Participants will examine real and hypothetical scenarios involving AI in education, identifying tensions between automation and agency, and exploring how design choices influence students' subjective experiences of learning. They will be guided to understand dilemmas not as part of a problem-solving continuum, but as spaces for pedagogical and political decision-making, requiring both imaginaries and lived experience in the teacher's role as co-auditors of AI-mediated practices.

Session 5: Invisible Scripts: AI, Student Selves, and Agency. This session intends to develop teacher agency through the deepening of their role as co-auditors of AI systems, fostering shared language and criteria for agency-respecting AIED use in schools. Key activities involve AI-Driven Feature Analysis, where teachers examine how AI features such as adaptive pacing or predictive analytics affect students' subjective agency (e.g., self-regulation).

The assessment strategies include self-reflection on subjective agency within AI-driven classroom practices. Peer feedback on AI features' impact on student autonomy. Facilitator analysis of teachers' understanding of AI's impact on subjective agency, through observation of the pair analysis activity and engagement during the subsequent group debrief.

Session 6: Reclaiming Agency and Meaning in AI-Classroom Interactions. This session aims to foster connection and ownership in AI-mediated learning environments by exploring the educational implications of how students may feel when interacting with systems they do not fully understand or cannot influence. It supports teachers in reclaiming agency within these environments, helping them design classroom experiences that enhance subjective agency through choice, goal alignment, and feedback fluency. Key activities include "Prompting as a Practice of Authorship", where context, instruction, questioning, sequencing, iteration, tone, and examples provide opportunities to bring meaning back to the centre of learning, through intentionality and planning. There is also an activity focused on engaging students in defining success based on their own terms and criteria, thereby creating spaces for autonomy.

The assessment strategies include self-reflection through a Task Diary on where agency could be reclaimed in AI use. Peer group assessment of contributions to a richer repertoire of assessment practices that can be considered and applied in the classroom. Facilitator evaluation of reflections on rebuilding opportunities for authorship within schools.

6.3.2.1.3. Module 3: Intersubjective Agency: Socialisation and Togetherness

This one-session module challenges dominant narratives of efficiency, automation and hyper-personalisation that often underpin the integration of AI in education. Instead, it foregrounds a pedagogy rooted in care, alterity, dialogue and the transformative potential of human encounters. The session invites participants to critically reflect on the kinds of relationships that AI systems promote or displace, and to imagine alternative digital futures in which intersubjectivity (the shared, reciprocal process of meaning-making between people) remains central to teaching and learning. Drawing on traditions of relational pedagogy, the module asks how AI tools may affect the ways students experience belonging, recognition and dialogue within educational spaces.

It examines the subtle shifts in teacher-student and peer-to-peer dynamics that can arise when algorithmic logics of prediction, efficiency and control are prioritised over messier, human-centred processes of learning. Learning goals include exploring the impact of AI on the emotional and relational dimensions of classroom life, developing a critical vocabulary for contrasting algorithmic and relational learning logics, and recognising the pedagogical and developmental value of uncertainty, discontinuity and open-ended dialogue.

Session 7: Intersubjective Agency: Socialisation and Togetherness. This session (Figure 13, below) supports teachers in identifying how intersubjective agency is challenged in AI-mediated educational environments, particularly in how system design and its application influence relationships, participation, and peer dynamics. It deepens teachers' understanding of how belonging, shared meaning-making, and situational freedoms are central to students' socialisation processes, and how these elements can be shaped by the integration of AI technologies in schools. A short, animated video illustrates how a person's sense of agency shifts depending on whether they believe they initiated an action or if it was driven by an external system, leading to a debate on the effects of automation and the responses that further enact distributed agency. This sets the stage for further framing agency as both a social and mental construct, including themes such as the impact of automated systems on responsibility and their implications for academic integrity. Participants will then break into small groups and rotate through four role-pair configurations involving a student and a teacher, addressing a scenario where a student used Generative AI for a school task, with varying responses from the teacher (e.g., "Full Copy, Negative Response" or "Critical Use, Constructive Response").

The session culminates in a Pedagogical Design Task (Shaping Agency with AI), which allows teachers to apply their understanding of the different layers of agency and educational purposes they aim to activate within their own teaching contexts. This task is designed to help teachers create curriculum-infused activities that align with their pedagogical goals while exploring the ways AI tools can support student agency.

The assessment strategies include self-Assessment through individual reflection on the design choices made for the planned activity. Peer Assessment through the evaluation of the designed activity plans by a designated teacher/pair. Facilitator Assessment through the review of the refined activity plans.

Figure 13

Example of the 7th Session Entry on the Course support Platform

Module 3. Intersubjective Agency

Session 7. Intersubjective agency: socialisation and togetherness

[Session Plan. M3.S7.](#)



6.3.2.1.4. Module 4: Distributed Agency: Sustaining Collective Praxis

This module consists of a single culminating session and moves beyond individualised and technocentric views of agency by foregrounding collectivity, interdependence, and ethical co-responsibility. It invites participants to imagine and enact educational practices in which agency is not centred on individuals or technologies alone, but co-constructed and shared across teachers, students, families, technologies, and institutions. Learning goals include exploring distributed agency as both a design principle, guiding how learning environments and AI tools are structured to support collaboration and shared authorship, and an education ethics, rooted in values such as care, participation, and mutual accountability. The module culminates in a collective reflection on classroom experimentation and the relational structures and processes that support meaningful agency practices. Participants will co-create a community of practice to sustain professional growth and innovation beyond the course, laying the foundation for ongoing collaboration in shaping the educational imaginaries for which we are always, collectively, responsible.

Session 8: Integration and Future Praxis. This final session serves as a capstone that consolidates teachers' learning across the programme and supports their transition from course

participants to active members of a sustained professional learning community. Drawing from the collective lived experiences of designing and implementing pedagogical activities using AI tools, the session focuses on reinforcing distributed agency and ethical responsibility as long-term educational commitments. The notion of distributed agency is revisited both as a design principle, an education ethic, and shape of thought, highlighting how teachers, students, and families, together with technologies and institutions, can act as co-agents in shaping a meaningful and ethically grounded pattern of collective agency in schools.

Teachers are invited to present the AI-integrated activities they implemented in their classrooms, succinctly sharing the subject, theme, selected AI tool, targeted layer of agency, educational purpose, and their key reflections on what worked, what surprised them, and what could be improved. The emphasis is on surfacing concrete lessons and principles that can inform future practice and lay the foundation for their emerging community of practice.

The assessment strategies include self-assessment through reflective commentary on the implementation of the AI activity, captured using the “Module 3: Evaluation of Draft Activity” form. Peer assessment through the evaluation of each other’s presented activities by designated feedback partners, focusing on pedagogical coherence, ethical integration of AI, and overall feasibility. Facilitator observation by monitoring conceptual clarity and collaborative engagement during presentations and the collective discussion, with a focus on distributed agency and community commitment.

6.4. Course Piloting through Walkthrough and Micro-evaluation

This section details the walkthroughs and micro-evaluation testing phases for the course. In the context of design research, evaluation plays a dual role: it supports the iterative development of the intervention while also generating insights that may be valuable to the broader scientific community. Such evaluation is intentional, structured, and embedded within reflective sub-cycles, each guided by its own internal logic. Before addressing the relevance of adopting both investigative and inventive mindsets, it is important to note that the evaluation and reflection phase involves empirical exploration of design concepts and prototype implementations ((McKenney & Reeves, 2018). This phase aims to deepen understanding of whether, how, and why specific elements of the intervention are effective.

The present study employed observations, interviews (walkthroughs) and questionnaires (micro-evaluation) as key methods. Importantly, these tools were not intended for impact assessment or comparative analysis, such as tracking changes in participants’ knowledge,

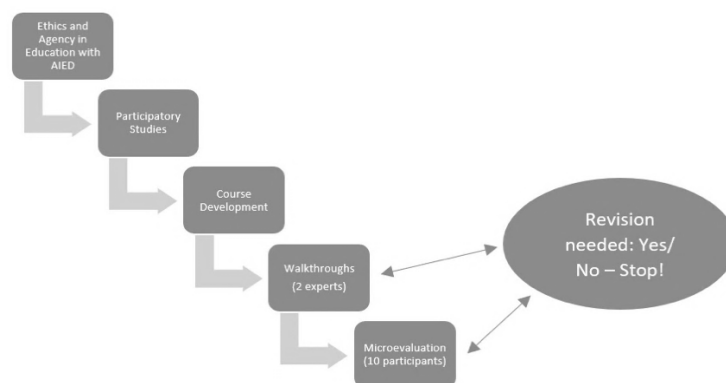
skills, or attitudes before and after the sessions. Instead, they serve as a mechanism for formative evaluation, primarily aimed at contributing to the improvement of the intervention and recognising its potential at a micro-evaluation stage to ensure it is ready for broader implementation. Although walkthroughs and micro-evaluations in EDR offer valuable insights into the expected practicality and effectiveness of an intervention, it is understood that a comprehensive demonstration of the course's actual, large-scale impact would necessitate more extensive field testing later in the research process (Plomp & Nieveen, 2013).

Walkthroughs often involve expert assessment and simulated scenarios, while micro-evaluations utilise small, purposive samples in controlled settings, neither of which fully replicating the complexities of real-world implementation with the full target group. This formative evaluation approach addresses these limitations by focusing on the generation of concrete suggestions for improvement over standardised data collection in the early stages. It highlights the richness of information, relevance, and adequacy of procedures in a new field of professional development, which required a solid establishment before full intervention.

By addressing these factors early in the development process, the walkthroughs and micro-evaluations enabled iterative agile adjustments (Figure 14, below), strengthening the course's alignment with its goals: reinforcing the broader purposes of education beyond qualification, while engaging with AI in ways that preserve educational ethics and foster collective agency.

Figure 14

Iterations of the Systematic Design Cycles



Note. Adapted from *Educational Design Research* (Vol. 1), by T. Plomp & N. Nieveen (Eds.), 2013, SLO. Adapted and used under fair use for educational purposes.

6.4.1. Walkthrough Stage

6.4.1.1. Walkthrough Rationale

The walkthrough phase involved a multi-faceted data collection strategy, which included directly observing how participants engaged, gathering qualitative feedback, and noting their interactions with the course materials. It allowed for the direct observation of participants' engagement, the identification of design strengths, and the pinpointing of areas requiring refinement prior to broader implementation. These walkthroughs represented a participatory and dialogic process, providing immediate, contextualised feedback to reinforce the course's pedagogical coherence and practical applicability. In the continuous cycles of design, implementation, and reflection inherent to Design Research, the walkthrough phase embodied the “implementation” and initial “reflection” components, providing a preliminary, but deep exploration of the course modules. Key objectives for this phase included:

1. Validating pedagogical flow: assessing whether the logical progression of concepts and activities aligned with the intended learning trajectory and facilitated meaningful understanding and participation.
2. Gathering initial feedback on content relevance and engagement: eliciting qualitative insights on the clarity, depth, attractiveness, and relevance of the course content, as well as the accuracy of specific activities in fostering personal and collective agency.
3. Identifying potential scaffold misalignments: uncovering discrepancies between the designed pedagogical intent and participants' actual experiences, particularly concerning the course's core stance against automated scaling and its emphasis on relational learning.
4. Identifying usability and accessibility: observing how participants/experts interacted with the course materials, particularly the session plans and the support platform on Canva. The aim was to ensure fluid navigation and easy access for educators with varying levels of technological proficiency.
5. Conducting in-depth analysis of specific course components: exploring particular course modules and learning activities with a small group of participants to uncover granular insights into their effectiveness, identify precise points of misunderstanding, and generate detailed suggestions for targeted improvements.

6.4.1.2. Methodology of Walkthrough Exploration

The walkthrough exploration was structured to simulate key dimensions of the course, prioritising qualitative data collection, direct engagement with participants, and real-time formative feedback. As such, the process was mainly developmental, dialogic, and iterative. Feedback was collected through think-aloud procedures, facilitated discussions, observation notes, and structured module feedback diaries. It adopted a deductive and design-informed thematic structure. This structure was shaped by the walkthrough instrument itself: each feedback diary prompted participant reflection across key dimensions: initial impressions, UX design, pedagogy, student impact, and timing. These dimensions were treated as analytic categories, allowing for the systematic redesign of the course through comparison of responses across modules and between participants, while recognising the distinct phases at which each participant engaged with the process; the second reflecting on a reconfigured version of the course shaped by the first participant's initial walkthrough.

While the feedback from the first participant was frequently delivered in real time, face-to-face or via voice call, and often resulted in immediate revisions, responses were also documented and retrospectively organised according to the same thematic structure. With the second participant, feedback was collected more summarily via completed forms and informal discussions, following the revised version of the course.

Although the walkthroughs functioned as embedded evaluation rather than as data collection for open coding, the resulting analysis still drew on a structured thematic framework, aligned with the key pedagogical and design dimensions under investigation. These thematic clusters are presented in Section 6.3, where each subsection (e.g., usability, engagement, pedagogical intent) reflects a targeted lens through which iterative improvements and course relevance were assessed.

6.4.1.2.1. Participants/Experts

A select group of two University lecturers in teacher education for K-12 participated in this intensive walkthrough process, at two different moments. They were purposefully chosen to represent different teaching experiences, subject areas, and prior familiarity with AI technologies. Both had participated in the initial Delphi Phase of this research, as experts.

Two K-12 educators were selected to participate in the intensive walkthrough process, which was conducted at two different points in time. The first participant is a professor of

Culture and History of Religions, in the 41-50 age range, with nearly two decades of experience as a lecturer and researcher. The second participant is a professor with extensive experience in lecturing future STEM teachers, in the 51-60 age range.

The rationale for selection of the two professors lies in their prior involvement in the Delphi Phase of the study (2019–2020), where they contributed their expertise to the foundational development of the research. This previous engagement ensured that they brought a strong understanding of the course themes and an iterative mindset to the walkthrough process. While they had offered critical insights during the earlier phase, they were not aware of how their input had shaped the course’s evolution, making their fresh evaluation both informed and unbiased.

Ensuring continuity of stakeholder perspective was another key reason for their selection. These professors were not simply new evaluators but long-term contributors who had helped articulate the ethical dilemmas and the importance of agency that would later inform the course design. Their consistent participation allowed for a coherent tracking of how the initial research insights were carried forward and materialised into educational content.

Following the evolution of the intervention with the same participants enabled the research team to monitor how the course emerged from the initial challenge identification stage. Because these individuals were involved from the beginning, they were uniquely positioned to assess the course not as an isolated product, but as the result of a developmental process they had helped initiate.

Providing informed feedback and validating the iterative design process was another rationale for selecting these professors. Their earlier involvement allowed them to evaluate whether the course effectively addressed the core issues surfaced in the Delphi Phase. Their feedback helped determine how well the course aligned with both their original insights and their evolving understanding of ethics and agency. In this way, their evaluations served not only as critiques but also as validations of the research methodology itself.

Going beyond a single discrete evaluation, this approach enabled the study to capture how the course aligned with the broader trajectory of the research project. The professors’ reflections helped reveal not only the current relevance of the course content, but also the necessary updates prompted by ongoing developments in the field, such as the emergence of GenAI and its implications for reframing ethics through the lens of agency.

6.4.1.2.2. Setting and Procedure

The walkthrough exploration was conducted in a blended format, mirroring the course's intended flexibility. The procedure followed for each participant is outlined below.

The first participant explored the course in a blended manner, reviewing the course session by session while simultaneously adopting the perspective of a potential teacher-trainee (the target audience) and a “back-office” trainer. He was encouraged to question any aspect of the course design at any time.

The initial sessions, comprising Module 1 and Module 2 (Sessions 1 to 5), were conducted in person over two days. There was a six-day interval between the delivery of the first three sessions and the fourth and fifth sessions. The final three sessions (Sessions 6, 7, and 8) were explored synchronously via WhatsApp chat. This approach was chosen to accommodate the professor's schedule, allowing for real-time observation and discussion.

The course was presented using the Canva platform as a supportive tool. The full session plan was revealed only at the end of each session, allowing for an assessment of how well it aligned with the participant's lived experience. All feedback from this participant was incorporated, and the course materials were revised accordingly before the walkthrough with the second participant.

Following the participant's initial review of the proposed activities and thematic flow, a collaborative discussion took place in which his expert perspective was central to shaping the micro-evaluation format of the course (see section 6.2, Micro-evaluation Stage). In collaboration with the main researcher of this thesis, five activities were identified as the most effective, impactful, and representative of the course's central themes. Additionally, slight adjustments were made to the session timings to facilitate smoother transitions and to allow more time for critical elements, particularly pedagogical transfer.

The second participant joined the study at a later stage, after all course sessions had been revised based on the feedback provided by the first participant. His role was primarily to evaluate the course as a potential target audience member or recipient, although he did not follow the natural pacing of the course.

He began by engaging face-to-face with the main researcher for the first two sessions, which took place over two days, and continued with the third session on a separate day. Over the following two weeks, he independently explored sessions 4 through 8, providing feedback at the end of his exploration and completing evaluation forms for each module.

The procedure for each walkthrough segment involved the aspects that are presented below.

- **Guided navigation:** the facilitator provided a brief overview of the sessions' objectives and guided participants through the digital resources on the Canva platform at the beginning of the process.
- **Activity engagement:** participants were encouraged to actively engage with the designated activities (e.g., Task Diary prompts, small-group dialogues, scenario analysis, character.ai prompting). While full completion of all written tasks was not requested for the sake of pacing, participants were prompted to articulate their thought processes and initial responses.
- **Think-aloud procedure:** participants were encouraged to verbalise their thoughts, questions, and reactions as they explored the content and attempted activities. This provided rich insight into their cognitive processes and immediate interpretations.
- **Facilitated discussion:** after each major activity or section with the first participant, the facilitator paused the walkthrough to initiate a structured discussion on the relevance, consistency, practicality, and potential effectiveness of the course, identifying points of confusion and insight.
- **Observation and note-taking:** the researcher systematically documented participant interactions, verbal feedback, non-verbal cues, and any technical issues encountered with the digital support tool on Canva.

6.4.1.2.3. Data Collection

Data collection during the walkthrough testing stage relied primarily on qualitative methods, consistent with the developmental and formative goals of the study. Key sources of data included facilitator observation notes, transcripts of facilitated discussions, and participants' reflections on the course's coherence, relevance, and pedagogical value. A significant instrument in the data collection process was the "Expert Feedback Diary", a structured form completed by participants following each module. The diary prompted targeted reflection across several key domains, including user experience (UX) design, pedagogical relevance, potential impact on learners, and overall module coherence.

Responses were open-ended and designed to capture evaluative and reflective commentary. The diary also prompted participants to suggest appropriate durations for each

session. This structured yet open format facilitated the articulation of expert insights in a consistent and analysable manner, supporting iterative course improvements.

All participants provided informed consent prior to data collection, with full awareness of the walkthrough's developmental purpose and how their feedback would inform the refinement of course content and design. Anonymity and confidentiality were maintained in all records and dissemination of findings. While participants granted permission to refer to general elements of their professional roles (e.g., teacher-trainer, subject expert), all personal identifiers were omitted or anonymised in the documentation.

6.4.1.5.1. Key Observations and Consequent Amendments

The walkthrough phase generated rich feedback, providing immediate insights into the course's strengths and areas for iterative improvement. Each module included a structured feedback diary, where participants responded to a series of targeted questions. One of the opening questions for each module was: “What is your first impression of Module 1? (You may comment on how it feels, what stands out, or any immediate thoughts.)”. This prompt aimed to elicit initial, holistic reactions to the module's design, tone, content relevance, and coherence before guiding the participant into more specific areas such as UX design, pedagogy, and learner impact.

An example of the type of response obtained through this instrument is the following reflection from the first participant. In response to this question for module 1, the participant described the session as “establish[ing] a solid foundation for the course's key themes, seamlessly integrating crucial concepts like agency that support further learning progression”. He noted that the module successfully moved “beyond theoretical frameworks”, instead engaging trainees “through interactive and participatory approaches to core discussion topics” and valued the way digital tools were “thoughtfully presented within a broader context rather than making them the central focus”. According to this feedback, the diverse learning activities may encourage “collaborative and intellectually stimulating discussions among participants”.

At the same time, this reflection pointed to a specific area for improvement. The participant observed that the “algorithmic section could benefit from a stronger connection between the main activity and the concluding discussion, particularly in illustrating how the practical exercise relates to the historical example presented”. In response to this input, the activity was entirely redesigned. The participant's reactions during the exploration also indicated that the original task (creating an algorithm based on a practical, everyday scenario, while drawing on

the original ritualistic and sociohistorical character of algorithms as a conceptual reference) could be overly demanding and feel abstract. The task's open-ended nature, according to this expert, could lead to highly diverse responses, making it difficult to connect effectively with the intended reflective objective and concluding discussion. To address this, a new task was designed, and additional facilitator guidance along with an explicit linking question were introduced to strengthen the coherence between the activity and its follow-up discussion.

This example illustrates the developmental nature of the walkthrough process. Feedback was incorporated immediately before the second participant's walkthrough, ensuring that successive iterations of the course reflected user-informed adjustments. Such responsiveness contributed to a more pedagogically aligned and experientially coherent learning environment.

Usability of the Canva Platform. Another area of feedback gathered through the expert feedback diary concerned the user experience (UX) design of the course platform. Participants were asked to reflect on the following aspects: “UX Design (e.g., First impression of layout. Was it easy to navigate? Where did you feel lost? Suggestions for UX.)”. These questions aimed to evaluate both the usability of the Canva platform and the coherence of navigation across modules and resources.

In response, participants generally found the Canva platform intuitive and accessible. One noted that although the interface initially felt somewhat static, it quickly became clear that the course was not intended for isolated, self-paced exploration. Instead, it was designed as a supportive tool for trainers to use in structured, presential, or blended delivery formats, guided by detailed session plans. This realisation aligned well with the course's pedagogical expectations and delivery model.

The integration of diverse media, such as videos, interactive prompts, and external links, was appreciated for enhancing engagement. Collaborative tasks, particularly those involving shared conceptual mapping, were also well-received, with participants highlighting their value in creating a shared learning environment distinct from more traditional LMS.

Regarding limitations, participants pointed out that some activities required user registration for external platforms, which momentarily interrupted the navigation flow. As the first participant explained: “My only moment of hesitation occurred when one of the activity applications requested user registration”. In response, it was recommended that trainers and their groups decide whether to engage with those platforms directly, with the suggestion to

create separate accounts for each course instance or participant group, thus maintaining flexibility without compromising privacy or usability.

Regarding Module 3, the second participant also gave a suggestion in terms of the navigational design: “because this module encourages so much reflection, the questions could be made more prominent or even expandable, so that they feel like natural pauses rather than embedded text”. Additionally, one suggestion from the first participant was the improvement of the navigation structure, by implementing a network of interconnected pages, allowing for smoother transitions and easier access to modular content.

Engagement with Content and Activities. High levels of engagement were observed across several activities, particularly those involving direct interaction with AI tools or ethically complex scenarios. Some examples are presented below.

- “Task diaries”: while requiring dedicated individual reflection, the Task Diary tasks were consistently viewed as valuable spaces for self-assessment and integrating new concepts with personal teaching visions, experiences, and contexts, facilitating insight, generalisation, and transfer.
- “Visual timelines” (Session 1): the historical perspective on EdTech imaginaries resonated, helping educators contextualise current AI trends within a broader narrative. Furthermore, participants were captivated by the unexpected dynamics and results achieved through the use of a collaborative tool like Miro, which they both felt held significant potential for project-based learning.
- “character.ai prompting” (Session 2): this activity generated significant interest and engaged reflections. Both experts found the simulated dialogues with a pre-set AI figure to be a powerful means of exploring the nuances of human agency in relation to AI. The direct, experiential nature of interacting with a Generative AI within a guided framework helped bridge the gap between abstract concepts and practical implications.
- "Safety vs. Surveillance debate" (Session 3): the semi-structured debates on ethical dilemmas like facial recognition or AI grading were felt as highly dynamic. Participants said that this activity had strong potential for eliciting critical reasoning and perspective-taking.
- "Prompting as a Practice of Authorship" (Session 6): this activity was also highlighted as having strong potential to generate engagement. Participants found the task of designing purposeful AI prompts for their specific curricular topics (e.g., science,

history, literature) highly relevant, as it helped foster meaningful connections with the subject matter. They noted that it encouraged teachers to create opportunities for students to address the alienation often associated with working with Generative AI, by fostering intentionality and promoting deeper engagement with the content.

- "A pedagogical design task (shaping agency with AI)" (Session 7): this culminating activity was perceived as highly relevant and integrative. It could be perceived as somewhat abrupt if the course had simply culminated in the sharing of that role-taking task. However, the first participant emphasised the importance of having linked that activity to the broader goal of creating a community of practice. The transition, outlining concrete steps for continued work, was recognised as an effort to foster collective agency. At the same time, in cases where an engaged community might not be sustained, it was seen as planting the seeds for teachers to build such communities within their own school contexts.
- Conversely, participants cautioned that an activity like the "Hashtag Reflection" (Session 3) required clarification regarding the expected depth of critical engagement beyond simple social media tagging. This led to a refinement of the prompt to emphasise analytical synthesis rather than merely descriptive responses.

In addition to assessing activity design, participants were asked to consider the impact on students, responding to the questions: "How does the session align with students' needs? In what ways might students benefit indirectly?" The first participant offered a particularly insightful reflection, stating:

I believe that deeper teacher reflection on AI and its electronic tools directly enhances their appropriate implementation, ultimately benefiting students. Understanding when and how to deploy these technologies, recognising their development processes, and considering their social, anthropological, ethical, and political implications provides significant pedagogical value. This comprehensive awareness strengthens teachers' sense of agency within educational dynamics, positioning them as informed participants rather than passive technology consumers. (Participant 1)

The second participant noticed that: "When teachers themselves explore such issues, they are better prepared to create classroom environments that honour student needs, vulnerability, and growth. (...) the framing of agency as both social and mental (...) reminds us that the way students perceive their own power within digital interactions matters".

These responses summarize how teacher-facing activities within the course were seen as indirectly but meaningfully supporting student learning. By exploring teachers' epistemic and collective agency, the course aimed to encourage more intentional pedagogical choices, choices that, in turn, help students shape more relational, dialogical, and ethical experiences.

Clarity of Pedagogical Intent. Both participants generally grasped the course's core concepts, including the typology of agency and Biesta's educational purposes. The emphasis on relational learning and resistance to "automated scaling" was clearly articulated and appreciated by the experts. The second participant expressed concerns about the depersonalisation often associated with technology-driven educational reforms. The course's non-traditional assessment approach (self, peer, facilitator feedback, qualitative rather than quantitative metrics) was perceived as refreshing and aligned with the values of professional development on such matters as agency and ethics. This coherence across content and form was highlighted as one of the most appreciable features of the course. Both walkthroughs confirmed that pedagogical integrity was maintained, and the "why" behind the course's design choices was largely understood.

One example of this understanding emerged from the response to the pedagogical reflection prompt: "Is this content relevant to your teaching? Does it support teacher agency? What was pedagogically strong or weak? How did it connect with your role as teacher?" Reflecting on Module 1, the first participant wrote:

At this stage, addressing this question is challenging; pedagogically, the first module prompted my reflection on teachers' autonomy and agency in their educational role. It helped me contextualise the often idealised or 'fossilised' visions of technology developers against the richer, more nuanced reality teachers face in actual classroom settings. This juxtaposition highlighted how educators' broader, humanistic mission often transcends the narrower frameworks envisioned by those designing electronic tools. (Participant 1)

This early-stage reflection shows how the course may foster critical awareness, particularly in relation to the contrast between top-down technological visions and lived pedagogical realities.

Later in the walkthrough process, in response to Module 3, the second participant demonstrated a very applied understanding of pedagogical design. Responding to the same guiding questions, he wrote:

Very much so. The emphasis on relational pedagogy resonates deeply with my own teaching values. In an age where there's mounting pressure to integrate technological tools into everything we do; this module provides much-needed language and structure to support critical choices. The role-play scenarios were pedagogically strong, especially in how they layered the complexity of responses, from punitive to constructive. It offered a non-dogmatic way to reflect on our own instincts as educators.
(Participant 2)

This response reflects a shift from conceptual awareness toward applied reflection and pedagogical decision-making, which aligns with the focus of the module under analysis. While the first quote is rooted in a critical recognition of systemic issues, the second builds on that by illustrating how course design choices supported the participant's personal pedagogical stance and classroom practice. Together, these considerations demonstrate that the course's pedagogical intent supported an evolving trajectory of reflection: from critique to intentional and systematic planning.

Feasibility and Adaptability. The estimated time required for the activities is considered feasible, although certain discussions (e.g., Session 2, “Agency, AI, and Shared Responsibility”) are expected to naturally extend due to the depth of potential participant engagement. The modular design and well-defined resource base were seen as highly adaptable. The proposed session timings were interpreted as indicative of the expected level of engagement and depth, rather than as rigid constraints.

Session seven was considered more demanding. Consequently, the expectations for the outputs of two of its main activities were lowered. This adjustment aimed to ensure that the activities primarily served to explore the capacity for maintaining affective dimensions of learning through lesson planning, even in contexts where automation and algorithms are present. This concern also related to the need to keep trainees/participants engaged by the end of the session, especially when a challenging and culminating task, such as preparing an integrative lesson plan, was requested from them. This also signified the need to propose more time between the seventh and final session, allowing teachers to deeply prepare their lesson plans, receive and give feedback to colleagues within a logic of collectively nurturing their community, and implementing their sessions in their schools.

Areas for Refinement. Based on the observations and feedback, several specific refinements were implemented. Those are documented below.

- Clarification of prompts: adjustments were made to the wording of certain prompts, particularly for activities requiring complex analytical thought, to ensure explicit instructions. Furthermore, a more impartial language was proposed in some activities, as a way to avoid transmitting any critical nuance before teachers engaged in their own exploration and reached their own conclusions.
- Additional examples/scenarios and conceptual presentations: more diverse examples were integrated into discussions, particularly for the application of agency types and educational purposes across different subject areas and grade levels. Additionally, a conceptual presentation of key terms was proposed. This change from the initial approach, which began with tasks designed for participants to gain insights independently before encountering theory, was made to provide an initial explanation for concepts like “agency”. This aimed to ensure teachers felt more secure in their understanding and less uncomfortable when concepts were quite new to them, thereby encouraging deeper exploration.
- Facilitator guidance: the session plans were integrated into the opening slide of each session, making the planning (1) easily accessible, (2) helpful in giving a clear overview of what would be explored next, and (3) supportive of a deeper understanding of each session. Beyond simply describing the activities, the plans also outlined the pedagogical strategies and learning processes involved, helping participants grasp the kind of engagement expected from them.

Technical Optimisations: minor adjustments were made to resource embedding within Canva to improve loading times and mobile responsiveness. For example, one video was directly embedded in the platform instead of relying on external online streaming. The walkthrough testing phase provided invaluable empirical grounding for the iterative refinement of "From Automation to Autonomy". The qualitative insights directly informed revisions to content, activities, and facilitator support materials, reinforcing the EDR cycle's commitment to continuous improvement. This preliminary, in-depth qualitative assessment confirmed the course's potential to deliver a nuanced, agency-focused professional learning experience, while also highlighting specific areas where additional clarity or scaffolding could enhance participant engagement and learning outcomes. It validated the deliberate choice to avoid a traditional LMS

platform and reinforced the emphasis on relational, dialogic learning as central to the course's success, consistently aligning form, content, and intentionality.

6.4.2. Micro-evaluation Stage

This section details a small-scale implementation trial that complements the expert walkthroughs, providing key insights from the perspective of the target user group. The trial aligns with the EDR's phase "Prototyping and Assessment of Preliminary Products & Theories" McKenney and Reeves' (2018) EDR methodology, specifically the phase titled. It offers early evidence regarding both the conceptual uptake and pedagogical relevance of the intervention.

The micro-evaluation was conducted by the main researcher of this thesis on June 12, 2025, with a group of ten K–12 teachers (four men and six women). Participants were naturally selected through an opt-in process during a professional development congress for teachers. The event offered career progression credit points and allowed participants to choose from one of three workshops: "Artificial Intelligence, Educational Imaginaries and Human Agency", "Mindfulness Strategies for Educators", and "School and Neurodiversity". The group evaluated in this trial voluntarily enrolled in the first workshop, "Artificial Intelligence, Educational Imaginaries and Human Agency", which corresponds to the course under study.

6.4.2.1. Methodological Rationale

The micro-evaluation adopted a mixed-method qualitative approach, grounded in formative and exploratory analysis. Conducted over a single day, the course spanned four hours and was structured into two separate 2-hour sessions, one in the morning and one in the afternoon. Five key activities were strategically selected for this trial; each aligned with a core theme of the course and identified in consultation with the lead expert from the initial walkthrough phase (cf. <https://docs.google.com/presentation/d/1LYpBxC3tJRb3T-w0bzZ7HmXLvWDeLWw/edit?usp=sharing&ouid=116166807221840325957&rtppof=true&sd=true>). These activities were chosen for their thematic diversity, conceptual depth, and representational value in reflecting the natural progression of the course. Minor adjustments were made to the session timings to ensure a coherent flow and to allocate more time to the final theme, which focused on pedagogical transfer and application.

The themes explored were:

- Educational Imaginaries, Socio-technical Technologies, and Futures (Session 1)

- Purposes of Education: Subjectification, Socialisation, Qualification (Session 2)
- Human Agency and AI: Subjective, Relational, Collective (Sessions 5 and 6)
- Pedagogical Strategies for Distributed Agency (Session 7)

Data collection relied on a brief post-session questionnaire (cf. <https://docs.google.com/forms/d/e/1FAIpQLSfCCkA0rSdribsjZyVrM5q0n3iz1suw0xg3sA6A2o1C2AsvxQ/viewform?usp=sharing&oid=107376701533573656457>), shared via QR code, designed to explore three key evaluative dimensions, as outlined below:

- **Conceptual Coherence:** exploring whether participants grasped the theoretical foundations of the course, particularly the concept of agency, and whether these were perceived as accessible and relevant to their educational contexts. Sample items included: "Which of these themes was most striking for you?"; "Which of these themes would you like to explore further because it is more challenging/demanding?"; "What was the most relevant idea or reflection during the session?"
- **Formative Potential:** assessing the workshop's ability to spark curiosity, engagement, critical awareness, and reflective thinking. Sample items included: "Which activity was most relevant for you?"
- **Recognition of Audience Needs:** evaluating how effectively the course addressed participants' perceived professional needs in the context of AI and education. Sample items included: "What is the relevance of a course like this, focused on AI, Human Agency and Education?"

The questionnaire combined both closed- and open-ended questions, enabling the collection of quantitative data (e.g., response frequencies) alongside narrative reflections that provide deeper insight into participant perceptions.

6.4.2.2. Data Analysis and Findings

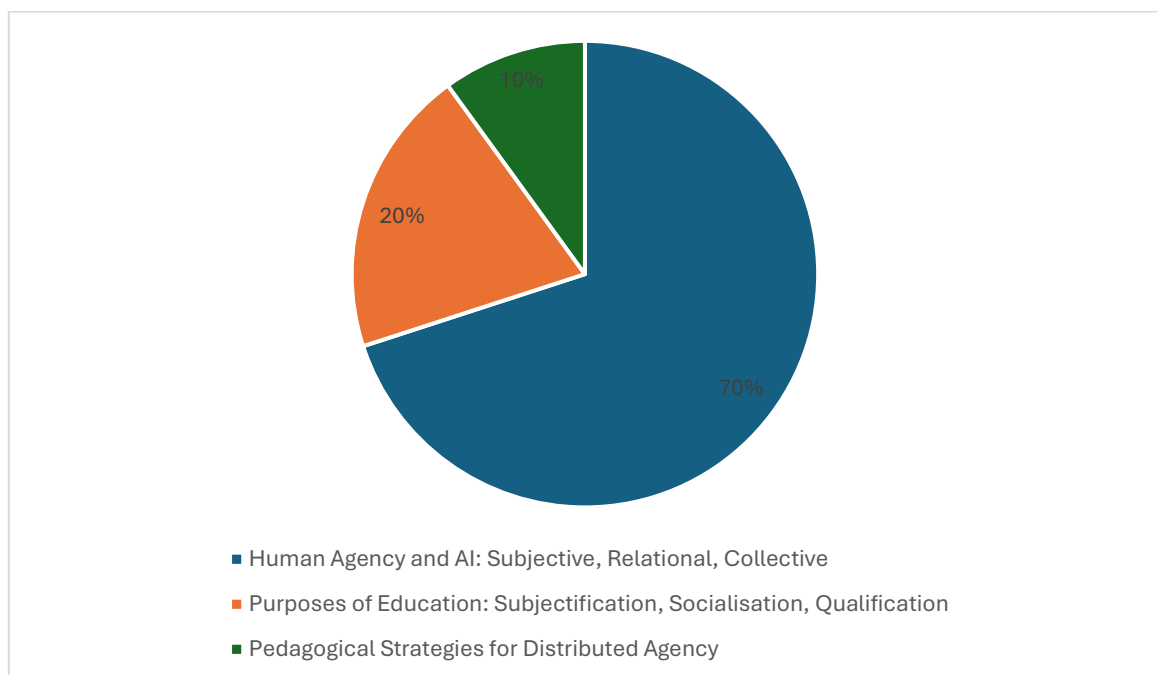
This section presents an integrated analysis of the micro-evaluation results, drawing on both closed- and open-ended questionnaire responses. Rather than isolating the three evaluative dimensions (conceptual coherence, formative potential, and response to audience needs), the analysis synthesises them into a cohesive interpretation. This approach reflects the inherent continuity in participants' experiences, in which theoretical and conceptual engagement, pedagogical relevance, and professional development were perceived as reciprocally shaping one another.

- First Question

In response to the item “Which of these themes was most striking for you?” (Figure 15, below), the majority of participants (70%) selected “Human Agency and AI: Subjective, Relational, Collective” as the most relevant theme. This was followed by “Purposes of Education: Subjectification, Socialisation, Qualification”, chosen by 20% of respondents, and “Pedagogical Strategies for Distributed Agency”, noted by 10%. Notably, the theme “Educational Imaginaries, Socio-technical Technologies, and Futures” did not receive any selections for this item.

Figure 15

Distribution of Responses to the First Question: Most Striking Theme



- Second Question

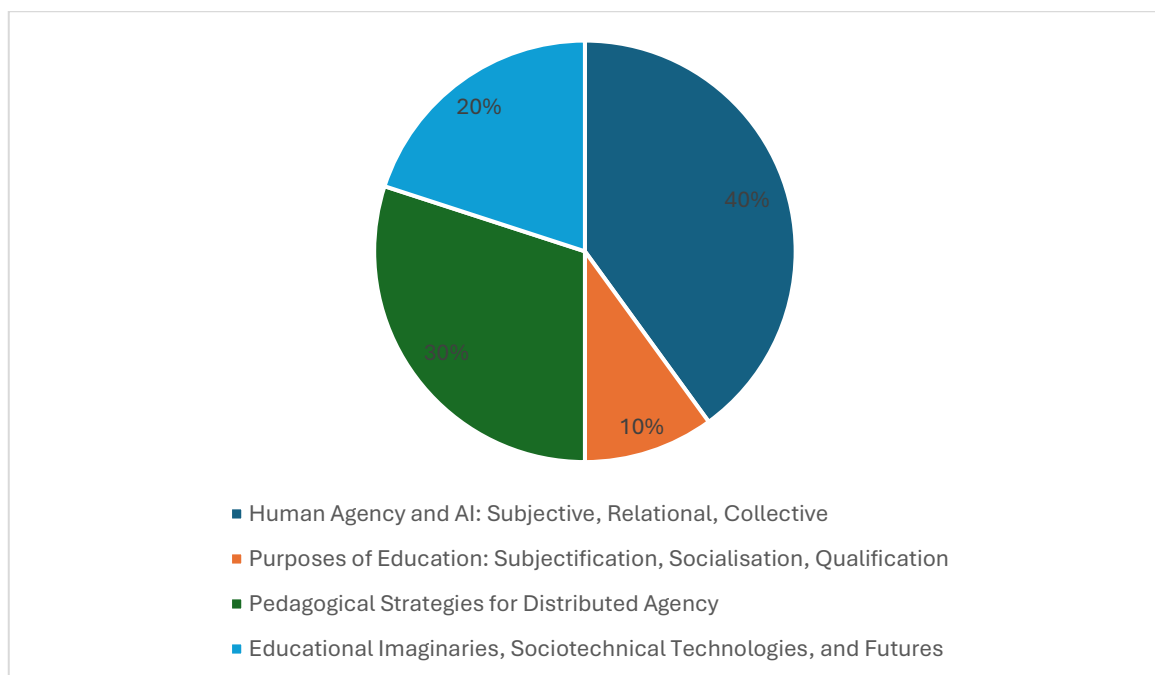
In response to the second item, “Which of these themes would you like to explore further because it is more challenging/demanding?” (Figure 16, below), 40% of participants identified “Human Agency and AI: Subjective, Relational, Collective” as the most demanding theme. This result reinforces its dual role as both striking and conceptually challenging, indicating its novelty and relevance as a critical reflection point within the AI and education landscape. “Pedagogical Strategies for Distributed Agency” was identified by 30% of participants as requiring further exploration, suggesting a perceived need to develop practical skills for translating theory into classroom application. This underscores the need for

structured, experiential learning environments, such as pedagogical role-play, that not only provide a safe space for professional experimentation but also scaffold the development of transferable, practice-oriented competencies.

Interestingly, while “Educational Imaginaries, Socio-technical Technologies, and Futures” was not selected as the most striking theme in the previous question, 20% of participants cited it as one they would like to explore further due to its conceptual complexity. Lastly, “Purposes of Education: Subjectification, Socialisation, Qualification” was selected by 10% of respondents as an area needing deeper engagement.

Figure 16

Distribution of Responses to the Second Question: Theme for Further Exploration



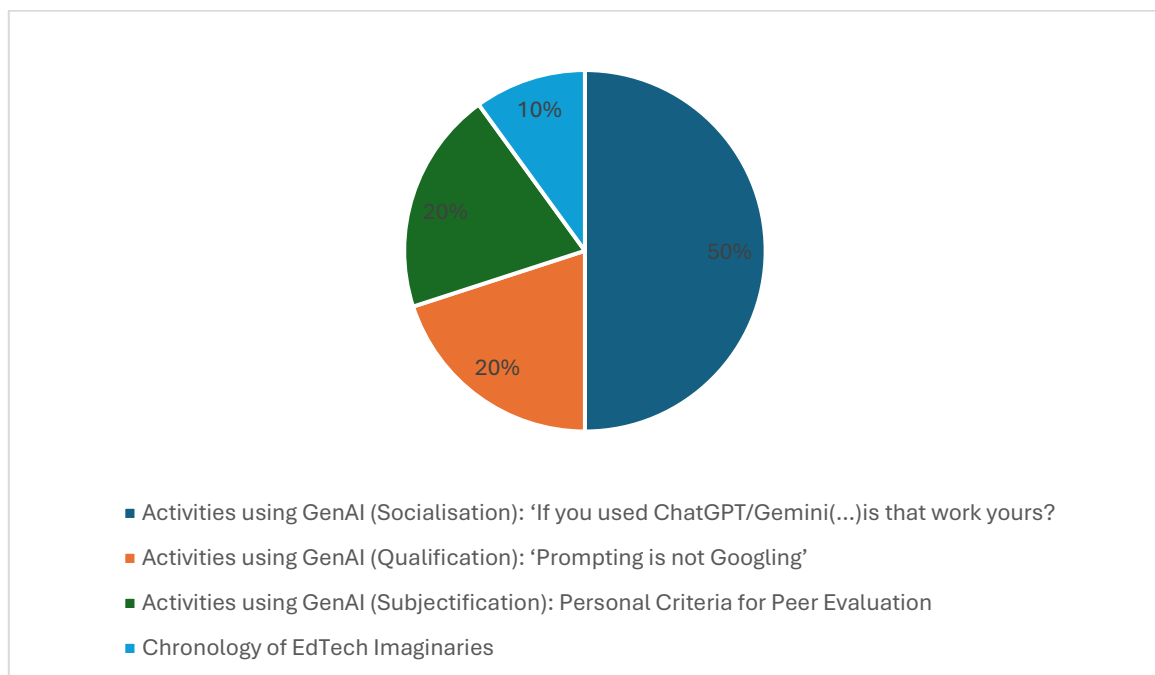
- Third Question

In response to the third item, “Which activity was most relevant for you?” (Figure 17, below), 50% of participants selected the activity “Activities using Generative AI (Socialisation): ‘If you used ChatGPT/Gemini to produce your work, is that work yours?’” as the most meaningful. This was followed by two other classroom-oriented activities: “Activities using Generative AI (Qualification): ‘Prompting is not Googling’” and “Activities using Generative AI (Subjectification): Personal Criteria for Peer Evaluation”, each chosen by 20% of the respondents.

These findings indicate that participants highly valued pedagogical activities directly applicable to their educational contexts. The preferred activities emphasised role-playing, ethical reflection, and hands-on engagement with Generative AI, all closely tied to subject matter. This reflects a clear preference for approaches that connect conceptual exploration with practical classroom use. In contrast, more abstract tasks, such as “Chronology of EdTech Imaginaries” (10% of responses) and “Balls and Balloons” (which received no selections), were seen as less relevant. While participants recognised the importance of theoretical foundations, they were more engaged by activities that allowed them to simulate affective-cognitive processes and envision their translation into practice.

Figure 17

Distribution of Responses to the Third Question: Most Relevant Activity



- Fourth Question

Regarding the open-ended questions, responses to the fourth item, "What was the most relevant idea or reflection during the session?" revealed several key themes. Twenty percent of participants highlighted the concept of human agency, while another 20% emphasised the importance of critically reflecting on the application of AI in education. Additional responses (each representing 10%) pointed to the ethical use of AI and the broader significance of authorship in AI-mediated learning environments, particularly through the crafting of

meaningful prompts. These prompts, guided by authorship criteria, not only foster alignment with subject matter but also re-engage students with disciplinary content and revive their enthusiasm for the learning process itself.

These reflections closely align with the patterns identified in earlier items, reinforcing the course's conceptual and pedagogical relevance. The prominence of human agency as a key takeaway reflects its identification as both the most striking and most challenging theme in the closed-ended questions, suggesting participants' strong engagement with this central concept. Similarly, teachers' emphasis on critical reflection regarding AI mirrors the earlier finding that activities promoting ethical inquiry and role-playing were considered most relevant. The focus on authorship and prompt crafting as meaningful pedagogical practices further indicates that participants are not only grasping theoretical constructs but are actively exploring how to apply them in their teaching contexts. Taken together, these insights suggest that conceptual understanding, ethical awareness, and pedagogical action are closely interwoven, confirming the course's formative value.

- Fifth Question

In response to the fifth open-ended question, "What is the relevance of a course like this, focused on AI, Human Agency and Education?" all participants underscored the course's importance. Sixty percent described it as "very interesting", "important nowadays", or even "indispensable" and "urgent". Additional responses pointed to specific rationales for its relevance: 10% emphasised the need "to update ourselves to work better with students", 10% highlighted the value of exploring "new practices and other learning modalities", another 10% noted the importance of adopting a critical stance toward the "massification and sophistication" of AI, and 10% stressed safeguarding "memories and human reflections" in educational contexts. This response reinforces the findings from the previous items by affirming both the conceptual relevance and pedagogical urgency of the course from the participants' perspective.

The strong consensus on the course's importance, described as urgent, indispensable, and very interesting, mirrors the high engagement reported earlier, particularly with themes like human agency and the role of AI in education.

Participants' emphasis on the need to "update ourselves to work better with students" aligns with earlier mentioned preferences for practical, classroom-applicable activities. This reinforces the notion that teachers are seeking not only theoretical insight but also opportunities for exploratory tinkering with AI in their professional practice. The mention of new practices

and other learning modalities points to pedagogically meaningful innovation, while the call for a critical stance toward the massification and sophistication of AI echoes previously noted concerns about its ethical implications, particularly those related to authorship, prompt crafting, and reflective engagement.

Finally, the reference to safeguarding “memories and human reflections” signals a strong commitment to preserving humanistic values in education. This aligns with the course’s central theme of distributed human agency and reveals a deeper engagement with the philosophical dimensions explored. Taken together, these responses indicate that the course activated not only teachers’ cognitive engagement with the subject matter, but also their affective responses, personal ethical stances, and an interest in behavioural transformation. This was reflected not only in their critical comments, but also in reflections that pointed to more comprehensive and emotionally grounded dimensions, such as the value of memory.

6.4.2.3. Discussion: Interpreting the Micro-Evaluation through the ik-model

The findings of this micro-evaluation are interpreted through the lens of the ik-model, developed by Mouta et al. (2015), which builds on the TPACK (Mishra & Koehler, 2006) framework by introducing diachronic dimensions to better understand technological integration in education. The use of the ik-model is methodologically coherent, given that the same model was used during the design phase of the course (Mouta et al., 2025b). Its reapplication in the evaluation phase allows for a consistent analytical lens that links course conception, implementation, and analysis. This alignment strengthens the interpretive validity of the study and supports a comprehensive understanding of how participants engaged with the course across conceptual, relational, technological, and processual dimensions. By highlighting the evolving and context-sensitive nature of pedagogical relationships and processes, this framework provides a deeper analytical structure for exploring how educational interventions respond to the complexity of organic educational contexts.

The model comprises four interrelated domains (technological, content, relational, and processes) and offers a comprehensive lens for evaluating how training initiatives align with educators’ professional needs, the socio-technical landscapes of schooling, and the evolving demands placed on pedagogical identity, epistemic agency, and collective practices as AI enters educational contexts. Its use in this study is particularly appropriate given the nature of the intervention: a professional development course designed to support educators in developing collective agency in the face of socio-technical transformation. Unlike frameworks

that prioritise static knowledge or predefined competencies, the ik-model emphasises the reciprocity of pedagogical relationships and learning processes as they evolve over time. In this context, it enables a comprehensive analysis of how the workshop not only initiated a disposition toward developing conceptual and technological understanding, but also opened space for reflection on professional roles, collaborative meaning-making, and the enactment of meaningful learning experiences.

The model's temporal and situated orientation resonates with the broader aims of teacher education, allowing the evaluation to move beyond superficial measures of satisfaction or utility and to reveal the extent to which teachers began to reconsider their pedagogical identities, ethical positions, and collective agency.

All four domains of the ik-model were organically activated in participant responses, confirming the pedagogical robustness of the training prototype.

- Technological Domain

The technological domain was addressed through participants' interaction with generative AI tools such as ChatGPT, Gemini, and Mentimeter. The activity "Activities using Generative AI (Socialisation): 'If you used ChatGPT/Gemini to produce your work, is that work yours?'" selected as the most relevant by respondents, reinforces the importance of scaffolding teacher practices through hands-on applications that are both accessible and ethically grounded. Such activities offer meaningful entry points into the pedagogical use of AI without imposing excessive workload, while simultaneously promoting ethical awareness (Cukurova et al., 2023). They also represent opportunities to exercise collective agency at a higher level. Engaging with these technologies in a tinkering, exploratory, and diachronic manner (Bardone, 2024) during training allows teachers to actively participate in reinterpreting and reconfiguring the educational role of AI tools, eventually contributing to their concretisation within practice (Simondon, 2008).

This approach takes place within dialogical pedagogical practices and can be extended to teachers' own learning environments. Tinkering aligns with situated learning theories, as it supports low-stakes experimentation and leaves room for improvisation. In this sense, tinkering provides a consistent and embodied way to exercise agency. It enables teachers to demystify complex technologies by creating meaningful space for exploration within new tools and digital landscapes. Tinkering also invites reflection and negotiation, which are two hallmarks

of collective agency, and aligns closely with the principles of educational design research and transformative learning, where practice is continuously revised in light of emerging insights.

- Content Domain

The theme “Human Agency and AI: Subjective, Relational, Collective” emerged as both the most striking and the most challenging for participants. This dual recognition reveals a critical pedagogical threshold: while the theme was intellectually stimulating and clearly relevant to teachers’ professional realities, it also proved sufficiently challenging to require further support and inquiry. This positioning suggests the topic entered participants’ zone of proximal development, generating interest and conceptual engagement while simultaneously indicating the need for structured scaffolding. These findings highlight the importance of guided, collaborative learning environments over self-directed approaches, especially when dealing with epistemologically and ethically layered concepts such as agency and AIED. Rather than oversimplifying theoretical content, teacher education on AI, ethics, and agency is expected to focus on pedagogical strategies that bring complex ideas to dialogic, reflective, and practice-oriented environments, enabling educators to meaningfully integrate these concepts into their evolving professional identities and classroom practices.

The role-playing activity on automation and responsibility, for instance, illuminated tensions around attribution, authorship, and the public nature of AI-generated content. The eliciting questions were around: "Is it the same to copy from a classmate, a book, or use AI?"; "If AI generated the text, would copying it and claiming it as mine be plagiarism? Why?"; "Who owns the authorship if I only interacted with the machine?"; "Am I able to speak differently about my responsibility depending on the feedback (positive or negative)?"

These questions provided a concrete entry point for connecting abstract themes, such as socio-technical imaginaries introduced earlier in the course, to real-world implications of AI in education, particularly in how those imaginaries materialise in data practices. This connection revealed the intentional learning continuum embedded in the course design. While more conceptual activities were not identified as the most preferred, they proved key in laying the cognitive and ethical groundwork for meaningful engagement with the hands-off components. In this way, participants were able to perceive the relevance and coherence of experiential tasks, which were enriched by the conceptual scaffolding established earlier in the workshop.

- Relational Domain

The relevance of the relational domain emerged through participants' preference for activities involving role-play, which proved particularly effective in surfacing tensions around differing attitudes toward the integration of agentic technologies in educational spaces. These exercises illustrated not just technical concerns but deeper ethical and relational implications. With the integration of generative AI, participants noted that difference, context, and dissonance tend to be diminished; conversations are often recentred in ways that appear homogenised, losing their richness and situated meaning. This concern aligns with Latour's (2005; 2014) notion that technologies should be recognised as actants (entities that influence human actions) before they solidify as actors, with fixed roles in socio-technical networks. Education, as a relational and symbolic practice, is therefore expected to maintain a critical stance toward AI, identifying its formative impacts early in its adoption trajectory. Zuboff (2019) similarly warns that dealing with such "unprecedented" configurations, requires heightened ethical imagination, vigilance, and democratic accountability.

In this light, the notion of agency takes on renewed meaning. Being agentic is not about asserting full control; rather, as one participant insightfully expressed, it is about being engaged in the process of reflection and debate, even when it is difficult, uncertain, or ambiguous. This stance repositions agency as inherently relational and processual, requiring continuous negotiation, dialogue, and situated interpretation (Archer, 2014, 2015), where professional development becomes less about acquiring fixed skills and more about a shared process of ethical becoming within institutional and interpersonal dynamics (Damşa, 2014).

One participant's reflection on the role of memory as central to the learning experience reconnects with a broader conversation held during the session, highlighting the importance of considering the relational axis as a core component in the intersection of AI and education. These symbolic tensions are not marginal; they shape how educational meaning is made and transmitted. They begin in language and extend across symbolic and epistemic processes. Several key tensions were identified: first, the opposition between history and testimony. While AI can reconstruct historical facts, it lacks the capacity for testimony, the personal, embodied, and ethically anchored recounting of lived experience. Without testimony, history may be stripped of its affective and moral dimensions. Second, the contrast between information and memory.

Although AI excels in storing and retrieving vast quantities of information, serving as external supports for human memory, or what Stiegler (1998) refers to as tertiary retention, it

does not replicate memory, which is inherently tied to emotion, affect, and narrative. AI can also offer interpretation through pattern recognition, but meaning arises only through human experience and contextual resonance, giving symbolic processes their depth and significance. Lastly, the tension between networking and forming bonds. While AI can facilitate networking (e.g., through recommendation-driven social platforms or matching and grouping tools), it does not cultivate bonds, the durable, affective ties that form the basis of communities of care, shared responsibility, and collective agency.

By enabling participants to rehearse, reflect, and negotiate meanings collaboratively, the workshop activated not only individual awareness but also a sense of shared ethical responsibility and pedagogical imagination. In doing so, it affirmed the importance of relational learning environments where educators can collaboratively engage with the socio-technical transformations shaping their professional lives.

- Processes Domain

In relation to the process domain, the course explicitly embedded processual dimensions into each activity's session plan, offering more than just targeted learning objectives. It provided a structured, diachronic vision of the affective, cognitive, and behavioural transitions anticipated to unfold throughout the course and, by extension, the workshop. The workshop was intentionally designed to engage with the often-overlooked liminal spaces of teacher learning, those characterised by uncertainty, resistance, and ambivalence toward change. This intention was validated by participants' appreciation for scenario-based, debate-driven, and prompting activities, which enabled them to explore ethical reasoning through situations that mirrored their own educational contexts, concerns, and needs. Such engagement not only stimulated reflection but also fostered a sense of agency through experimentation, an approach that research has shown can positively reinforce agency itself (Obhi et al., 2012).

By embedding scaffolded, experiential, and dialogic activities throughout the session, the training aimed to foster a supportive environment in which teachers could collaboratively construct their own inquiries and explore the complexity of AI-related themes in ways that felt personal, meaningful, and professionally relevant. Rather than confronting teachers with abstract theory or merely instrumental tools, the course offered hands-on, reflective experiences, such as an embodied activity using a ball and a balloon, that allowed teachers to grasp the concept of agency through physical interaction. This playful, low-risk, and peer-supported environment enabled them to experiment with new concepts and pedagogical

imaginaries in an engaging and meaningful way, ultimately making the concept of agency feel highly relevant to reflecting on educational practices with AI.

This analysis offered a multilayered understanding of how the technological, content, relational, and processes domains were activated in the workshop. The model's diachronic and relational orientation not only enriched the interpretation of participants' experiences but also validated the coherence between the course's epistemological foundations and its pedagogical enactment. The findings revealed that teachers are not merely seeking technical proficiency, but rather ethically and professionally meaningful ways of integrating AI into their pedagogical imaginaries and narratives, being attuned with their evolving professional identities.

6.5. Conclusion

This chapter presents the conception, development, and evaluation of a professional development course that represents the culminating outcome of the EDR project conducted throughout this doctoral thesis. Designed to address the evolving intersections of Artificial Intelligence, human agency, and education ethics, the course synthesises theoretical inquiry, iterative design processes, empirical grounding, and context-responsiveness employed throughout the entire research journey.

The chapter traced the step-by-step articulation of the course, its modular structure, pedagogical rationale, activity design, forms of assessment, and evaluative walkthrough phases, demonstrating how these components were deliberately constructed to support meaningful integration into actual educational settings. Particular attention was paid to the alignment between conceptual frameworks and pedagogical action, ensuring that key ideas such as distributed agency, ethical reflection, and socio-technical critique were not simply addressed as content but embedded into the structure and logic of the learning experience itself.

The evaluation of the workshop provided critical insights into how the course fosters agentic pedagogical practices across the technological, content, relational, and processes dimensions, in ways that meaningfully resonate with teachers' professional contexts. This alignment between theoretical constructs and practical uptake reflects the core aim of design-based research: to bridge the divide between knowledge production and its contextualised application.

Importantly, the course design actively engages with the affective and liminal dimensions of teacher learning. It recognises that professional development, especially amid the rapid integration of AI, is not merely a rational or technical process, but one that involves discomfort,

ambivalence, and the need for reorientation. Through scenario-based activities, embodied metaphors, and dialogic encounters, the course is able to foster forms of engagement that allow teachers to explore this uncertainty in generative ways. These experiences support not only conceptual understanding but also the lived enactment of agency, fostering positive attitudes toward the unprecedented challenges posed by generative AI, particularly in relation to the educational dimensions of subjectification, socialisation, and qualification.

Some limitations must be acknowledged. The course was delivered in a condensed format, intended as a preliminary phase within the broader EDR process prior to scaling. As such, this version provides only a snapshot of the course's potential impact, while indicating possibilities for sustained engagement through the building of a community of practice and follow-up. Additionally, although the activities were carefully designed to promote transversality and context-sensitive reflection, the preliminary findings stem from a small, self-selected group of participants, which may limit the generalizability of results.

The course presents a promising prototype for fostering ethically grounded, critically engaged, and pedagogically situated approaches to professional development in education ethics, human agency, and AI. By planning for actual use, the course also reflects a fundamental principle of EDR: the necessity of designing with future implementation and adaptability in mind. Its modular, replicable format, combined with the layered scaffolding of concepts, activities, and assessment strategies, allows it to be adapted across contexts while maintaining coherence with its core ethical and pedagogical objectives.

Ultimately, this work affirms that AI in education must not be treated as a technological challenge, but as a profoundly pedagogical and ethical one: one that calls for agentic collectives and cultures of engagement. This means recognising not only who is acting (collectives), but also how action is cultivated and sustained (cultures), through environments, norms, and ongoing practices that support meaningful participation and critical reflection.

The course developed throughout this research offers both a practical contribution to teacher education and a theoretical proposition for how educators might be supported in experiencing these new terrains, not through prescriptions, but through participation; not by reducing complexity, but by learning to live and act within it. It is in such dialogic, situated, and ethically attuned environments that agency becomes the ground from which ethics can emerge.

CHAPTER SEVEN

Conclusions, Implications, and Future Directions

7.1. Conclusion

The integration of Artificial Intelligence into educational ecosystems presents a complex and rapidly evolving landscape, characterised by ethical ambiguities, significant pedagogical implications, and opportunities to shape our collective agency in its development. Scholars have argued that there is a pressing need for curriculum reform that promotes critical thinking and prepares educational communities to explore the benefits of these technologies while mitigating potential risks (García-Peñalvo, 2023). In response, this doctoral dissertation undertakes a critical inquiry into this emergent domain, grounded in a conceptual framework that contends preserving the core purposes of education (qualification, socialisation, and subjectification) requires an agency-centred approach. Far from being a neutral innovation to be simply adopted at scale or meaningfully integrated, AI, particularly Generative AI, introduces novel forms of agency into educational contexts, demanding for a renegotiation of meanings, purposes, and pedagogical practices. This research was thus conceived to address a critical gap: the absence of comprehensive, ethically informed, and agency-oriented frameworks that support K–12 educators in articulating their role, the purpose of schooling, and the aims of education in the context of AI.

Educational design research provided a uniquely suited framework for addressing the research questions posed, namely: the identification and critical examination of the ethical considerations involved in integrating socio-technical AI systems into educational settings; the co-design of a conceptual framework that bridges AI, ethics, and educational agency to inform pedagogical practices that uphold and strengthen human agency; and the development and implementation of a professional training course for K–12 teachers, grounded in ethical inquiry and responsive to the complexities of teaching and learning with AI. EDR participatory, iterative, and practice-oriented nature allowed for a continuous dialogue between scientific investigation and the systematic design of the course. This methodological choice proved foundational, as it enabled the co-construction of ethical meaning in action and ensured that the resulting intervention was collectively negotiated and meaningfully aligned with the voices of educational actors and the specificities of their local contexts.

The thesis progressed through a series of interconnected phases, each building upon the preceding one to deepen understanding and cumulatively inform the production of two key outcomes: (1) a framework for problematising and understanding agency in the context of AI in education, and (2) a professional development course for K–12 educators focused on AI,

ethics, and agency. Each research phase is further elaborated in relation to its corresponding research question, outlining specific objectives, methodologies, and key contributions to the overall progression of the study.

Addressing Research Question 1: What are the ethical considerations associated with the integration of socio-technical AI systems in educational settings?

The initial phase of this thesis comprehensively explored the ethical considerations arising from the integration of socio-technical AI systems in educational settings. Chapter 2, "State of the Art: Systematic Literature Review (SLR)", presented the research landscape through a systematic literature review conducted in accordance with PRISMA guidelines. This review explored the ethical dimensions of AI in education, analysing research from 2011 to 2022, critically examining its applications while identifying gaps in the existing literature.

The study focused on several key aspects:

- How ethical concerns, referring to cultural differences, inclusion, and emotions, have been addressed.
- The extent to which capacity-building efforts, ethical guidelines, and frameworks exist for educators.
- The role of teachers in ensuring ethical AI implementation in educational settings.
- Outcome: synthesised overview of the current state of research.

The review explored a theoretical model to rethink education ethics in the context of AI and identified significant blind spots in existing AIED ethical frameworks. These include: (1) epistemic ambiguities, where many frameworks presume a universal ethical stance without clarifying whether it is utilitarian, deontological, virtue-based, or dialogic; (2) theoretical fragmentation, with conceptual overlaps and vague boundaries between ethical principles, educational aims, and practical digital competencies; (3) implementation gaps, as many policies are developed at macro levels with little actionable guidance for schools, municipalities, or educators; (4) the limited availability of ethics-specific training for teachers (and the broader public) despite the growing impact of AI in education; and (5) the problematic transposition of ethical models from legal or technological domains that lack pedagogical grounding, making them poorly suited for the educational context. In response, the review emphasised the need for more context-sensitive ethical guidelines, teacher training, and critical engagement with AI technologies that foreground educational agency.

This study has been published in the “International Journal of Artificial Intelligence in Education” and is attached as Appendix A.

Building on this foundational mapping, Chapter 3, “Futures Studies”, further advanced the investigation by exploring imaginaries of AI’s pedagogical implications. Using the Delphi Method, this phase involved consultations with educational experts to imagine, anticipate, and critically evaluate AI’s potential impact on education.

Key components of this study included:

- An examination of how AI’s ethical challenges reflect broader concerns about the ethics of education itself.
- The construction of eight hypothetical future scenarios, derived from expert insights, to explore ethical concerns related to AI’s integration.
- Outcome: an actionable toolkit for educators as an outcome, developed through a participatory process that itself reflects an ethical approach to education.

The study underscored the importance of inclusive, diverse, participatory, and informed processes, bringing together a range of roles, professional profiles, visions, and narratives, to ensure that the questions themselves opened space for new imaginaries and broader, more open horizons.

This research has been published in the “Education and Information Technologies” journal, with the full text included in Appendix B.

Addressing Research Question 2: How can a conceptual framework on AI, ethics, and educational agency be co-designed to inform pedagogical practices that preserve and strengthen human agency in educational environments where AI is present?

This stage involved a four-session focus group study, which served two main purposes: to inform the strategic, processual, relational, and content axes of course design, and to support the drafting of a conceptual framework addressing agency and AI within the context of education ethics. The study, presented in Chapter 5, critically examined how AI technologies influence agency dynamics (including subjective, intersubjective, and collective agency) based on discussions with teacher trainers. It highlighted concerns that AI may diminish the role of teachers or reshape traditional power structures in education. It also revealed that while teachers demonstrate a nuanced understanding of students’ individual agency, they tend to focus less on AI’s broader intersubjective and collective impacts. This narrowing of perspective may indicate an epistemic shift toward dystopian narratives that risk alienating educators from

critical school bonds and collaborative agentic frameworks. Therefore, the study offered insights for educators, researchers, and policymakers, addressing medium and long-term implications of AI on human agency, in the context of education.

Main Insights and Outcomes Emerging from the Study included:

- AI systems in education must be critically assessed for both their benefits and their risks to students' reasoning, authorship, moral development, and sense of belonging.
- Teachers show deep awareness of students' subjective agency but often overlook intersubjective and collective dimensions, possibly due to dominant techno-solutionist narratives.
- There is a perceived risk of professional devaluation as AI replaces core pedagogical tasks, challenging teacher authority, professional status, and identity.
- AI is seen as a proxy agent but it requires clear boundaries to preserve human-led ethics and pedagogical intentionality.
- Teachers stress the irreplaceable role of human scaffolding in moral, emotional, and socio-cognitive development.
- Schools are reaffirmed as democratic spaces, requiring inclusive, participatory AI decision-making.
- AI may foster a “foreclosed agency”, limiting genuine authorship and decision-making.
- The study advocates for training grounded in dialogic and andragogic approaches that address agency as a layered construct.
- Outcome: three-layered framework of agency in educational settings, comprising subjective, intersubjective, and collective dimensions, whose layers dynamically intersect, illustrating how growth in one dimension can strengthen the others.

This research has been published in a Topical Collection of “Digital Society” journal, with the full text included in Appendix C.

Addressing Research Question 3: How can a training course for K-12 teachers on the ethics of education with AI be designed and implemented using an Educational Design Research approach?

Chapter 4 explored deeper into the role of teachers by conducting focus groups with educators involved in K-12 teacher education. This phase built on the findings from previous studies to co-design a professional learning programme on AI and education ethics.

Key findings and outcomes from this research included:

- Use AI not only to support course design, delivery, and assessment but also to redefine and concretise AI tools within dialogical practices.
- Consider it an ethical decision in itself to discuss AIED frameworks as either supporting tools or constraints to meaningful practices and pertinent innovation in technogeographies.
- Provide andragogic opportunities for addressing teachers' concerns, needs, and uncertainty during transition processes throughout course's design and implementation phases.
- Foster individual and collective agency on AIED ethical issues through informal and self-directed learning channels.
- Incorporate meaningful experiential learning to stimulate dialogic ethics in AIED-related discussions.
- Invest more in professional development to enable teachers to critically assess and reshape the values associated with education ethics and AI in education.
- Outcome: guidelines outlining the key organisational dimensions that shape teacher professional development in the field of AI, education ethics, and agency.

This study has been published in "Education and Information Technologies" journal and is attached as Appendix D.

The culmination of this thesis was the professional development course, "From Automation to Autonomy: Educational Agency with AI", the subject of Chapter 6, which directly addressed the design and formative implementation of the training course for K-12 teachers. This course represents a tangible, formative educational response to the ethically complex landscape of AI in education, meticulously synthesising the research theoretical analysis, iterative design processes, conceptual and empirical grounding, and context-responsiveness.

The design principles were directly informed by the preceding research stages. Agency is the central modular principle, operationalising the three-layered framework of subjective, intersubjective, and collective agency developed in Chapter 5 to guide teachers in reclaiming their roles as shapers. A principled choice was made to forgo conventional Learning Management Systems with their emphasis on individualised, algorithm-driven paths, in favour of a Canva-based platform, selected to align more coherently with the agentic capabilities the course aimed to foster. This strategic decision directly built upon the insights from the focus

groups, which emphasised the importance of dialogical practices and nurturing collective agency over isolated learning. The chapter systematically explores the importance of deliberate course architecture that fosters epistemic and collective agency through andragogic practices.

Learning within the course is intentionally integrated with AI technologies, to enable role-play and tinkering, a way of questioning learning design, assessment practices, and teacher development, while also grappling with deeper ambivalences related to automation, agency, and the erosion of pedagogical judgment. Tools like Synthesia, Character.AI, Mentimeter, Wayground, and ChatGPT/Gemini are integrated not as ends in themselves, but as vehicles for exploring collective and distributed agency in practice. The course is structured across distinct modules, each designed to progressively deepen understanding and engagement with AI's ethical and pedagogical implications, moving from broad imaginaries to specific classroom applications. Accompanying these modules, a three-layered assessment strategy, comprising personal reflection, peer feedback, and facilitator guidance, was implemented. This multi-faceted approach to assessment, valuing qualitative insights and continuous dialogue over quantitative metrics, further reinforced the course's relational pedagogy and its commitment to professional development as a process of ethical becoming, directly echoing the andragogic principles and concerns regarding teacher needs and uncertainty identified in the focus groups research stage.

The implementation and evaluation phase adhered to EDR principles, undergoing iterative refinement through structured walkthroughs and a micro-evaluation, aligning with the "Prototyping and Assessment of Preliminary Products & Theories" phase. This formative assessment, utilising qualitative observations, interviews, and mixed-method questionnaires, provided critical insights. The interpretation, guided by the ik-model across its technological, content, relational, and processual domains, confirmed the course's conceptual coherence, strong formative potential, and large alignment with teachers' expressed needs.

Main Insights and Outcomes Emerging from the Study include:

- Course design must explicitly incorporate ethical reflection and critical dialogue about AI, moving beyond mere technical adaptation or the allure of innovation narratives.
- AI can support instructional tasks, but uncritical use may compromise pedagogical intentionality and professional judgment.
- Assessment practices risk becoming mechanistic if AI-driven feedback is not balanced with relational and context-sensitive evaluation.

- Teachers-in-training often approach AI with ambivalence, recognising its potential while expressing concern over loss of creativity, autonomy, ethical reasoning, and emotional depth.
- Structured opportunities for dialogic, experiential, and reflective learning can foster more critical and empowered engagements with AI.
- Andragogic principles related to task relevance, salience and epistemic agency, are key for meaningful teacher development in AI-mediated contexts.
- Educators benefit from collaborative design processes that position them as active agents and emancipated practitioners-researchers in shaping how AI is integrated into educational practices, from classroom pedagogy to institutional democratic culture.
- There is a need to support educators in developing not only a repertoire of capabilities, but also a critical vocabulary to meaningfully engage with AI's socio-technical entanglements and its long-term educational implications.
- Responsible course design entails a balance between experimentation and ethical guardrails, reinforcing agency, justice, and pedagogical depth.
- Outcome: an 8-session professional development course (designed for a minimum of 12 hours, plus role-based implementation), supported by a custom-built Canva platform and an accompanying eBook containing detailed session plans and supplementary materials.

This doctoral research led to four main concrete outcomes. First, it mapped the evolving ethical landscape of AI in education, drawing on a decade of academic literature and continuously updated to reflect the emergence of Generative AI across the different phases of the research, as documented in the introductions to each study and corresponding chapter. Second, the study produced an ethical dilemma toolkit: a set of eight practical dilemmas developed through a participatory process. The toolkit is intended to foster meaningful dialogue in both school and policy settings, offering accessible entry points into complex ethical questions surrounding AI in education. Third, the research introduced a three-layered framework of agency, designed specifically for educational contexts to better understand and support collective forms of agency in education. Lastly, the thesis culminated in the design and formative implementation of an eight-session professional development course, supported by a Canva platform and an accompanying eBook with detailed session plans and materials.

A further and more foundational outcome of this doctoral dissertation lies in exposing the ways AI in education shapes not only practice, but also the very language through which we think, teach, and imagine. Across its phases, this research traced how socio-technical systems embed and reproduce specific imaginaries, about learning, about students, about teachers, about schools and education, often privileging prediction, standardisation, and surveillance over meaning, care, and relationality. In doing so, they subtly reshape the lexicon of education, narrowing the symbolic and epistemic resources available to name and nurture human experience.

This is not a merely semantic concern. When the lexicon of education is reduced to data points, behavioural markers, or optimisation metrics, it undermines the conditions for the open-endedness of becoming. Preserving the educational lexicon is a form of rewilding education: safeguarding its unpredictability, contradiction, silence, and imagination, which are the very conditions that encourage new beginnings and the often-invoked promise of meaningful innovation.

This work has shown that the challenges posed by AIED are not confined to the realm of legality. Ethical concerns must be reframed to account for deeper tensions, such as those between history and testimony, information and memory, networking and bonding. These are not merely semantic or abstract; they reflect deeper symbolic displacements increasingly obscured by AIED-driven jargon. In many cases, automated systems risk flattening these distinctions, replacing lived, relational concepts with operationalised, computational proxies. Rewilding education, in this sense, entails reclaiming these vocabularies and making visible how such tensions manifest in the socio-technical imaginaries of AI in education.

Testimony gives history its moral and affective depth; what ethics can there be if history is made without it? What is information without memory? Even if the concept of tertiary memory reminds us that data can be stored in externalised, technical supports interpretation remains dependent on memory as an experiential layer, one that redistributes meaning through situated, and embodied forms. It is within this layer that personal narratives take shape, where symbolic and affective resonances are disclosed, and where specific pedagogical judgments or democratic school decisions are made. Far from being a neutral repository, memory mediates the transformation of data into meaning through its entanglement with subjectivity, emotion, context, and time. Networking has become a blunt substitute for encounter, bonds, and care. While AI systems can facilitate connections through optimised networks, they cannot foster

trust or care. They can identify patterns, but they cannot testify. They can remember in storage, but not in narrative. How do we expect students to grow within systems that measure without asking, respond without listening, and give without care?

The professional development course developed in this study engages directly with these dynamics, offering educators the conceptual and experiential tools to explore these distinctions into the compositional field of thinking and practicing education with AI.

The concept of collective agency, explored throughout this study, depends precisely on these symbolic, relational dimensions. It cannot be automated. It must be enacted, in schools, in conversations, in decisions that are slow, negotiated, situated, and often uncomfortable. The challenge ahead is not simply to regulate AI or develop technical competencies, but to protect the fragile, meaningful spaces in which education remains open to subjectification, dissent, and the formation of ethical selves.

At stake is also a deeper question about desire and futurity. What kinds of futures can we meaningfully imagine without a present that nourishes desire? In a socio-technical culture increasingly shaped by visibility and performativity, where secrecy, anonymity, and intimacy are often eroded or pathologized, education must remain a space where opacity is allowed, where not being seen is also a right, and where ethical formation includes the freedom to resist being tracked, scored, or anticipated.

Agency is a compositional phenomenon: one that is fragile, layered, and deeply entangled with the psychic and societal ecologies of our time. When diminished at one level, it risks erosion at others, cutting away at the very foundations of schools as places for the co-construction of subjective and collective futures. In this sense, agency, as suggested by this research, can be understood as protention, the capacity to orient oneself toward a future that is not yet fully knowable but remains open to creation. For this to occur, technologies themselves must be rethought: they must be unsettled, redesigned, and repurposed in ways that prioritise care, relationality, and distributed agency, grounded in human, negotiated, and dialogic ethics.

Ultimately, this dissertation invites educators, researchers, and policymakers to consider that the most important task ahead is not simply about integrating AI, but about rearticulating what education is for, who it serves, and how it can continue to make space for others, for difference, for the not-yet, through a lexicon that resists closure and keeps human agency meaningfully alive. This research advances the view that ethics is not a fixed set of universal principles, but a historically constructed and socially embedded process. This aligns with

contemporary critical views in education that reject the depoliticising and deprofessionalising effects of technocratic approaches. In this light, ethical practice is inseparable from power, from pedagogy, and from the conditions under which professional judgement is exercised, inviting teachers to resist uncritical automation and algorithmic normativisation.

In the face of an expanding culture of immediacy and algorithmic determination, where the collapse of desire is met with a frantic urgency to appear active, the rewilding of education begins with renewed attention to its lexicon, to the ethical and affective vocabularies that sustain individuation, relationality, conviviality, and the imagination of meaningful futures in which the planet, also entangled in AI's extractive circuits, still matters and endures.

7.2. Limitations of the Current Research

While this thesis managed to achieve its goals of developing a framework to reframe the discourse on ethics through a prototype for agency-focused AI education, and of designing a course based on this collaboratively developed framework, certain limitations inherent in its design and scope must be acknowledged.

Firstly, the evaluation of the course relied on a small-scale micro-evaluation and expert walkthroughs. The course was implemented in a condensed, single-day format with a limited group of ten K-12 teachers. While this setting offered valuable qualitative insights, aligning with the methodological soundness of educational design research and responding to the study's conceptual novelty and the field's inherent uncertainty, it provided only a partial view of the course's long-term impact. The condensed timeframe, though informative for assessing affective engagement and participant dynamics, did not allow for a full exploration of pedagogical transfer.

Secondly, the generalizability of findings is constrained. Participants in the micro-evaluation were self-selected and already interested in the intersection of AI and education. This group may not fully reflect the broader teaching population in terms of openness to the theme, levels of digital literacy, or pedagogical context. Although the course was designed for transversality and adaptability, the findings remain contextually bound, a characteristic challenge in EDR.

Thirdly, the EDR methodology entails close researcher involvement in both design and implementation. While this proximity supported iterative refinement and responsiveness to participant needs, it also introduces a degree of interpretative bias. Although efforts were made

to mitigate this through triangulation and external feedback, the researcher's positionality remains an inherent feature of the methodology.

Fourthly, the measurement of complex constructs such as agency, or ethical reasoning poses methodological challenges, especially within a short, formative trial. The mixed-method approach helped capture emerging perceptions and conceptual shifts, but long-term impacts on teacher practice or sustained ethical positioning could not be fully assessed in this phase.

Furthermore, while the thesis addressed AI in education in a broad sense, it did not explicitly differentiate between types of AI systems. Future research could meaningfully examine whether co-creative AI tools, which might invite reflection, iteration, and dialogue, foster agency more effectively than prescriptive systems that risk narrowing learner autonomy. This distinction is crucial to further unpack how specific AI designs either support or constrain educational agency.

Lastly, the study did not focus on policy or regulatory frameworks. While the research emphasised the pedagogical and ethical dimensions of AI integration, it also highlighted the absence of regulatory structures explicitly designed to support agency-centred educational practices. Future work is needed to design governance models that embed participatory, democratic, and ethical values at the systemic level.

7.3. Future Directions for Research and Development

Building on the conceptual and practical groundwork established by this doctoral dissertation, several significant paths for future research and development are proposed.

First, longitudinal studies are to assess the sustained impact of the developed course on teacher practice, classroom dynamics, and broader school culture. Such studies could track how educators integrate AI ethics and agency-focused practices into curriculum design, assessment strategies, and dialogic pedagogical approaches over time. Such work would contribute to a deeper understanding of how agency, conceived in this study as subjective, intersubjective, and collective, evolves in response to dynamic socio-technical conditions. These longitudinal studies could also draw from existing research in educational psychology and cognitive science, particularly in areas such as sense of control, intentional binding, and agency development, while expanding into fields such as political psychology, educational sciences, critical pedagogy, and democratic theory. This interdisciplinary approach would offer a more integrated and contextually grounded understanding of human–AI interaction in education,

linking individual experiences of agency with broader institutional and societal negotiations of agency.

Second, broader implementation across varied educational settings is crucial to explore the adaptability and scalability of the course and framework. Future studies could examine the course's relevance across different grade levels, subject areas, and sociocultural contexts. Deploying the training in international or underrepresented contexts may offer valuable insights into how agency is interpreted and embodied within diverse educational imaginaries and ethical traditions.

Third, there is a pressing need to develop strategies for ethical dissemination and pedagogical scalability. While the course is intentionally designed to resist automated delivery, wider adoption can be supported through the development of train-the-trainer programmes, robust facilitator guides, and open-access educational resources. These tools should preserve the course's dialogic, participatory ethos while expanding its reach. Moreover, research into the formation of sustained professional learning networks and communities of practice could deepen its impact and support continuous professional development in shared agency across the diverse "trading zones" where people, education, technology, and policy intersect. This implies understanding agency as situated within specific geographies, shaped by the tools, spaces, and systems that surround learners, and which they actively co-construct and participate in shaping through processes of signification.

Fourth, future research could directly investigate the student perspective. While this dissertation has centred on teacher agency, students' interactions with AI systems deserve further study. Exploring how students perceive, represent, question, and respond to AI socio-technical systems could meaningfully contribute to expanding the current model. This direction would also support the refinement of the agency framework by grounding it in observed developmental trajectories within student learning contexts.

Fifth, ongoing refinement of the course and conceptual model is necessary to remain responsive to technological and pedagogical change. As AI capabilities evolve, particularly in areas such as multimodal interaction, predictive personalisation, and generative content creation, new ethical challenges and contested understandings of agency will emerge. Future iterations of the course could include updated examples of AIED technologies, mapped explicitly onto the three-layered framework of agency. A more representational and visual

articulation of these layers, integrated with relevant technologies and associated developmental processes, could support both practitioner comprehension and further theoretical development.

In sum, future research is expected to advance the central agenda of this dissertation: that collective agency, ethics, and pedagogy must remain at the core of educational engagement with AI.

CHAPTER EIGHT

Dissemination of Ph.D. Research

This thesis was developed as part of a doctoral research project funded by the Portuguese Foundation for Science and Technology (FCT), under the reference 2020.06258.BD and funding DOI <https://doi.org/10.54499/2020.06258.BD>.

Throughout the course of this doctoral research, significant efforts have been made to disseminate the findings and contribute to the academic community. The following sections outline the diverse range of scholarly outputs resulting from this work, including publications in peer-reviewed journals, proceedings from international conferences, conference presentations, involvement as a reviewer for various academic journals, advisory board memberships, participation in research networks, and research stays. These contributions highlight the ongoing engagement with the academic community and the impact of this research across multiple platforms and networks.

8.1 Publications in Peer-Reviewed Journals

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2025). “Where is agency moving to?”: Exploring the interplay between AI technologies in education and human agency. *Digital Society*, 4(49). <https://doi.org/10.1007/s44206-025-00203-9>

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2025). Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12946-y>

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2024). Design of a future scenarios toolkit for an ethical implementation of Artificial Intelligence in education. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12229-y>

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2024). Uncovering blind spots in education ethics: Insights from a systematic literature review on Artificial Intelligence in education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-023-00384-9>

8.2 Publications in Conference Proceedings

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2021). Blending machines, learning, sense of agency, and ethics: Designing an in-depth framework with experts using the Delphi Method approach. In M. Alier & D. Fonseca (Eds.), *Proceedings of the Ninth*

International Conference on Technological Ecosystems for Enhancing Multiculturality (pp. 665–670). Association for Computing Machinery.

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2020). Sense of agency in times of automation: A teachers' professional development proposal on the ethical challenges of AI applied to education. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial intelligence in education: AIED 2020* (Vol. 12164, pp. 405–408). *Lecture Notes in Computer Science*. Springer. https://doi.org/10.1007/978-3-030-52240-7_73

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2019). Blending machines, learning, and ethics. In M. González, F. Sedano, C. Llamas, & F. García-Peñalvo (Eds.), *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 993–998). Association for Computing Machinery. <https://doi.org/10.1145/3362789.3362909>

8.3. Conference Presentations

Mouta, A., Pinto Llorente, A. M., & Torrecilla Sánchez, E. M. (2025, February 20). *From mastery to networks: Socio-technical AI systems and human agency* [Conference presentation]. Education After the Algorithm: Co-designing Critical and Creative Futures, Dublin City University, Online.

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2024, August 27–30). *Professional learning and teacher's agency in AI integration* [Conference presentation]. Encuentros Doctorales UIMP 2024: Desafíos en la Investigación Interdisciplinar, Universidad Internacional Menéndez Pelayo, Santander, Spain.

Mouta, A., Pinto Llorente, A. M., & Torrecilla Sánchez, E. M. (2023, September 20–22). *Hacia un uso ético de las tecnologías de IA en la educación: Los docentes como agentes éticos de cambio* [Conference presentation]. The Values of the Commons in the Digital Society, Santiago de Compostela, Spain.

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2022, May 26). *Participatory processes for governance in AIED* [Conference presentation]. Digital Education Governance Beyond International Comparative Assessments, Online.

Mouta, A., Pinto Llorente, A. M., & Torrecilla Sánchez, E. M. (2021, October 27–29). *Blending machines, learning, sense of agency and ethics: Designing an in-depth framework with experts using the Delphi Method approach* [Conference presentation]. TEEM'21: Technological Ecosystems for Enhancing Multiculturality, Online.

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2021, November 4). *Inteligencia artificial en educación y sentido de agencia: Riesgos y oportunidades* [Conference presentation]. IRED 2021: Conferencia Internacional de Investigación en Educación, Online.

Mouta, A., Torrecilla Sánchez, E., & Pinto-Llorente, A. M. (2020, July 12). *Sense of agency in times of automation: A teachers' professional development proposal on the ethical challenges of AI applied to education* [Conference presentation]. 21st International Conference on Artificial Intelligence in Education, Online.

Mouta, A., Torrecilla Sánchez, E., & Pinto-Llorente, A. (2019, October 17). *Blending machines, learning, and ethics* [Conference presentation]. Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality, Association for Computing Machinery, Salamanca, Spain.

8.4. Journal Reviewer

- Acta Psychologica (Arts and Humanities Q1)
- Critical Education (Institute for Critical Education Studies, The University of British Columbia)
- Discover Education (Springer)
- Education and Information Technologies (Education Q1)
- Frontiers in Education (Education Q2)
- Fuori Logo (Università degli Studi di Napoli Federico II)
- Heliyon, Elsevier (Multidisciplinary Q1)
- Innoeduca: International Journal of Technology and Educational Innovation (Universidad de Málaga)
- International Journal of Artificial Intelligence in Education (Education; E-learning Q1)
- Social Sciences & Humanities Open (Social Sciences Q2)

8.5. Advisory Board Member

- 26th International Conference on Artificial Intelligence in Education
- 25th International Conference on Artificial Intelligence in Education
- 24th International Conference on Artificial Intelligence in Education
- EAIT 2024: 5th International Conference on Education and AI Technologies
- ICAITE 2024: International Conference on AI and Teacher Education

- EARLI23 20th Biennial EARLI Conference – Education as a Hope in Uncertain Times

8.6. Research Network

- 2025: Local coordinator for the project “Problematizing Education and Digital Technology”, CSET 2025, Critical Studies of Education and Technology, led by Neil Selwyn
- 2025: Member of the Online Consultation on the Draft Committee of Ministers “Recommendation on AI Literacy for Human Rights, Democracy and Social Agency”
- 2025: Team Member of the ASEFCClassNet18 Faculty Collaboration on “Effective, Inclusive, and Ethical Artificial Intelligence in Education (AIED) Design” - 7-month long hybrid knowledge and capacity-building project for teachers-trainers, academics, researchers and teacher trainers across Asia and Europe
- 2024: Member of the "2nd Working Conference: Regulating the Use of AI Systems in Education”, at the Council of Europe
- 2024: Became a member of the Research Group “Critical Studies on AI and Education”, led by Wayne Holmes, International Research Centre on AI, UNESCO and College London

8.7. Publications Resulting from Research Network

Holmes, W., Mouta, A., Hillman, V., ... Yeo, B. (2025, June 26). *Critical Studies of Artificial Intelligence and Education. Putting a Stake in the Ground*. SSRN.

<https://doi.org/10.2139/ssrn.5391793>

Selwyn, Neil et al. (2025). *Critical studies of education and technology ... reasons to be hopeful? Monash University*. Report. <https://doi.org/10.26180/29265038.v1>

Mouta, A., & Bardone, E. (in press). Neither the helm nor the wheel: Charting the landscape of agency in highly automated education. In W. Holmes (Ed.), *Handbook of critical studies of AI in education*. Edward Elgar Publishing Ltd. [Manuscript in press].

8.8. Research Stays

- Research stay at the Centre for Educational Technology, University of Tartu, under the supervision of Emanuele Bardone. Period: 10 January 2025 – 10 February 2025.

- Research stay at the Observatório Social para a Inteligência Artificial e Dados Digitais, under the supervision of Paulo Nuno Vicente. Period: 17 February 2025 – 17 April 2025.

An extended thesis abstract in Spanish is included in Appendix F.

CHAPTER NINE

References

- Abakah, E., Widin, J., & Ameyaw, E. K. (2022). Continuing professional development (CPD) practices among basic school teachers in the Central Region of Ghana. *SAGE Open*, 12(2). <https://doi.org/10.1177/21582440221094597>
- Abboud, R., Arya, A., & Pandi, M. (2020). Redefining the digital divide in the age of AI: The harvest of the 25th anniversary. In L. Gómez Chova, A. López Martínez, & I. Candel Torres (Eds.), *INTED2020 Proceedings* (pp. 4483–4492). IATED Academy. <https://doi.org/10.21125/inted.2020.1241>
- Adams, C., Pente, P., Lemermeyer, G., Turville, J., & Rockwell, G. (2022). Artificial intelligence and teachers' new ethical obligations. *The International Review of Information Ethics*, 31(1). <https://doi.org/10.29173/irie483>
- AI4T. (2024a). *Artificial intelligence for and by teachers*. <https://www.ai4t.eu/>
- AI4T. (2024b). *Work Package 1: Analysis of traces – Final version*. https://www.ai4t.eu/wp-content/uploads/2024/01/AI4T_WP1_Analysis-of-traces_Final-version.pdf
- AI4T. (2024c). *Work Package 3: Comparative European evaluation report*. https://www.ai4t.eu/wp-content/uploads/2024/07/AI4T_WP3_D3.3_Comparative-European-Evaluation-Report.pdf
- Akin, S., Calik, B., & Demir, C. (2017). Students as change agents in the community: Developing active citizenship at schools. *Educational Sciences: Theory and Practice*, 17(3), 809–834. <https://doi.org/10.12738/estp.2017.3.0176>
- Ali, S., Williams, R., Payne, B., Park, H., & Breazeal, C. (2019). Constructionism, ethics, and creativity: Developing primary and middle school artificial intelligence education. *Proceedings of the International Workshop on Education in Artificial Intelligence K-12*. <https://www.media.mit.edu/publications/constructionism-ethics-and-creativity/>
- Alier Forment, M., García-Peñalvo, F.J., Casañ Guerrero, M. J., Pereira, J. A., & Llorens Largo, F. (2024, October 8). *Safe AI in Education Manifesto* (Version 0.4.0). *Safe AI in Education Manifesto*. <https://manifesto.safeaieducation.org>
- Al-Omran, G., Al-Abdulhadi, S., & Jan, M. R. (2019). Ethics in artificial intelligence. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 940–949. <https://www.ieomsociety.org/gcc2019/papers/337.pdf>
- Alrahwawi, H. A., Jamiat, N., Umar, I. N., & Abu-Naser, S. S. (2023). Improvement of students' achievement by using intelligent tutoring systems: A bibliometric analysis

- and reviews. *Journal of Theoretical and Applied Information Technology*, 101(10), 3793–3800. <https://www.jatit.org/volumes/Vol101No10/11Vol101No10.pdf>
- Anders, G. (2001). *Wir Eichmannsöhne: Offener Brief an Klaus Eichmann* (2^a ed.). C.H. Beck.
- Andreotti, V., Stein, S., Ahenakew, C., & Hunt, D. (2015). Mapping interpretations of decolonization in the context of higher education. *Decolonization: Indigeneity, Education & Society*, 4(1), 21–40. <https://representing-education.gertrudecotter.info/wp-content/uploads/2016/08/andreotti-stein-ahenakew-hunt-decolonization.pdf>
- Antoniou, J. (2021). Dealing with emerging AI technologies: Teaching and learning ethics for AI. In J. Antoniou (Ed.), *EAI/Springer Innovations in Communication and Computing* (pp. 79–93). Springer. https://doi.org/10.1007/978-3-030-52559-0_6
- Archer, M. S. (2014). *Structure, agency and the internal conversation*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139087315>
- Archer, M. S. (2015). How agency is transformed in the course of social transformation: Don't forget the double morphogenesis. In M. S. Archer (Ed.), *Generative mechanisms transforming late modernity* (pp. 135–157). Springer. https://doi.org/10.1007/978-3-319-10788-6_6
- Arendt, H. (1958). *The human condition*. University of Chicago Press.
- Arteaga-Martínez, B., Macías-Sánchez, J., Pla-Castells, M., & Ramírez-García, M. (2021). Design and validation of an observation instrument for mathematics classes in secondary education: Expert panel and Delphi process. *RELIEVE*, 27(2), Article 3. <https://doi.org/10.30827/relieve.v27i2.21812>
- Aurobindo, S. (1933). *The riddle of this world*. Sri Aurobindo Ashram Publication.
- Bahadorestani, A., Naderpajouh, N., & Sadiq, R. (2020). Planning for sustainable stakeholder engagement based on the assessment of conflicting interests in projects. *Journal of Cleaner Production*, 242, 118402. <https://doi.org/10.1016/j.jclepro.2019.118402>
- Ball, S. J. (2013). *Foucault, power, and education*. Routledge.
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193–209. https://doi.org/10.1207/s15327957pspr0303_3
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>

- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1, 164–180. <https://doi.org/10.1111/j.1745-6916.2006.00011.x>
- Bandura, A. (2023). *Social cognitive theory: An agentic perspective on human nature*. Wiley.
- Barab, S., & Squire, K. (2004). Design-Based Research: Putting a Stake in the Ground. *Journal of the Learning Sciences*, 13(1), 1–14. https://doi.org/10.1207/s15327809jls1301_1
- Bardone, E., Mõttus, P., & Eradze, M. (2024). Tinkering as a complement to design in the context of technology integration in teaching and learning. *Postdigital Science and Education*, 6(1), 114–134. <https://doi.org/10.1007/s42438-023-00416-6>
- Bates, R. A. (2011, June). AI & SciFi: Teaching writing, history, technology, literature, and ethics [Paper presentation]. *ASEE Annual Conference & Exposition*, Vancouver, BC. <https://doi.org/10.18260/1-2-17433>
- Bell, W. (2004). *Foundations of Futures Studies II: Values, Objectivity and the Good Society*. Taylor & Francis.
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, 3(21). <https://doi.org/10.1126/scirobotics.aat5954>
- Bengston, D. N. (2016). The futures wheel: A method for exploring the implications of social–ecological change. *Society & Natural Resources*, 29(3), 374–379. <https://doi.org/10.1080/08941920.2015.1054980>
- Bennett, J. (2010). *Vibrant Matter. A Political Ecology of Things*. Duke University Press.
- Berberian, B., Sarrazin, J. C., Le Blaye, P., & Haggard, P. (2012). Automation technology and sense of control: A window on human agency. *PLoS ONE*, 7, e34075. <https://doi.org/10.1371/journal.pone.0034075>
- Biesta, G., & Tedder, M. (2006). *How is agency possible? Towards an ecological understanding of agency-as-achievement (Working Paper Five)*. *Learning Lives: Learning, Identity, and Agency in the Life Course*. Exeter: Teaching and Learning Research Program.
- Biesta, G. (2013). *Beautiful Risk of Education*. Routledge. <https://doi.org/10.4324/9781315635866>
- Biesta, G. (2020). Risking ourselves in education: Qualification, socialization, and subjectification revisited. *Educational Theory*, 70(1), 89–104. <https://doi.org/10.1111/edth.12411>

- Bjerstedt, A. (Ed.). (1982). Future Consciousness and the School. *Didakometry & Sociometry*, 14, 3-4. Department of Educational and Psychological Research, School of Education.
- Blease, C., Kharko, A., Annoni, M., Gaab, J., & Locher, C. (2021). Machine learning in clinical psychology and psychotherapy education: A mixed methods pilot survey of postgraduate students at a Swiss university. *Frontiers in Public Health*, 9, 623088. <https://doi.org/10.3389/fpubh.2021.623088>
- Bogina, V., Hartman, A., Kuflik, T., & Shulner-Tal, A. (2021). Educating software and AI stakeholders about algorithmic fairness, accountability, transparency and ethics. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00248-0>
- Bond, M., Marín, V. I., & Bedenlier, S. (2021). International collaboration in the field of educational research: A Delphi study. *Journal of New Approaches in Educational Research*, 10(2), 190–213. <https://doi.org/10.7821/naer.2021.7.614>
- Borges, C. D., & Santos, M. A. (2005). Aplicações da técnica do grupo focal: Fundamentos metodológicos, potencialidades e limites. *Revista da SPAGESP*, 6(1), 74–80.
- Bourdieu, P. (1990). *The logic of practice*. Stanford University Press.
- Bourdieu, P. (1999). *Language and symbolic power*. Harvard University Press.
- Bozkurt, A., Karadeniz, A., Bañeres, D., Guerrero-Roldán, A., & Rodríguez, M. E. (2021). Artificial intelligence and reflections from educational landscape: A review of AI studies in half a century. *Sustainability*, 13, 800. <https://doi.org/10.3390/su13020800>
- Brey, P. (2017). Ethics of emerging technologies. In S. O. Hansson (Ed.), *Methods for the ethics of technology* (pp. 175–192). Rowman & Littlefield International.
- Bridges, W. (1991). *Managing transitions: Making the most of change*. Da Capo Press.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences*, 2(2), 141–178. https://doi.org/10.1207/s15327809jls0202_2
- Bucea-Manea-Țoniș, R., Kuleto, V., Gudej, S. C. D., Lianu, C., Lianu, C., Ilić, M. P., & Păun, D. (2022). Artificial intelligence potential in higher education institutions: Enhanced learning environment in Romania and Serbia. *Sustainability*, 14(10), 5842. <https://doi.org/10.3390/su14105842>

- Burr, C., Taddeo, M., & Floridi, L. (2020). The ethics of digital well-being: A thematic review. *Science and Engineering Ethics*, 26, 2313–2343. <https://doi.org/10.1007/s11948-020-00175-8>
- Cabero Almenara, J., & Barroso Osuna, J. (2013). La utilización del juicio de experto para la evaluación de TIC: El coeficiente de competencia experta. *Bordón. Revista de Pedagogía*, 65(2), 25–38. <https://recyt.fecyt.es/index.php/BORDON/article/view/brp.2013.65202>
- Cabero Almenara, J., & Infante Moro, A. (2014). Empleo del método Delphi y su empleo en la investigación en comunicación y educación. *EDUTEC. Revista Electrónica de Investigación Educativa*, 48, 1–16. <https://doi.org/10.21556/edutec.2014.48.187>
- Call for fellow nominations. (2023, March 31). *AIED 2023*. https://www.aied2023.org/c_mentoring_fellowship.html
- Cave, S. (2020). The problem with intelligence: Its value-laden history and the future of AI. In A. Markham, J. Powles, T. Walsh, & A. L. Washington (Eds.), *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)* (pp. 29–35). Association for Computing Machinery. <https://doi.org/10.1145/3375627.3375813>
- Charisi, V., Malinverni, L., Rubegni, E., & Schaper, M. M. (2020). Empowering children's critical reflections on AI, robotics and other intelligent technologies. In I. Šmorgun & G. Berget (Eds.), *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (pp. 1–4). Association for Computing Machinery. <https://doi.org/10.1145/3419249.3420090>
- Chichekian, T., & Benteux, B. (2022). The potential of learning with (and not from) artificial intelligence in education. *Frontiers in Artificial Intelligence*, 5, 903051. <https://doi.org/10.3389/frai.2022.903051>
- Chiu, T. K. F., Meng, H., Chai, C., King, I., Wong, S., & Yam, Y. (2021). Creation and evaluation of a pre-tertiary artificial intelligence (AI) curriculum. *IEEE Transactions on Education*, 65(1), 30–39. <https://doi.org/10.1109/TE.2021.3085878>
- Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring teachers' perceptions of artificial intelligence as a tool to support their practice in Estonian K-12 education. *International Journal of Artificial Intelligence in Education*, 32, 725–755. <https://doi.org/10.1007/s40593-021-00243-5>

- Christakopoulou, S., Dawson, J., & Gari, A. (2001). The community well-being questionnaire: Theoretical context and initial assessment of its reliability and validity. *Social Indicators Research*, *56*, 319–349. <https://doi.org/10.1023/A:1012478207457>
- Cochran-Smith, M., Craig, C. J., Orland-Barak, L., Cole, C., & Hill-Jackson, V. (2022). Agents, agency, and teacher education. *Journal of Teacher Education*, *73*(5), 445–448. <https://doi.org/10.1177/00224871221123724>
- Coeckelbergh, M. (2020). *AI ethics*. MIT Press.
- Coenen, M., Coenen, T., Stamm, A., Stucki, G., & Cieza, A. (2012). Individual interviews and focus groups with patients with rheumatoid arthritis: A comparison of two qualitative methods. *Quality of Life Research*, *21*, 359–370. <https://doi.org/10.1007/s11136-011-9943-2>
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good proctor or “Big Brother”? Ethics of online exam supervision technologies. *Philosophy & Technology*, *34*, 1581–1606. <https://doi.org/10.1007/s13347-021-00476-1>
- Copur-Gencturk, Y., Li, J., Cohen, A. S., & Orrill, C. H. (2024). The impact of an interactive, personalized computer-based teacher professional development program on student performance: A randomized controlled trial. *Computers and Education*, *210*, 104963. <https://doi.org/10.1016/j.compedu.2023.104963>
- Córdova, P. R., & Vicari, R. M. (2022). Practical ethical issues for artificial intelligence in education. In A. Reis, J. Barroso, P. Martins, A. Jimoyiannis, R. Y. M. Huang, & R. Henriques (Eds.), *Technology and innovation in learning, teaching and education. TECH-EDU 2022. Communications in computer and information science* (pp. 437–445). Springer. https://doi.org/10.1007/978-3-031-22918-3_34
- Costa, A. P., Charpiot, L., Lera, F. J., Ziafati, P., Nazarihorram, A., van der Torre, L., & Steffgen, G. (2018). More attention and less repetitive and stereotyped behaviors using a robot with children with autism. In J. J. Cabibihan, F. Mastrogiovanni, A. K. Pandey, S. Rossi, & M. Staffa (Eds.), *27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 534–539). IEEE. <https://doi.org/10.1109/ROMAN.2018.8525747>
- Coulter, D., & Wiens, J. R. (2002). Educational judgment: Linking the actor and the spectator. *Educational Researcher*, *31*(4), 15–25. <https://doi.org/10.3102/0013189X031004015>

- Council of Europe. (2022). *Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law*. <https://rm.coe.int/prems-092922-gbr-2517-ai-and-education-txt-16x24-web/1680a956e3>
- Council of Europe. (2024, October 23). *Conference on regulating the use of AI systems in education*. <https://www.coe.int/en/web/education/-/conference-on-regulating-the-use-of-ai-systems-in-education-1>
- Crary, J. (2022). *Scorched earth: Beyond the digital age to a post-capitalist world*. Verso.
- Cukurova, M., Miao, X., & Brooker, R. (2023). Adoption of artificial intelligence in schools: Unveiling factors influencing teachers' engagement. In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova (Eds.), *Proceedings of the International Conference on Artificial Intelligence in Education (AIED 2023), Lecture Notes in Computer Science* (pp. 151–163). Springer Nature. <https://doi.org/10.48550/arXiv.2304.00903>
- Dai, C.-P. (2024). Designing learning support for simulation-based preservice teacher learning with artificial intelligence-powered virtual agents. *International Journal for Educational Media and Technology*, 17(2). <https://ijemt.org/index.php/journal/article/view/312>
- Dalkey, N., & Helmer, O. (1963). An experimental application of the Delphi method to the use of experts. *Management Science*, 9(3), 458–467. <https://doi.org/10.1287/mnsc.9.3.458>
- Damşa, C. (2014). Shared epistemic agency and agency of individuals, collaborative groups, and research communities. In J. L. Polman, E. A. Kyza, D. K. O'Neill, I. Tabak, W. R. Penuel, A. S. Jurow, K. O'Connor, T. Lee, & L. D'Amico (Eds.), *Learning and Becoming in Practice: The International Conference of the Learning Sciences (ICLS) 2014* (pp. 440–447). International Society of the Learning Sciences.
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.
- Davis-Kean, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19 (2), 294–304. <https://doi.org/10.1037/0893-3200.19.2.294>
- Dean, M. (2019). *Scenario planning: A literature review*. <https://doi.org/10.13140/RG.2.2.12629.24802>

- Deleuze, G., & Guattari, F. (1980). *A thousand plateaus: Capitalism and schizophrenia*. University of Minnesota Press.
- Deleuze, G., & Guattari, F. (1995). *Mil platôs: Capitalismo e esquizofrenia* (Vols. 1–5). Editora 34.
- Dever, D. A., Wiedbusch, M. D., Romero, S. M., & Azevedo, R. (2024). Investigating pedagogical agents' scaffolding of self-regulated learning in relation to learners' subgoals. *British Journal of Educational Technology*, 55, 1290–1308. <https://doi.org/10.1111/bjet.13432>
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. Macmillan.
- Dewey, J. (1933). *How we think: A restatement of the relation of reflective thinking to the educative process*. D.C. Heath and Company.
- Dieterle, E., Dede, C., & Walker, M. (2022). The cyclical ethical effects of using artificial intelligence in education. *AI & Society*, 1–11. Advance online publication. <https://doi.org/10.1007/s00146-022-01497-w>
- Dignum, V. (2020). AI is multidisciplinary. *AI Matters*, 5(4), 18–21. <https://doi.org/10.1145/3375637.3375644>
- Dignum, V. (2021). The role and challenges of education for responsible AI. *London Review of Education*, 19(1), Article 1. <https://doi.org/10.14324/LRE.19.1.01>
- Dinges, M., Wang, A., & Schuch, K. (2020). Using the Delphi method in evaluations: Incorporating a future-oriented perspective in evaluations. *Fteval Journal for Research and Technology Policy Evaluation*, 50, 44–50. <https://doi.org/10.22163/fteval.2020.469>
- Dixit, P., & Pathak, U. (2023). Students' learning outcomes and emerging practices of blended learning: A case study. In N. J. Ahuja, A. Kumar, & A. Nayyar (Eds.), *Sustainable blended learning in STEM education for students with additional needs. Contributions to Environmental Sciences & Innovative Business Technology*. Springer. https://doi.org/10.1007/978-981-99-3497-3_11
- Dobrosovestnova, A., & Hannibal, G. (2020). Teachers' disappointment: Theoretical perspective on the inclusion of ambivalent emotions in human-robot interactions in education. In T. Belpaeme, J. Young, H. Gunes, & L. Riek (Eds.), *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI)* (pp. 225

- 471–480). Association for Computing Machinery.
<https://doi.org/10.1145/3319502.3374816>
- Eliade, M. (2016). *O sagrado e o profano*. Relógio D'Água.
- Elias, N. (2007). *Time: An essay*. University College Dublin Press.
- Ellul, J. (1964). *The technological society*. Vintage Books.
- Elyagutu, D. C. (2016). *Importance of art education (Sanat eğitiminin önemi)*. *Art-Sanat*, 6, 243–247. <https://dergipark.org.tr/en/download/article-file/229873>
- Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology*, 103, 962–1023. <https://doi.org/10.1086/231294>
- Epp, C. D., & Makos, A. (2013). Using simulated learners and simulated learning environments within a special education context. *CEUR Workshop Proceedings*, 1009, 1–10. <https://ceur-ws.org/Vol-1009/0401.pdf>
- European Commission. (2019). *Ethics guidelines for trustworthy AI*. Publications Office of the European Union. <https://data.europa.eu/doi/10.2759/346720>
- European Schoolnet. (2021). *Artificial intelligence role in K12 education: Agile collection of information*. <https://resetedu.eu/wp-content/uploads/2021/11/4.-LIDIJAKRALJ-Group-Discussion.pdf>
- European Union. (2018). *Supporting school innovation across Europe*. Publications Office of the European Union. <https://doi.org/10.2766/466312>
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <https://doi.org/10.1016/j.system.2024.103254>
- Fawns, T. (2022). An entangled pedagogy: Looking beyond the pedagogy–technology dichotomy. *Postdigital Science and Education*, 4, 711–728. <https://doi.org/10.1007/s42438-022-00302-7>
- Firminger, L. (2003). *Trend analysis: Methods and problems, strategic planning services*. Swinburne University of Technology, TAFE Division.
- Floridi, L., & Cowls, J. (2019). A Unified Framework of Five Principles for AI in Society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- Foucault, M. (1975). *Discipline and punish: The birth of the prison*. Vintage Books.
- Foucault, M. (1980). *Power/knowledge: Selected interviews and other writings, 1972–1977* (C. Gordon, Ed.). Pantheon Books.

- Freire, P. (1996). *Pedagogy of the oppressed*. Penguin. (Original work published 1970)
- Fuchs, T. (2018). The cyclical time of the body and its relation to linear time. *Journal of Consciousness Studies*, 25(7–8), 47–65.
- Gallastegui, G., Miguel, L., & Forradellas, F. R. (2021). Business methodology for the application in university environments of predictive machine learning models based on an ethical taxonomy of the student’s digital twin. *Administrative Sciences*, 11(4), 118. <https://doi.org/10.3390/admsci11040118>
- Gangadean, A. (2006). A planetary crisis of consciousness: The end of ego-based cultures and our dimensional shift toward a sustainable global civilization. *World Futures*, 62, 441–454. <https://doi.org/10.1080/02604020600798627>
- Gaona, J., López, S. S., & Montoya-Delgadillo, E. (2022). Prospective mathematics teachers learning complex numbers using technology. *International Journal of Mathematical Education in Science and Technology*, 1–30. <https://doi.org/10.1080/0020739X.2022.2133021>
- García-Peñalvo, F. J., Corell, A., Abella-García, V., & Grande-De-Prado, M. (2021). Recommendations for mandatory online assessment in higher education during the COVID-19 pandemic. In D. Burgos, A. Tlili, & A. Tabacco (Eds.), *Radical solutions for education in a crisis context: COVID-19 as an opportunity for global learning* (pp. 85–89). Springer. https://doi.org/10.1007/978-981-15-7869-4_6
- García-Peñalvo, F. J., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in generative AI. *International Journal of Interactive Multimedia and Artificial Intelligence*, 8(4), 7-16. <https://doi.org/10.9781/ijimai.2023.07.006>
- García-Peñalvo, F. J., Llorens-Largo, F., & Vidal, J. (2024). The new reality of education in the face of advances in generative artificial intelligence. *RIED: revista iberoamericana de educación a distancia*, 27(1), 9-39. <https://doi.org/10.5944/ried.27.1.37716>
- García-Peñalvo, F.J., Casañ-Guerrero, M., Alier-Forment, M., & Pereira, J. (2025). La ética de la inteligencia artificial generativa en educación a debate: Perspectiva desde el desarrollo de un caso de estudio teórico-práctico. *Revista Española de Pedagogía*, 83, Article 4577. <https://doi.org/10.22550/2174-0909.4577>
- Garrett, N., Beard, N., & Fiesler, C. (2020). More than “If time allows”: The role of ethics in AI education. In A. Markham, J. Powles, T. Walsh, & A. L. Washington (Eds.),

- Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)* (pp. 272–278). Association for Computing Machinery. <https://doi.org/10.1145/3375627.3375868>
- Gary, E. (2019). Ethics to prepare educators for professional service robots in classrooms. In E. Miyoshi, E. G. Kowch, J. C. Liu, Q. Jin, Z. Li, & H. Zhang (Eds.), *International Joint Conference on Information, Media, and Engineering Proceedings* (pp. 478–484). IEEE. <https://doi.org/10.1109/IJCIME49369.2019.00102>
- Gasevic, D., Siemens, G., & Sadiq, S. (2023). Empowering learners for the age of artificial intelligence. *Computers and Education: Artificial Intelligence*, 4, 100130. <https://doi.org/10.1016/j.caeai.2023.100130>
- Gauld, C., Micoulaud-Franchi, J., & Dumas, G. (2021). Comment on Starke et al.: ‘Computing schizophrenia: Ethical challenges for machine learning in psychiatry’: From machine learning to student learning: Pedagogical challenges for psychiatry. *Psychological Medicine*, 51(14), 2509–2511. <https://doi.org/10.1017/S0033291720003906>
- Gertz, N. (2016). Designing responsibility. In M. van der Velden, M. Strano, H. Hrachovec, J. Abdelnour Nocera, & C. Ess (Eds.), *Culture, technology, communication: Common worlds, different futures? Proceedings of the Tenth International Conference on Culture, Technology, Communication* (pp. 191–201). https://www.researchgate.net/publication/312591793_Designing_Responsibility
- Ghotbi, N., & Ho, T. (2021). Moral awareness of college students regarding artificial intelligence. *Asian Bioethics Review*, 13(4). <https://doi.org/10.1007/s41649-021-00182-2>
- Gibson, J. M. E. (1998). Using the Delphi to identify the content and context of nurses’ continuing professional development needs. *Journal of Clinical Nursing*, 7, 451–459. <https://doi.org/10.1046/j.1365-2702.1998.00175.x>
- Giddens, A. (1991). *Modernity and self-identity: Self and society in the late modern age*. Stanford University Press.
- Gidley, J. M. (1998). Youth futures: Transcending violence through the artistic imagination. In S. Inayatullah & P. Wildman (Eds.), *Futures studies: Methods, emerging issues and civilizational visions*. Prosperity Press. <https://wfsf.org/wp-content/uploads/2020/04/GIDLEY-and-INAYATULLAH-2002-Youth-futures-book-file.pdf>

- Gidley, J. M., Bateman, D., & Smith, C. (2004). *Futures in education: Principles, practices and potential* (Australian Foresight Institute Monograph Series, No. 5). Swinburne Press.
- Gidley, J. M. (2017). *The future: A very short introduction*. Oxford University Press.
- Giroux, H. A. (1988). *Teachers as intellectuals: Toward a critical pedagogy of learning*. Praeger.
- Godwin-Jones, R. (2024). Distributed agency in language learning and teaching through generative AI. *Language Learning & Technology*, 28(2), 5–30. <https://doi.org/10125/73570>
- Goldsmith, J., Burton, E., Dueber, D., Goldstein, B., Sampson, S., & Toland, M. (2020). Assessing ethical thinking about AI. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(9), 13525–13528. <https://doi.org/10.1609/aaai.v34i09.7075>
- Gong, X., Tang, Y., Liu, X., Jing, S., Cui, W., Liang, J., & Wang, F. Y. (2020). K-9 artificial intelligence education in Qingdao: Issues, challenges and suggestions. *2020 IEEE International Conference on Networking, Sensing and Control*, 1–6. <https://doi.org/10.1109/ICNSC48988.2020.9238087>
- Gordon, T. (2003). *Cross-impact analysis* [CD-ROM]. The Millennium Project.
- Gorur, R., Hoon, L., & Kowal, E. (2020). Computer science ethics education in Australia – A work in progress. In H. Mitsuhashi, Y. Goda, Y. Ohashi, M. M. T. Rodrigo, J. Shen, N. Venkatarayalu, G. Wong, M. Yamada, & C. U. Lei (Eds.), *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)* (pp. 945–947). IEEE. <https://doi.org/10.1109/TALE48869.2020.9368375>
- Gozli, D. (2019). Sense of Agency. In D. Gozli (Ed.), *Experimental Psychology and Human Agency* (pp.137-164). Springer, Cham. https://doi.org/10.1007/978-3-030-20422-8_7
- Green, N. L. (2021). An AI ethics course highlighting explicit ethical agents. In M. Fourcade, B. Kuipers, S. Lazar, & D. Mulligan (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society* (pp. 519–524). Association for Computing Machinery. <https://doi.org/10.1145/3461702.3462552>
- Green, R. (2014). The Delphi technique in educational research. *SAGE Open*, 4(2), 1–8. <https://doi.org/10.1177/2158244014529773>
- Greeuw, S. C. H., Asselt, M. B. A., Grosskurth, J., Storms, C., Rijkens-Klomp, N., Rothman, D. S., & Rotmans, J. (2000). *Cloudy crystal balls: An assessment of recent European*

- and global scenario studies and models*. European Environment Agency.
https://www.eea.europa.eu/publications/Environmental_issues_series_17
- Gudiño Paredes, S., Jasso Peña, F. D., & de La Alcazar, F. J. M. (2021). Remote proctored exams: Integrity assurance in online education? *Distance Education*, 42(2), 200–218.
<https://doi.org/10.1080/01587919.2021.1910495>
- Guest, G., Namey, E., & McKenna, K. (2016). How many focus groups are enough? Building an evidence base for non-probability sample sizes. *Field Methods*, 29(1), 3–22.
<https://doi.org/10.1177/1525822X166639015>
- Guilló, M. (2013). Futures, communication and social innovation: Using participatory foresight and social media platforms as tools for evaluating images of the future among young people. *European Journal of Futures Research*, 1, Article 17.
<https://doi.org/10.1007/s40309-013-0017-2>
- Gulya, J. (2023). Rebranding originality for the age of AI. *International Journal of Emerging and Disruptive Innovation in Education: VISIONARIUM*, 1(1), Article 5.
<https://doi.org/10.62608/2831-3550.1014>
- Habermas, J. (1985). *The Theory of Communicative Action, Volume 2: Lifeworld and System: A Critique of Functionalist Reason*. Boston: Beacon Press.
- Habermas, J. (1991). *Moral Consciousness and Communicative Action*. MIT Press.
- Habermas, J. (1998). *Between Facts and Norms: Contributions to a Discourse Theory of Law and Democracy*. MIT Press.
- Hagendijk, R., & Irwin, I. (2006). Public deliberation and governance: Engaging with science and technology in contemporary Europe. *Minerva*, 44(2), 167–184.
<https://doi.org/10.1007/s11024-006-0012-x>
- Haggard, P., & Tsakiris, M. (2009). Peripersonal space: A psychological perspective on the multisensory aspects of the body. *Psychological Bulletin*, 135(3), 440–464.
<https://doi.org/10.1007/s00221-022-06324-8>
- Hall, G. E., & Hord, S. M. (1987). *Change in schools: Facilitating the process*. State University of New York Press.
- Han, B. C. (2014). In M. S. Pereira (Ed.), *Psicopolítica* (Trans.). Relógio D'Água.
- Haraway, D. (1985). A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In *Simians, cyborgs, and women: The reinvention of nature* (pp. 149–181). Routledge.

- Haraway, D. J. (2016). *Staying with the trouble: Making kin in the Chthulucene*. Duke University Press. <https://doi.org/10.1515/9780822373780>
- Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing*, 32, 1008–1015. <https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x>
- Hennink, M. M., Kaiser, B. N., & Weber, M. B. (2019). What influences saturation? Estimating sample sizes in focus group research. *Qualitative Health Research*, 29(10), 1483–1496. <https://doi.org/10.1177/1049732318821692>
- Herzog, C., Leinweber, N. A., Engelhard, S. A., & Engelhard, L. H. (2022). Autonomous ferries and cargo ships: Discovering ethical issues via a challenge-based learning approach in higher education. *IEEE International Symposium on Technology and Society (ISTAS)*, 2022, 1–6. <https://doi.org/10.1109/ISTAS55053.2022.10227124>
- Hicks, D., & Holden, C. (1995). *Visions of the Future: Why we Need to Teach for Tomorrow*. Trentham Books.
- Hilbert, M., Miles, I., & Othmer, J. (2009). Foresight tools for participative policy-making in inter-governmental processes in developing countries: Lessons learned from the eLAC Policy Priorities Delphi. *Technological Forecasting and Social Change*, 76(2), 880–896. <https://doi.org/10.1016/j.techfore.2009.01.001>
- Hines, A. (2019). When did it start? Origin of the foresight field. *World Futures Review*, 12(1), 4–11. <https://doi.org/10.1177/1946756719889053>
- Holmes, W., Bektik, D., Whitelock, D., & Woolf, B. (2018). Ethics in AIED: Who cares? [Workshop presentation]. *19th International Conference on Artificial Intelligence in Education (AIED '18)*, London. https://oro.open.ac.uk/53443/1/AIED_2018_paper_14%20%283%29.pdf
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in education: Towards a community-wide framework. *International*

- Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00239-1>
- Holmes, W., Persson, J., Chounta, I.-A., Wasson, B., & Dimitrova, V. (2022). *Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law*. Council of Europe Publishing. <https://rm.coe.int/artificial-intelligence-and-education-a-critical-view-through-the-lens/1680a886bd>
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57(4), 542–570. <https://doi.org/10.1111/ejed.12533>
- Hood, D., Lemaignan, S., & Dillenbourg, P. (2015). When children teach a robot to write: An autonomous teachable humanoid which uses simulated handwriting. In J. Adams, W. Smart, B. Mutlu, & L. Takayama (Eds.), *ACM/IEEE International Conference on Human-Robot Interaction* (pp. 83–90). Association for Computing Machinery. <https://doi.org/10.1145/2696454.2696479>
- Hrastinski, S., Olofsson, A. D., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., Jaldemark, J., Ryberg, T., Öberg, L.-M., Fuentes, A., Gustafsson, U., Humble, N., Mozelius, P., Sundgren, M., & Utterberg, M. (2019). Critical imaginaries and reflections on artificial intelligence and robots in postdigital K-12 education. *Postdigital Science and Education*, 1(2), 427–445. <https://doi.org/10.1007/s42438-019-00046-x>
- Hudlicka, E. (2016). Virtual affective agents and therapeutic games. In D. D. Luxton (Ed.), *Artificial intelligence in behavioral and mental health care* (pp. 81–115). Elsevier.
- Illich, I. (1971). *Deschooling society*. Harper & Row.
- Inayatullah, S. (1990). Deconstructing and reconstructing the future: Predictive, cultural and critical epistemologies, *Futures*, 22(2), 115-141. [https://doi.org/10.1016/0016-3287\(90\)90077-U](https://doi.org/10.1016/0016-3287(90)90077-U)
- Inayatullah, S. (2013). Futures studies: Theories and methods. In J. Minkinen, A. Toivonen, & J. Heikka (Eds.), *There's a future: Visions for a better world* (pp. 36–66). Telos.
- Initiative for Climate Action Transparency (ICAT). (2020). *Stakeholder participation guide: Supporting stakeholder participation in design, implementation and assessment of policies and actions*. Climate, Community & Biodiversity Alliance and Verra. <https://climateactiontransparency.org/wp-content/uploads/2020/04/Stakeholder-Participation-Assessment-Guide.pdf>

- Jasanoff, S., Kim, S.-H., & Sperling, S. (2007). Socio-technical imaginaries and science and technology policy: A cross-national comparison. *Science as Culture*, 16(3), 135–153. <https://doi.org/10.1080/09505430701463847>
- Javed, R. T., Nasir, O., Borit, M., Vanhee, L., Zea, E., Gupta, S., Vinuesa, R., & Qadir, J. (2022). Get out of the BAG! Silos in AI ethics education: Unsupervised topic modeling analysis of global AI curricula. *Journal of Artificial Intelligence Research*, 73, 933–965. <https://doi.org/10.1613/jair.1.13550>
- Jobin, A., Ienca, M., & Vayena, E. (2019). Artificial intelligence: The global landscape of ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>
- Johnson, L. (1987). Children’s visions of the future. *The Futurist*, 2(1), 36-40.
- Johnson, W. L., & Lester, J. C. (2016). Face-to-face interaction with pedagogical agents, twenty years later. *International Journal of Artificial Intelligence in Education*, 26, 25–36. <https://doi.org/10.1007/s40593-015-0065-9>
- Jones, L. (2007). *The student-centered classroom*. Cambridge University Press.
- Kant, I. (2005). *A metafísica dos costumes*. Fundação Calouste Gulbenkian.
- Karaca, O., Çalışkan, S., & Demir, K. (2021). Medical artificial intelligence readiness scale for medical students (MAIRS-MS) – Development, validity and reliability study. *BMC Medical Education*, 21, Article 565. <https://doi.org/10.1186/s12909-021-02546-6>
- Katznelson, G., & Gerke, S. (2021). The need for health AI ethics in medical school education. *Advances in Health Sciences Education*, 26(4), 1447–1458. <https://doi.org/10.1007/s10459-021-10040-3>
- Kelly, A. E. (2006). Quality criteria for design research: Evidence and commitments. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 107–118). Routledge.
- Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable artificial intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Khurana, M. P. (2020). Keeping pace: The need for digital health education in medical schools. *Academic Medicine*, 95(10), 1629–1630. <https://doi.org/10.1097/acm.0000000000003672>

- Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Journal of Educational Technology*, 50(6), 2855–2870. <https://doi.org/10.1111/bjet.12868>
- Kong, S.-C., Cheung, W. M.-Y., & Zhang, G. (2023). Evaluating an artificial intelligence literacy program for developing university students' conceptual understanding, literacy, empowerment and ethical awareness. *Educational Technology & Society*, 26(1), 16–30. <https://www.jstor.org/stable/48707964>
- Kosow, H., & Gaßner, R. (2008). *Methods of future and scenario analysis: Overview, assessment, and selection criteria*. German Development Institute. <https://www.idos-research.de/en/studies/article/methods-of-future-and-scenario-analysis-overview-assessment-and-selection-criteria/>
- Kowalczuk-Waledziak, M., Korzeniecka-Bondar, A., Danilewicz, W., & Lauwers, G. (Eds.). (2019). *Rethinking teacher education for the 21st century: Trends, challenges and new directions*. Verlag Barbara Budrich. <https://library.oapen.org/bitstream/handle/20.500.12657/23733/1006411.pdf?sequence=1&isAllowed=y>
- Kreibich, R. (2007). Wissenschaftsverständnis und Methodik der Zukunftsforschung. *Zeitschrift für Semiotik*, 29(2–3), 177–198.
- Krueger, R. A., & Casey, M. A. (2015). *Focus groups: A practical guide for applied research* (5th ed.). Sage Publications.
- Kukulska-Hulme, A., Beirne, E., Conole, G., Costello, E., Coughlan, T., Ferguson, R., FitzGerald, E., Gaved, M., Herodotou, C., Holmes, W., Mac Lochlainn, C., Nic Giolla Mhichíl, M., Rienties, B., Sargent, J., Scanlon, E., Sharples, M., & Whitelock, D. (2020). *Innovating pedagogy 2020: Open University innovation report 8*. The Open University.
- Lados, G., Hegedus, G., & Kovacs, Z. (2022). The role of identity and ontological (in)security in return migration: An empirical perspective from Hungary. *Journal of International Migration and Integration*, 24(2), 1–20. <https://doi.org/10.1007/s12134-022-00964-z>
- Lameras, P., & Arnab, S. (2022). Power to the teachers: An exploratory review on artificial intelligence in education. *Information*, 13(1), 14. <https://doi.org/10.3390/info13010014>
- Landeta, J. (1999). *El método Delphi: Una técnica de previsión para la incertidumbre*. Ariel.
- Latham, A., & Goltz, S. (2019). A survey of the general public's views on the ethics of using AI in education. In S. Isotani et al. (Eds.), *Artificial Intelligence in Education: 20th*

- International Conference, AIED 2019 Proceedings, Part I* (pp. 194–206). Springer.
https://doi.org/10.1007/978-3-030-23204-7_17
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- Latour, B. (2014). Agency at the time of the Anthropocene. *New Literary History*, 45(1), 1–18.
<https://doi.org/10.1353/nlh.2014.0003>
- Latour, B. (2018). *Down to earth: Politics in the new climatic regime* (C. Porter, Trans.). Polity Press.
- Lazarus, M. D., Truong, M., Douglas, P., & Selwyn, N. (2022). Artificial intelligence and clinical anatomical education: Promises and perils. *Anatomical Sciences Education*. Advance online publication. <https://doi.org/10.1002/ase.2221>
- Learning Forward. (2022). *Standards for professional learning*.
<https://www.learningforward.org/standards-for-professional-learning>
- Lee, I., Ali, S., Zhang, H., DiPaola, D., & Breazeal, C. (2021, March). Developing middle school students' AI literacy [Paper presentation]. *SIGCSE '21: Proceedings of the 52nd ACM Technical Symposium on Computer Science Education*, Virtual Event, USA.
<https://doi.org/10.1145/3408877.3432513>
- Lee, N. (2014). *Digital Da Vinci: Computers in the arts and sciences*. Springer.
<https://doi.org/10.1007/978-1-4939-0965-0>
- Leonidis, A., Korozi, M., Margetis, G., Ntoa, S., Papagiannakis, H., Antona, M., & Stephanidis, C. (2012). A glimpse into the ambient classroom. *Bulletin of the IEEE Technical Committee on Learning Technology*, 14(1), 3–6.
- Levinas, E. (1969). *Totality and infinity: An essay on exteriority* (A. Lingis, Trans.). Duquesne University Press.
- Li, Y., Li, S., & Wang, L. (2021). The integration development of artificial intelligence and education. In *Proceedings of the 16th International Conference on Computer Science & Education* (pp. 994–997). IEEE.
<https://doi.org/10.1109/ICCSE51940.2021.9569551>
- Liu, D. (2023). How Sydney educators are building 'AI doubles' of themselves to help their students. *Teaching@Sydney Homepage*. <https://educational-innovation.sydney.edu.au/teaching@sydney/how-sydney-educators-are-building-ai-doubles-of-themselves-to-help-their-students/>

- Liu, S., & Kunnan, A. (2016). Investigating the application of automated writing evaluation to Chinese undergraduate English majors: A case study of WriteToLearn. *CALICO Journal*, 33(1), 71–91. <https://doi.org/10.1558/cj.v33i1.26380>
- Loftus, M., & Madden, M. G. (2020). A pedagogy of data and artificial intelligence for student subjectification. *Teaching in Higher Education*, 25(4), 456–475. <https://doi.org/10.1080/13562517.2020.1748593>
- Long, T., & Johnson, M. (2000). Rigour, reliability and validity in qualitative research. *Clinical Effectiveness in Nursing*, 4(1), 30–37. <https://doi.org/10.1054/cein.2000.0106>
- Lucic, A., Bleeker, M., Jullien, S., Bhargav, S., & de Rijke, M. (2022). Reproducibility as a mechanism for teaching fairness, accountability, confidentiality, and transparency in artificial intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(11), 12792–12800. <https://doi.org/10.1609/aaai.v36i11.21558>
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.
- Luo, S. (2019). Research on the change of educational management in the era of artificial intelligence. In Z. Hou (Ed.), *Proceedings of the 12th International Conference on Intelligent Computation Technology and Automation* (pp. 442–445). IEEE. <https://doi.org/10.1109/ICICTA49267.2019.00101>
- Lynch, K. (2022). *Care and capitalism*. Polity Press.
- Malabou, C. (2021). *Métamorphoses de l'intelligence*. Presses Universitaires de France/Humensis.
- Mannermaa, M. (1992). *Evolutionary futures research: Searching qualifications of futures research's paradigms and methodologies* [in Finnish]. *Acta Futura Fennica* (No. 2, pp. 23–29). VapK-Kustannus.
- Marras, M., Boratto, L., Ramos, G., & Fenu, G. (2022). Equality of learning opportunity via individual fairness in personalized recommendations. *International Journal of Artificial Intelligence in Education*, 32, 636–684. <https://doi.org/10.1007/s40593-021-00271-1>
- Martelli, A. (2014). *Models of Scenario Building and Planning: Facing Uncertainty and Complexity*. Palgrave Macmillan.
- Marx, K., & Engels, F. (1976). *The German ideology* (C. J. Arthur, Ed.). International Publishers.
- Masini, E. (1993). *Why Future Studies?* Grey Seal.

- Mason, J., Peoples, B. E., & Lee, J. (2020). Questioning the scope of AI standardization in learning, education, and training. *Journal of ICT Standardization*, 8(2), 107–122. <https://doi.org/10.13052/jicts2245-800X.822>
- McDonald, N., & Pan, S. (2020). Intersectional AI: A study of how information science students think about ethics and their impact. *Proceedings of the ACM on Human-Computer Interaction*, 4(2), 1–19. <https://doi.org/10.1145/3415218>
- McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.). Routledge. <https://doi.org/10.4324/9781315105642>
- Metzinger, T. (2019, April 8). *Ethics washing made in Europe*. Der Tagesspiegel. <https://www.tagesspiegel.de/politik/ethics-washing-made-in-europe-5937028.html>
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- Miller, R., & Tuomi, I. (2022). Making the futures of AI in education: Why and how imagining the future matters. *European Journal of Education*, 57(4), 537–541. <https://doi.org/10.1111/ejed.12529>
- Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, 108(6), 1017–1054. <https://doi.org/10.1111/j.1467-9620.2006.00684.x>
- Mohammed, P. S., & Watson, E. (2019). Towards inclusive education in the age of artificial intelligence: Perspectives, challenges, and opportunities. In J. Knox, Y. Wang, & M. Gallagher (Eds.), *Artificial intelligence and inclusive education: Speculative futures and emerging practices* (pp. 17–37). Springer. https://doi.org/10.1007/978-981-13-8161-4_2
- Mollenhauer, K. (2014). *Forgotten connections: On culture and upbringing* (N. Friesen, Trans.). Routledge. (Original work published 1983)
- Mollick, E. R., Mollick, L., Bach, N., Ciccarelli, L. J., Przystanski, B., & Ravipinto, D. (2024). AI agents and education: Simulated practice at scale. *The Wharton School Research Paper*. <https://doi.org/10.2139/ssrn.4871171>
- Montuori, A. (1999). Planetary culture and the crisis of the future. *World Futures: The Journal of General Evolution*, 54(4), 297–311. <https://doi.org/10.1080/02604027.1999.9972761>

- Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology*, 7, Article 1272. <https://doi.org/10.3389/fpsyg.2016.01272>
- Morais, A. M., Neves, I. P., & Afonso, M. (2005). *Teacher training processes and the transformation of pedagogic discourse: A contribution to the development of teacher education*. EDUCA. <https://repositorio.ulisboa.pt/handle/10451/4391>
- Mousavinasab, E., Zarifsanaiey, N., Niakan Kalhori, R. S., Rakhshan, M., Keikha, L., & Ghazi Saeedi, M. (2018). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, 29(1), 142–163. <https://doi.org/10.1080/10494820.2018.1558257>
- Mouta, A., Paulino, A., Ferreira, J., & Couto, F. (2015). Pedagogy as a techné: Meaningful ICT integration in formal learning scenarios. In G. Roman, J. Komuves, & R. Hahn (Eds.), *Proceedings of the European Distance and E-Learning Network 2015 Annual Conference* (pp. 64–72). EDEN.
- Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2019). Blending machines, learning, and ethics. In M. González, F. Sedano, C. Llamas, & F. García-Peñalvo (Eds.), *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 993–998). Association for Computing Machinery. <https://doi.org/10.1145/3362789.3362909>
- Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2020). Sense of agency in times of automation: A teachers' professional development proposal on the ethical challenges of AI applied to education. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial Intelligence in Education* (pp. 405–408). Springer. https://doi.org/10.1007/978-3-030-52240-7_73
- Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2021). Blending machines, learning, sense of agency, and ethics: Designing an in-depth framework with experts using the Delphi Method approach. In M. Alier & D. Fonseca (Eds.), *Proceedings of the Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 665–670). Association for Computing Machinery. <https://doi.org/10.1145/3486011.3486545>
- Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2024a). Uncovering blind spots in education ethics: Insights from a systematic literature review on Artificial

- Intelligence in education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-023-00384-9>
- Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2024b). Design of a future scenarios toolkit for an ethical implementation of Artificial Intelligence in education. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12229-y>
- Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2025a). “Where is agency moving to?”: Exploring the interplay between AI technologies in education and human agency. *Digital Society*, 4(49). <https://doi.org/10.1007/s44206-025-00203-9>
- Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2025b). Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12946-y>
- Nagle, B., & Williams, N. (2013). Methodology brief: Introduction to the focus group. *Center for Assessment, Planning and Accountability*. <http://www.mmgconnect.com/projects/userfiles/file/focusgroupbrief.pdf>
- Nesta. (2019, January 22). *Mapping global approaches to AI governance*. <https://www.nesta.org.uk/blog/mapping-global-approaches-ai-governance/>
- Ng, D. T. K., Lee, M., Tan, R. J. Y., Hu, X., Downie, J. S., & Chu, S. K. W. (2023). A review of AI teaching and learning from 2000 to 2020. *Education and Information Technologies*, 28, 8445–8501. <https://doi.org/10.1007/s10639-022-11491-w>
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2022). Ethical principles for artificial intelligence in education. *Educational Technology Research and Development*. <https://doi.org/10.1007/s10639-022-11316-w>
- Nussbaum, M. C. (2011). *Creating capabilities: The human development approach*. Belknap Press of Harvard University Press.
- Nuwan, P. M. M. C., Perera, B. A. K. S., & Dewagoda, K. G. (2021). Development of core competencies of construction managers: The effect of training and education. *Tech Know Learn*, 26, 945–984. <https://doi.org/10.1007/s10758-020-09474-2>
- Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, 25(2), 177–203. <https://doi.org/10.1007/s40593-014-0028-6>

- Nygaard, K., Graversgaard, M., Dalgaard, T., Jacobsen, B. H., & Schaper, S. (2021). The role of stakeholder engagement in developing new technologies and innovation for nitrogen reduction in waters: A longitudinal study. *Water*, *13*(22), 3313. <https://doi.org/10.3390/w13223313>
- Obhi, S. S., & Hall, P. (2011). Sense of agency in joint action: Influence of human and computer co-actors. *Experimental Brain Research*, *211*(3–4), 663–670. <https://doi.org/10.1007/s00221-011-2662-7>
- Obhi, S. S., Swiderski, K. M., & Brubacher, S. P. (2012). Induced power changes the sense of agency. *Consciousness and Cognition*, *21*(3), 1547–1550. <https://doi.org/10.1016/j.concog.2012.06.008>
- OECD. (2012). *PISA 2012 database*. <http://www.oecd.org/pisa/pisaproducts/pisa2012database-downloadabledata.htm>
- OECD. (2018). *Curriculum flexibility and autonomy in Portugal – An OECD review*. [https://www.poch.portugal2020.pt/pt-pt/Candidaturas/Documents/CurriculumFlexibility-and-Autonomy-in-Portugal-an-OECD-Review%20\(002\).pdf](https://www.poch.portugal2020.pt/pt-pt/Candidaturas/Documents/CurriculumFlexibility-and-Autonomy-in-Portugal-an-OECD-Review%20(002).pdf)
- OECD. (2021). *OECD Digital Education Outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD Publishing. <https://doi.org/10.1787/589b283f-en>
- OECD. (2023). *OECD Digital Education Outlook 2023: Towards an effective digital education ecosystem*. OECD Publishing. <https://doi.org/10.1787/c74f03de-en>
- Ogan, A., Yarzebinski, E., Fernández, P., & Casas, I. (2015). Cognitive tutor use in Chile: Understanding classroom and lab culture. In C. Conati, N. Heffernan, A. Mitrovic, & M. F. Verdejo (Eds.), *Artificial Intelligence in Education* (Vol. 9112). Springer.
- Ogbeifun, E., Agwa-Ejon, J., Mbohwa, C., & Pretorius, J. H. (2016, March). *The Delphi technique: A credible research methodology* [Conference presentation]. 2016 International Conference on Industrial Engineering and Operations Management, Kuala Lumpur, Malaysia.
- O'Shea, L. (2019). *Future histories: What Ada Lovelace, Tom Paine, and the Paris Commune can teach us about digital technology*. Verso.
- Ottenbreit-Leftwich, A., Glazewski, K., Jeon, M., Jantaraweragul, K., Hmelo-Silver, C., Scribner, A., Lee, S., Mott, B., & Lester, J. (2023). Lessons learned for AI education

- with elementary students and teachers. *International Journal of Artificial Intelligence in Education*, 33, 267–289. <https://doi.org/10.1007/s40593-022-00304-3>
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ*, 372, n71. <https://doi.org/10.1136/bmj.n71>
- Pammer-Schindler, V., & Rosé, C. (2022). Data-related ethics issues in technologies for informal professional learning. *International Journal of Artificial Intelligence in Education*, 32, 609–635. <https://doi.org/10.1007/s40593-021-00259-x>
- Pantelimon, F.-V., Bologa, R., Toma, A., & Posedaru, B.-S. (2021). The evolution of AI-driven educational systems during the COVID-19 pandemic. *Sustainability*, 13(23), 13501. <https://doi.org/10.3390/su132313501>
- Papa, R., & Jackson, K. M. (2021). Enduring questions, innovative technologies: Educational theories interface with AI. In K. Arai (Ed.), *Intelligent computing – Proceedings of the 2021 computing conference* (pp. 725–742). Springer. https://doi.org/10.1007/978-3-030-80126-7_51
- Pasquinelli, M. (2023). *The eye of the master: A social history of artificial intelligence*. Verso Books.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE.
- Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons.
- Pham, P., & Wang, J. (2017). AttentiveLearner2: A multimodal approach for improving MOOC learning on mobile devices. In E. André, R. Baker, X. Hu, M. Rodrigo, & B. du Boulay (Eds.), *Artificial intelligence in education. AIED 2017. Lecture notes in computer science* (pp. 561–564). Springer. https://doi.org/10.1007/978-3-319-61425-0_64
- Piaget, J. (1952). *The origins of intelligence in children*. International Universities Press.
- Pinkwart, N. (2016). Another 25 years of AIED? Challenges and opportunities for intelligent educational technologies of the future. *International Journal of Artificial Intelligence in Education*, 26, 771–783. <https://doi.org/10.1007/s40593-016-0099-7>

- Plomp, T., & Nieveen, N. (2013). *Educational design research* (Vol. 1). SLO.
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Porayska-Pomsta, K., Holmes, W., & Nemorin, S. (2023). The ethics of AI in education. In *Handbook of Artificial Intelligence in Education* (pp. 571–604). Edward Elgar Publishing. <https://www.elgaronline.com/edcollchap/book/9781800375413/book-part-9781800375413-38.xml>
- Powell, C. (2003). The Delphi technique: Myths and realities. *Journal of Advanced Nursing*, 41(4), 376–382. <https://doi.org/10.1046/j.1365-2648.2003.02537.x>
- Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury.
- Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room: The obligation to act. In M. Hatala (Ed.), *Proceedings of the seventh international learning analytics & knowledge conference* (pp. 46–55). Association for Computing Machinery. <https://doi.org/10.1145/3027385.3027406>
- Qin, F., Li, K., & Yan, J. (2020). Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology*, 51(5), 1693–1710. <https://doi.org/10.1111/bjet.12994>
- Radford, J., Richard, G., Richard, H., & Serrurier, M. (2021). Detecting dyslexia from audio records: An AI approach. In C. Gehin, B. Wacogne, A. Fred, & H. Gamboa (Eds.), *Proceedings of the 14th international joint conference on biomedical engineering systems and technologies – HEALTHINF* (pp. 58–66). SciTePress. <https://doi.org/10.5220/0010196000580066>
- Raise. (2024). *Responsible AI for social empowerment and education*. <https://raise.mit.edu/>
- Ramírez-Montoya, M., & García-Peñalvo, F. (2018). Co-creation and open innovation: Systematic literature review [Co-creación e innovación abierta: Revisión sistemática de literatura]. *Comunicar*, 54, 9–18. <https://doi.org/10.3916/C54-2018-01>
- Rausch, E. (2003). *Simulation and Games in Futuring and Other Uses* [CD ROM]. The Millennium Project.
- Redman, E., Wiek, A., & Redman, A. (2018). Continuing professional development in sustainability education for K–12 teachers: Principles, programme, applications,

- outlook. *Journal of Education for Sustainable Development*, 12(1), 59–80.
<https://doi.org/10.1177/2455133318777182>
- Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. *European Journal of Operational Research*, 291(3), 906–917.
<https://doi.org/10.1016/j.ejor.2020.09.020>
- Rello, L., Williams, K., Ali, A., White, N. C., & Bigham, J. P. (2016). Dydetective: Towards detecting dyslexia across languages using an online game. In L. Rello & V. Conway (Eds.), *Proceedings of the 13th Web for All Conference* (pp. 1–4). Association for Computing Machinery. <https://doi.org/10.1145/2899475.2899491>
- Renz, A., & Vladova, G. (2021). Reinvigorating the discourse on human-centered artificial intelligence in educational technologies. *Technology Innovation Management Review*, 11, 5–16. <https://doi.org/10.22215/timreview/1438>
- Renzi, A., & Freitas, S. (2015). The Delphi method for future scenarios construction. *Procedia Manufacturing*, 3, 5785–5791. <https://doi.org/10.1016/j.promfg.2015.07.826>
- Restrepo, E. G., Baldassarre, M., & Boticario, J. G. (2019). Accessibility, biases and ethics in chatbots and intelligent agents for education. In L. G. Chova, A. L. Martínez, & I. C. Torres (Eds.), *EDULEARN19 proceedings* (pp. 8824–8833). IATED Academy. <https://doi.org/10.21125/edulearn.2019.2196>
- Richards, D., & Dignum, V. (2019). Supporting and challenging learners through pedagogical agents: Addressing ethical issues through designing for values. *British Journal of Educational Technology*, 50(6), 2885–2901. <https://doi.org/10.1111/bjet.12863>
- Roberts, K., Dowell, A., & Nie, J. B. (2019). Attempting rigour and replicability in thematic analysis of qualitative research data: A case study of codebook development. *BMC Medical Research Methodology*, 19, 66. <https://doi.org/10.1186/s12874-019-0707-y>
- Rogers, C. R. (1983). *Freedom to learn for the 80's*. Charles E. Merrill Publishing Company.
- Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.
- Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26, 582–599.
<https://doi.org/10.1007/s40593-016-0110-3>
- Rubie-Davies, C. M. (2015). *Becoming a high expectation teacher: Raising the bar*. Routledge.

- Rubin, A. (1996). *Unfolding tomorrow: Adolescents' images of the future as strategies for coping with transition* [Paper presentation]. Cultural Alienation Conference, Losarvi, Finland.
- Salgado, M., & Gilbert, N. (2013). Agent based modelling. In T. Teo (Ed.), *Handbook of quantitative methods for educational research* (pp. 247-265). Sense Publishers. https://doi.org/10.1007/978-94-6209-404-8_12
- Saltman, K. J. (2020). Artificial intelligence and the technological turn of public education privatization: In defense of democratic education. *London Review of Education*, 18(2), 196–208. <https://doi.org/10.14324/LRE.18.2.04>
- Sanromà-Giménez, M., Lázaro Cantabrana, J. L., Usart Rodríguez, M., & Gisbert-Cervera, M. (2021). Design and validation of an assessment tool for educational mobile applications used with autistic learners. *Journal of New Approaches in Educational Research*, 10(1), 101–121. <https://doi.org/10.7821/naer.2021.1.574>
- Santos, J. M. (2012). *Introdução à ética*. Sistema Solar.
- Sanusi, I. T., & Olaleye, S. A. (2022). An insight into cultural competence and ethics in K-12 artificial intelligence education. *IEEE Global Engineering Education Conference (EDUCON), 2022*, 790–794. <https://doi.org/10.1109/EDUCON52537.2022.9766818>
- Sartre, J. P. (1943). *Being and nothingness: A phenomenological essay on ontology* (H. E. Barnes, Trans.). Washington Square Press.
- Schiff, D. (2021). Education for AI, not AI for education: The role of education and ethics in national AI policy strategies. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00270-2>
- Schultz, W. (2015). *A brief history of futures*. *World Futures Review*, 7(4), 321–334. <https://doi.org/10.1177/1946756715627646>
- Schwenker, B., & Wulf, T. (2013). *Scenario-Based Strategic Planning: Developing Strategies in an Uncertain World*. Springer Gabler.
- Seker, S. E. (2015). Computerized argument Delphi technique. *IEEE Access*, 3, 368–380. <https://doi.org/10.1109/ACCESS.2015.2424703>
- Selwyn, N. (2016). *Education and Technology: Key Issues and Debates* (2nd ed.). Bloomsbury Academic.

- Serholt, S., Barendregt, W., Vasalou, A., Alves-Oliveira, P., Jones, A., Petisca, S., & Paiva, A. (2017). The case of classroom robots: Teachers' deliberations on the ethical tensions. *AI & Society*, 32(4), 613–631. <https://doi.org/10.1007/s00146-016-0667-2>
- Severance, S., Penuel, W., Sumner, T., & Leary, H. (2016). Organising for teacher agency in curricular co-design. *Journal of the Learning Sciences*, 25, 531–564. <https://doi.org/10.1080/10508406.2016.1207541>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Houghton Mifflin.
- Shih, P.-K., Lin, C.-H., Wu, L., & Yu, C.-C. (2021). Learning ethics in AI—Teaching non-engineering undergraduates through situated learning. *Sustainability*, 13(7), 3718. <https://doi.org/10.3390/su13073718>
- Shum, S. B. (2018). Transitioning education's knowledge infrastructure: Shaping design or shouting from the touchline? In J. Kay & R. Luckin (Eds.), *Proceedings of the 13th International Conference of the Learning Sciences* (p. 5). International Society of the Learning Sciences. <https://hdl.handle.net/10453/133232>
- Silva, I. S., Veloso, A. L., & Keating, J. B. (2014). Focus group: Considerações teóricas e metodológicas. *Revista Lusófona de Educação*, 26, 175–190.
- Silva, P. C., & Fortunato, M. (2021). Modera, observa, escuta e foca-te na conversa de grupo – uma reflexão crítica. In P. Sá, A. P. Costa, & A. Moreira (Eds.), *Reflexões em torno de metodologias de investigação: Recolha de dados* (Vol. 2, pp. 9–10). UA Editora. https://ria.ua.pt/bitstream/10773/30772/3/Metodologias%20investigacao_Vol2_Digital.pdf
- Simondon, G. (2008). *Imagination et invention (1965–1966)*. La Transparence.
- Simonova, P., Cincera, J., Kroufek, R., Krepelkova, S., & Hadjichambis, A. (2019). Active Citizens: Evaluation of a Community-Based Education Program. *Sustainability*, 11(3), 663. <https://doi.org/10.3390/su11030663>
- Sjödén, B. (2020). When lying, hiding and deceiving promotes learning – A case for augmented intelligence with augmented ethics. In I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, & E. Millán (Eds.), *Artificial intelligence in education* (pp. 291–295). Springer. https://doi.org/10.1007/978-3-030-52240-7_53
- Slaughter, R. (1998). The knowledge base of futures studies. In D. Hicks & R. Slaughter (Eds.), *World yearbook of education 1998: Futures education* (pp. 73–84). Kogan Page.

- Slaughter, R. (1999). Professional standards in futures work. *Futures*, 31, 835–851. [https://doi.org/10.1016/S0016-3287\(99\)00039-7](https://doi.org/10.1016/S0016-3287(99)00039-7)
- Slaughter, R. (2002). From rhetoric to reality: The emergence of futures into the educational mainstream. In J. Gidley & S. Inayatullah (Eds.), *Youth futures: Comparative research and transformative visions* (pp. 175–186). Praeger. https://richardslaughter.com.au/wp-content/uploads/2018/03/Emergence_of_Futs_into_Edn_2004.pdf
- Slaughter, R. (2003). *Futures beyond dystopia: Creating social foresight*. Routledge.
- Slaughter, R. (2019). Futures studies as a quest for meaning. *World Futures Review*, 12(1), 26–39. <https://doi.org/10.1177/1946756719870277>
- Slaughter, R. A. (2020). Futures studies as a quest for meaning. *World Futures Review*, 12(1), 26–39. <https://doi.org/10.1177/1946756719870277>
- Smakman, M. H. J., Konijn, E. A., Vogt, P., & Pankowska, P. (2021). Attitudes towards social robots in education: Enthusiast, practical, troubled, sceptic, and mindfully positive. *Robotics*, 10(1), 24. <https://doi.org/10.3390/robotics10010024>
- Son, H. (2015). The history of Western futures studies: An exploration of the intellectual traditions and three-phase periodization. *Futures*, 66, 120–137. <https://doi.org/10.1016/j.futures.2014.12.013>
- Srinivasan, R., & Uchino, K. (2021). The role of arts in shaping AI ethics. In A. Fokoue, C. Agunwa, K. Lee, L. T. Quigley, & S. Hobson (Eds.), *CEUR workshop proceedings* (pp. 1–6). CEUR. https://ceur-ws.org/Vol-2812/RDAI-2021_paper_3.pdf
- Stein, Z., Engberg, J., & Brown, T. (2017). The seductions of quantification: Measuring human rights, gender violence, and sex trafficking. *Signs: Journal of Women in Culture and Society*, 42(2), 397–422. <https://doi.org/10.1086/690937>
- Steiner, R. (1982). *The kingdom of childhood: Lectures, 1924*. Anthroposophic Press.
- Steiner, R. (1996). *Education of the child: And early lectures on education*. Anthroposophic Press.
- Stiegler, B. (1998). *Technics and time, 1: The fault of Epimetheus* (R. Beardsworth & G. Collins, Trans.). Stanford University Press.
- Stiegler, B. (2015). *States of shock: Stupidity and knowledge in the 21st century* (D. Ross, Trans.). Polity Press.

- Sudina, E., & Plonsky, L. (2024). The effects of frequency, duration, and intensity on L2 learning through Duolingo: A natural experiment. *Journal of Second Language Studies*, 7(1), 1–43. <https://doi.org/10.1075/jsls.00021.plo>
- The Institute for Ethical AI in Education. (2021). *The ethical framework for AI in education*. <https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>
- Thompson, G. (2021). *The global report on the status of teachers 2021*. Education International. <https://www.ei-ie.org/file/230>
- Timms, M. J. (2016). Letting artificial intelligence in education out of the box: Educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education*, 26, 701–712. <https://doi.org/10.1007/s40593-016-0095-y>
- Toffler, A. (1974). *Learning for Tomorrow: The Role of the future in Education*. Vintage Books.
- Tolon, K. (2011). *The American futures studies movement (1965–1975): Its roots, motivations, and influences* (Publication No. 12029) [Doctoral dissertation, Iowa State University]. Iowa State University Digital Repository. <https://lib.dr.iastate.edu/etd/12029>
- Trindade, R., & Cosme, A. (2010). *Educar e aprender na escola: Questões, perspectivas e respostas pedagógicas*. Fundação Manuel Leão.
- Trindade, R. (2012). *O movimento da educação nova e a reinvenção da escola: Da afirmação de uma necessidade aos equívocos de um desejo*. U. Porto Editorial.
- Tundrea, E. (2020). Artificial intelligence in higher education: Challenges and opportunities. In L. G. Chova, A. L. Martínez, & I. C. Torres (Eds.), *INTED2020 proceedings* (pp. 2041–2049). IATED Academy. <https://doi.org/10.21125/inted.2020.0644>
- Tuomi, I. (2018). The impact of artificial intelligence on learning, teaching, and education: Policies for the future. In M. Cabrera, R. Vuorikari, & Y. Punie (Eds.), *Joint Research Centre (European Commission)*. <https://doi.org/10.2760/12297>
- Tuomi, I. (2024, September 12). *Disturbing futures: Towards a theory of futures literacy* [Conference presentation]. Anticipation 2024 Conference, Oslo, Norway.
- Tuovinen, L., & Rohunen, A. (2021). Teaching AI ethics to engineering students: Reflections on syllabus design and teaching methods. In J. Koskinen, M. M. Rantanen, A.-M. Tuikka, & S. Knaapi-Junnila (Eds.), *Proceedings of the conference on technology ethics* (pp. 19–33). CEUR. https://ceur-ws.org/Vol-3069/FP_02.pdf

- UNESCO. (2012). *International standard classification of education (ISCED) 2011*. United Nations Educational, Scientific and Cultural Organization. <https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>
- UNESCO. (2015). *SDG 4 indicators*. <https://www.education-progress.org/en/indicators>
- UNESCO. (2019a). *Beijing consensus on artificial intelligence and education*. <https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- UNESCO. (2019b). *Steering AI and advanced ICTs for knowledge societies: A rights, openness, access, and multistakeholder perspective*. <https://unesdoc.unesco.org/ark:/48223/pf0000372132>
- UNESCO. (2021). *Reimagining our futures together: A new social contract for education*. <https://creativecommons.org/licenses/by-sa/3.0/igo/>
- UNESCO. (2022). *K–12 AI curricula: A mapping of government-endorsed AI curricula*. <https://unesdoc.unesco.org/ark:/48223/pf0000380602>
- UNESCO. (2023a). *Guidance for generative AI in education and research*. <https://doi.org/10.54675/EWZM9535>
- UNESCO. (2023b). *Draft AI competency frameworks for teachers and for school students*. <https://www.unesco.org/en/digital-education/ai-future-learning/competency-frameworks>
- UNESCO. (2023c). *AI competency framework for teachers*. UNESCO. <https://doi.org/10.54675/ZJTE2084>
- Ungerer, L., & Slade, S. (2022). Ethical considerations of artificial intelligence in learning analytics in distance education contexts. In R. K. Atkinson, C. A. Lee, M. D. Brown, & H. R. Collier (Eds.), *Proceedings of the 2022 International Conference on Advanced Learning Technologies* (pp. 97–105). Springer. https://doi.org/10.1007/978-981-19-0786-9_8
- UNHCR. (2016). *Connecting refugees: How internet and technology can improve refugee well-being and self-reliance*. https://www.unhcr.org/innovation/wp-content/uploads/2018/02/20160707-Connecting-Refugees-Web_with-signature.pdf
- U.S. Department of Education, Office of Educational Technology. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*.

- Washington, DC: U.S. Department of Education.
<https://www.ed.gov/sites/default/files/documents/ai-report/ai-report.pdf>
- Vallor, S. (2016). *Technology and the virtues: A philosophical guide to a future worth wanting*. Oxford University Press.
- van Notten, P. (2006). Scenario development: A typology of approaches. In OECD (Ed.), *Think scenarios, rethink education* (pp. 69–92). OECD Publishing.
- Vera, M., Le Blanc, P., Taris, T., & Salanova, M. (2014). Patterns of engagement: The relationship between efficacy beliefs and task engagement at the individual versus collective level. *Journal of Applied Social Psychology, 44*, 133–144.
<https://doi.org/10.1111/jasp.12219>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. & Trans.). Harvard University Press.
- Wang, H., Tlili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X., & Fei, C. (2023). Examining the applications of intelligent tutoring systems in real educational contexts: A systematic literature review from the social experiment perspective. *Education and Information Technologies*. Advance online publication.
<https://doi.org/10.1007/s10639-022-11555-x>
- Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences, 13*(11), 6716. <https://doi.org/10.3390/app13116716>
- Warner, B. P., & Elser, M. (2015). How Do Sustainable Schools Integrate Sustainability Education? An Assessment of Certified Sustainable K–12 Schools in the United States. *The Journal of Environmental Education, 46*(1), 1–22.
<https://doi.org/10.1080/00958964.2014.953020>
- Watters, A. (2021). *Teaching machines: The history of personalized learning*. The MIT Press.
- Weber, A. (2020). Ethics concerns in artificial intelligence use in education. In L. Gómez Chova, A. López Martínez, & I. Candel Torres (Eds.), *INTED2020 Proceedings* (pp. 4539–4544). IATED Academy. <https://doi.org/10.21125/inted.2020.1262>

- Westman, S., Kauttonen, J., Klemetti, A., Korhonen, N., Manninen, M., Mononen, O., Niittymäki, S., & Paananen, H. (2021). Artificial intelligence for career guidance – Current requirements and prospects for the future. *IAFOR Journal of Education: Technology in Education*, 9(4), 43–62. <https://doi.org/10.22492/ije.9.4.03>
- Williams, R., Machado, C. V., Druga, S., Breazeal, C., & Maes, P. (2018). “My doll says it’s OK”: A study of children’s conformity to a talking doll. In M. Høiseth, G. E. Baykal, A. Peter, & B. Farschian (Eds.), *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 625–631). Association for Computing Machinery. <https://doi.org/10.1145/3202185.3210788>
- Williams, R., & Breazeal, C. (2020, July). How to train your robot: A middle school AI and ethics curriculum [Session]. *International Workshop on Education in Artificial Intelligence K–12 (EDUAI '20)*, Virtual Event. https://dam-prod.media.mit.edu/x/2020/06/19/EduAI_2020.pdf
- Williams, T., Zhu, Q., & Grollman, D. (2020). An experimental ethics approach to robot ethics education. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(9), 13428–13435. <https://doi.org/10.1609/aaai.v34i09.7067>
- Williamson, B. (2022). Big EdTech. *Learning, Media and Technology*, 47(2), 157–162. <https://doi.org/10.1080/17439884.2022.2063888>
- Williamson, B. (2024). The social life of AI in education. *International Journal of Artificial Intelligence in Education*, 34(1), 97–104. <https://doi.org/10.1007/s40593-023-00342-5>
- Wright, D., Finn, R., Gellert, R., Gutwirth, S., Schütz, P., Friedewald, M., Venier, S., & Mordini, E. (2014). Ethical dilemma scenarios and emerging technologies. *Technological Forecasting and Social Change*, 87, 325–336. <https://doi.org/10.1016/j.techfore.2013.12.008>
- Wright, E. O. (2010). *Envisioning real utopias*. Verso.
- Xu, L. (2020). The dilemma and countermeasures of AI in educational application. In Y. Tan (Ed.), *Proceedings of the 2020 4th International Conference on Computer Science and Artificial Intelligence* (pp. 289–294). Association for Computing Machinery. <https://doi.org/10.1145/3445815.3445863>
- Xu, S., Zhang, X., & Qin, L. (2024). EduAgent: Generative student agents in learning. *arXiv*. <https://doi.org/10.48550/arXiv.2404.07963>

- Yau, K. W., Chai, C. S., Chiu, T. K., Meng, H. M., King, I., & Yam, Y. (2022). A phenomenographic approach on teacher conceptions of teaching artificial intelligence (AI) in K–12 schools. *Education and Information Technologies*, 28, 1041–1064. <https://doi.org/10.1007/s10639-022-11161-x>
- Zagouras, C., Egarchou, D., Skiniotis, P., & others. (2022). Face to face or blended learning? A case study: Teacher training in the pedagogical use of ICT. *Education and Information Technologies*, 27, 12939–12967. <https://doi.org/10.1007/s10639-022-11144-y>
- Zhang, H., Lee, I., Ali, S., DiPaola, D., Cheng, Y., & Breazeal, C. (2023). Integrating ethics and career futures with technical learning to promote AI literacy for middle school students: An exploratory study. *International Journal of Artificial Intelligence in Education*, 33, 290–324. <https://doi.org/10.1007/s40593-022-00293-3>
- Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research and future directions. *Computers and Education: Artificial Intelligence*, 2, 100025. <https://doi.org/10.1016/j.caeai.2021.100025>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.
- Zumeta, L., Oriol, X., Telletxea, S., Amutio, A., & Basabe, N. (2016). Collective efficacy in sports and physical activities: Perceived emotional synchrony and shared flow. *Frontiers in Psychology*, 6, Article 1960, 1–12. <https://doi.org/10.3389/fpsyg.2015.01960>

APPENDICES

APPENDIX A

PUBLISHED MANUSCRIPT IN INTERNATIONAL JOURNAL OF ARTIFICIAL INTELLIGENCE IN EDUCATION

Mouta, A., Pinto-Llorente, A.M. & Torrecilla-Sánchez, E.M. (2024). Uncovering Blind Spots in Education Ethics: Insights from a Systematic Literature Review on Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-023-00384-9>

Journal Quality Indicators:

JCR – Q1. JIF (2024): 8.5. JIF Rank: 15/175. Percentile: 91.7%

SJR – Q1. SJR (2024): 1.96. H-Index: 68



Uncovering Blind Spots in Education Ethics: Insights from a Systematic Literature Review on Artificial Intelligence in Education

Ana Mouta¹ · Ana María Pinto-Llorente² · Eva María Torrecilla-Sánchez²

Accepted: 20 November 2023
© The Author(s) 2023

Abstract

In the last decade, research on the use of artificial intelligence technologies in education has steadily grown. Many studies have demonstrated the potential of these technologies to improve school administration processes, enhance students' learning experiences, simplify teachers' daily tasks, and broaden opportunities for life-long learning. However, the enthusiasm surrounding these possibilities may overshadow the ethical challenges posed by these systems. This systematic literature review is designed to explore the ethical dimensions surrounding the utilisation of these technologies within the defined timeframe (2011–2022) in the field of education. It undertakes a thorough analysis of various applications and objectives, with a particular focus on pinpointing any inherent shortcomings within the existing body of literature. The paper discusses how cultural differences, inclusion, and emotions have been addressed in this context. Finally, it explores the capacity building efforts that have been put in place, their main targets, as well as guidelines and frameworks available for the ethical use of these systems. This review sheds light on the research's blind spots and provides insights to help rethink education ethics in the age of AI. Additionally, the paper explores implications for teacher training, as educators play a critical role in ensuring the ethical use of AI in education. This review aims to stimulate ethical debates around artificial intelligence that recognise it as a non-neutral tool, and to view it as an opportunity to strengthen the debates on the ethics of education itself.

✉ Ana Mouta
mouta.ana@usal.es

Ana María Pinto-Llorente
ampintoll@usal.es

Eva María Torrecilla-Sánchez
emt@usal.es

¹ PhD Programme in Education in the Knowledge Society, University of Salamanca, Salamanca, Spain

² Instituto Universitario de Ciencias de La Educación (IUCE), University of Salamanca, Salamanca, Spain

Keywords Education · Artificial intelligence · Ethics · Literature review

Introduction

In the late 1950s, artificial intelligence (AI) emerged as both a promise and a threat based on the expectation of having a powerful agent that would facilitate everyday tasks while exercising its decision-making power over people. As an autonomous, adaptive, and interactive human-designed software system, AI today can make decisions in complex situations relying on perception, interpretation, and reasoning based on data (Dignum, 2021). AI in education (AIED), in particular, is being used in administration, to track the goals of the school in conjunction with policies, to pursue student interests, to assist teachers in daily tasks, and to facilitate lifelong learning (Miao et al., 2021). Promises surrounding these systems include identifying and enhancing competencies and talents, freeing up teachers from workload, addressing student diversity, predicting underachievement in students and institutions, facilitating the transition to professional life, enabling access to cheaper and better-quality education for poorer people, and more engaging learning experiences. With the COVID-19 pandemic, debates about the role of these resources in facilitating online assessment and student learning experience have increased (García-Peñalvo et al., 2021). For example, a study conducted in Romania showed that due to the educational needs caused by the pandemic, the use of AI-driven platforms increased between 2019 and 2020 among both teachers and students, even in less developed geographic areas (Pantelimon et al., 2021). Discussions on whether AI can offer a fair learning experience in health emergencies are going on. The suitability of AIED for cases of forced migration resulting from environmental or social crises may also emerge as a concern in the near future, underscoring the pressing need for discussions on AIED ethics. So, what are the trade-offs between all the opportunities and solutions these technologies can offer to some and the challenges posed by its design, implementation practices and non-universality in terms of access? And what type of changes can we expect in education and in the role of its actors, when introducing these technologies into pedagogical practice?

Lurking within all these advances are ideologies, fantasies, and projections about how the future should be or is likely to be (Mouta et al., 2020). As early as 1960, Weiner (1960) observed that our ability to keep up with and understand what is happening decreases as machines become more powerful and technologies become smarter. Therefore, thinking about AI is expected to consider compromises between ethics and the movements of transhumanism in order to understand how AI participates in this pursuit of perfection that Western philosophy sees as human's innate nature (Byk, 2021). And once perfectibility is considered the ultimate social and human goal of education, are intelligent machines given a free pass in the process? Ethical reasoning seems to be able to offer such considerations and possible transitions a sustainable conceptual support, especially in education where the "when, how, and to what end" of the pedagogical use of AI should always be asked (Holmes et al., 2021).

When thinking about AIED, it is important to understand what concept of intelligence lives on and has prevailed in these AI advances. Is it inclusive enough? Is it ethically considered to suit learning diversity and respect emotional expression? And why is intelligence the artificial layer we are promoting and what message does it bring to education? Cave (2020) argues that intelligence is value-laden, with links to colonialism, racism and patriarchy. Another reason for concern about the rapid advances in AI has to do with the complexity that has been added to discussions about the digital divide: now the gap is not just between users, but between those generating the data – users who supposedly should own it – and tech corporations who process it – those that objectively own it (Abboud et al., 2020). Disparities are no more just about access, but about ownership and rights, and the abilities to deliver, process, and access particular data. According to Miao et al. (2021), this divide encompasses developed and developing countries, socio-economic groups within countries, owners, and users of technologies, and those having jobs enhanced by AI against those who may be replaced by it.

An analysis of the role of education in AI policies conducted by Schiff (2021) found that global policy does not address the ethics of AIED and that it mainly focuses on education as an instrumental strategy to guarantee the supply of AI talent to ensure manpower. There is also a shortage of education and training opportunities for teachers, parents, and the general public on this matter (Miao et al., 2021). Moreover, a systematic review approach of AIED from 1970 to 2020 presents the main research topics of the field: 1) adaptive learning and personalisation, 2) deep learning and machine learning algorithms used online, 3) educational human-AI interaction, 4) AI-generated data educational use, and 5) AI in higher education. Ethics on AIED seemed to be an almost absent field of research (Bozkurt et al., 2021). Having AI systems detecting students' emotions could support emotional preparedness for the educational process. However, this may bring concerns related to emotional privacy, the elicitation of emotions, and virtual connections between a person and a digital assistant (Hudlicka, 2016). Remembering children's rights is a call to reflect on educational ethics through a cross-cultural and mutual perspective in a globalised world (Nizhnikov, 2018). Thus, AIED studies are expected to discuss the decisions made regarding pedagogy, purposes of learning, the role of technology in relation to teachers, and access to education (Holmes et al., 2021). According to Weber (2020), AIED ethics should take into account: 1) legal frameworks for the use of AIED in learning institutions; 2) cloud-based data clear terms for privacy, security and trust; 3) power asymmetries and misuse by malicious actors; 4) equity and social justice; 5) machine liability and accountability; and 6) educational programmes for students on technology ethical dimensions.

The aim of these reflections is to prioritise contextual pedagogical factors in any potential reconfiguration of learning practices with the emergence of AIED. How can ethical frameworks for the use of these systems incorporate dialogic ethics that reflect the principles of participation from diverse educational communities? Is it possible to incorporate these technologies in a manner that not only supports personal ethical development but also fosters an environment in which educational stakeholders can embrace and respect the diverse interests and entities within the educational landscape? This encompasses not only various educational systems but

also extends to the different stakeholders such as governance bodies, the technology industry, and policy-makers. Can the integration of these technologies create conditions that cater to the needs and values of diverse educational communities, balancing interests, while promoting ethical growth? To explore these issues, the following lines will examine published literature on AIED and ethics from 2011 to 2022, considering conceptual assumptions and the impact assessment of programmes that use AI technologies.

Method

Following the discipline of a systematic literature review (SLR), this research screens the most relevant data from the existing literature, responding to pre-defined eligibility criteria (Ramírez-Montoya & García-Peñalvo, 2018). The next lines acknowledge the nature of the debates on AIED ethics and how ethics is screened in AIED studies.

Planning

As shown in Table 1, the planning phase began with four mapping questions (MQs) and a set of four inclusion and exclusion criteria to globally screen what has been published in AIED and ethics over the last 10 years. To ensure a comprehensive review considering the global scope of this study, the researchers' language proficiency was taken into account and multiple languages were included. Due to the time frame for completing the review, it was not possible to add more than four languages, as translation would depend on resources not immediately available. To allow for a more in-depth analysis, five research questions (RQs) were added. These research questions are intricately connected to the themes and concerns raised in the introduction. They provide the necessary direction and focus to fulfill the paper's primary motivation, which is a thorough examination of the shortcomings in the literature, limited to the chosen timeframe, regarding the ethical aspects of AIED and their impact on educational practices. These questions offer a structured approach to identifying current applications and the challenges they may bring (RQ1), disparities in terms of data ownership, access, implementation quality and contextualisation (RQ2), emotion surveillance (RQ3), focus and targets of capacity building for an ethical use of AIED (RQ4) and available regulations to safeguard the compliance of ethical principles in AIED design and implementation (RQ5). As stated in the introduction, the knowledge obtained from these RQs would be beneficial for promoting ethical responsibility through dialogic practice, as it addresses several unresolved aspects in AIED use, including the balance between access and quality (of implementation), the tension between universality and context, the interplay of cognition and emotions, and the normative guidance provided by targeted training and regulations.

The PICOC method (Petticrew & Roberts, 2008) was used to accurately structure this review. The deliberate use of broader keywords in this SLR was

Table 1 Systematic review: Planning Phase

Goal	To understand how ethics has been covered in AIED studies
PICOC	<ul style="list-style-type: none"> • Population (P): AIED and Ethics studies • Intervention (I): Conceptual models and projects on ethics for using AIED • Comparison (C): No comparison • Outcomes (O): Frameworks and projects/programmes of AIED and ethics • Context(C): k-12 Education, Higher Education, LLL
Mapping Questions	<p>MQ1. What has been the evolution of the number of research articles concerning ethics in AIED since 2011?</p> <p>MQ2. Who are the key authors in AIED ethics?</p> <p>MQ3. Which countries exhibit the most significant productivity in terms of disseminating conceptual or theoretical approaches, frameworks, or interventions in the realm of AIED ethics?</p> <p>MQ4. From which subject areas are coming the main studies in AIED ethics?</p>
Research Questions	<p>RQ1. Which AIED technologies, targets and applications are being considered and what challenges do they pose in general?</p> <p>RQ2. How is AIED considering cultural differences and inclusion?</p> <p>RQ3. How does AIED monitor emotions and what ethical challenges may be at stake?</p> <p>RQ4. How is capacity building on AIED ethics being covered?</p> <p>RQ5. What principles, regulations and frameworks are there for AIED?</p>
Sources & String search	<p>SCOPUS</p> <p>(TITLE-ABS-KEY (artificial AND intelligence) AND TITLE-ABS-KEY (education) OR TITLE-ABS-KEY (learning) OR TITLE-ABS-KEY (teaching) AND TITLE-ABS-KEY (ethics)) AND PUBYEAR > 2010 AND PUBYEAR < 2022 AND (LIMIT-TO (LANGUAGE, "English") OR LIMIT-TO (LANGUAGE, "French") OR LIMIT-TO (LANGUAGE, "Portuguese") OR LIMIT-TO (LANGUAGE, "Spanish"))</p> <p>WoS</p> <p>artificial intelligence in education (All Fields) or artificial intelligence in learning (All Fields) or artificial intelligence in teaching (All Fields) and ethics (All Fields) and English OR French OR Portuguese OR Spanish (Language)</p>
Inclusion Criteria	<p>IC1. Publication about AIED and ethics</p> <p>IC2. Data range 2011–2022</p> <p>IC3. English, French, Portuguese, Spanish</p> <p>IC4. Peer-reviewed</p>
Exclusion Criteria	<p>EC1. Publication about AIED or ethics but not both</p> <p>EC2. Data range before 2011 or over 2022</p> <p>EC3. Other language than English, French, Portuguese, Spanish</p> <p>EC4. Not peer-reviewed</p>

necessary because including research question terms directly in the search strings limited the results. Broader keywords were chosen to encompass the terminology variations used by authors when addressing the same research questions. This approach allowed for a comprehensive exploration of the AIED ethics field and the flexibility to fine-tune the search based on the results obtained. Considering the comprehensive focus of the research, Scopus and Web of Science were the selected databases as they correspond to all-inclusive digital libraries that cover a wide range of academic disciplines within the scope of this study; they were

also selected for having logical expressions and both full-length and field-specific articles.

Conducting and Analysing

To manage the final cluster of articles, the results were gathered into a spreadsheet (cf. <https://docs.google.com/spreadsheets/d/1u6ArXbQ5w4bxhHbKAcRUCZLUSACK0AoG/edit?pli=1#gid=1005218435>), which presents the analysis in several stages (cf. Figure 1). In the identification phase, 250 publications were retrieved from Scopus and 249 from WoS. The duplicate papers were eliminated, and a selection of 410 papers was made based on their title and abstract. For the preliminary screening phase, the alignment of the paper's topic with the research goal, the PICOC elements (population, intervention, comparison, outcomes, and context), the paper's implicit (e.g., using terms such as "AIED challenges", "FATE") or explicit relation to AIED ethics, and the fulfilment of the inclusion/exclusion criteria (cf. Table 1) were taken into account. If in doubt, the entire article was read. This preliminary screening phase (abstracts) enabled the gathering of 156 contributions. After this stage, a quality assessment screening was conducted for several purposes, including ensuring relevance (i.e., that the selected papers directly align with the research goals, objectives, and topic), assessing methodological rigor (i.e., whether the selected papers meet specific quality criteria), evaluating reliability (using a predefined scoring system to measure the assessment's reliability), ensuring a high-quality and relevant sample selection, and verifying data integrity (i.e., that the

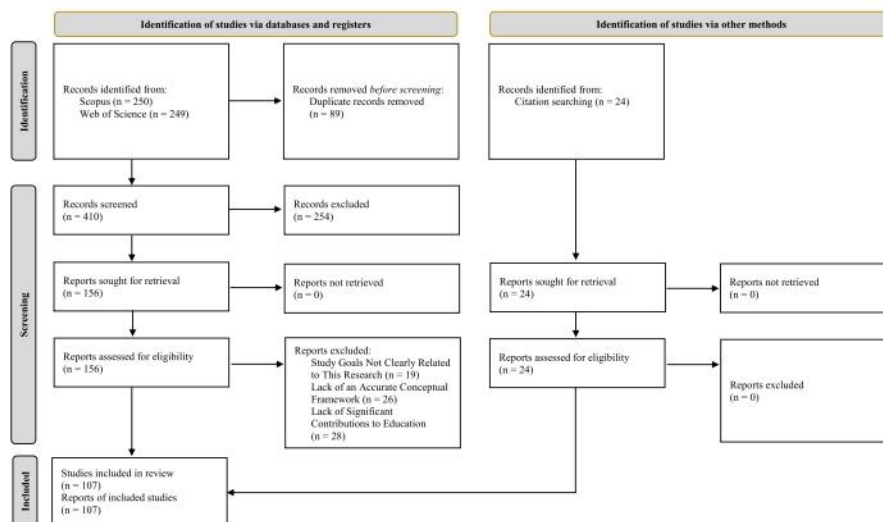


Fig. 1 PRISMA 2020 flow diagram for the SLR process: identification of AIED ethics papers. Note: Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 2021;372:n71. <https://doi.org/10.1136/bmj.n71>. For more information, visit: <http://www.prisma-statement.org/>

selected papers are reliable and meet established quality standards). In the eligibility stage, the retrieved articles were thoroughly read and analysed according to the five-criteria quality assessment checklist presented in Table 2. Three options were available for each criterion: Yes (1 point), Partially (0.5 point), and No (0 point). Five points were awarded to items that fully met the defined quality parameters. All papers with four or more points were considered; since four is the cut-off point, all papers with a lower score were excluded from the final sample and from the corresponding RQs. This quality assessment phase allowed the collection of 83 papers. Finally, 24 articles were added to this cluster: they were first or second references, chosen whenever they engaged in relevant work surrounding the RQs, particularly on topics requiring a deeper understanding (e.g., AIED and cultural differences) or further examples (e.g., AIED technologies in use). The reasons for choosing these papers are presented in Table 3. As a result of this process, 107 papers make up this sample.

Mapping Analysis

MQ1. What has been the evolution of the number of research articles concerning ethics in AIED since 2011?

As Fig. 2 shows, between 2011 and 2018, the number of publications on this subject exhibited some fluctuations. This period can be seen as a foundational phase, where discussions around the ethical dimensions of AI in education began to take shape. However, starting in 2019, there was a noticeable shift in the trajectory. AIED research began to experience exponential growth, although there was a small retraction in 2022, with a remarkable surge in the number of articles in 2021. This explosion in publications can be attributed to several factors. Firstly, AI technologies continued to advance rapidly, permeating various educational settings. As these technologies gained traction, there was a corresponding surge of interest in their ethical implications. Additionally, the global adoption of AI in education became more widespread, with researchers and educators worldwide focusing on the ethical aspects that should accompany its integration. This broader reach led to a surge in research outputs, as experts from various geographies contributed to the discourse.

Table 2 AIED and Ethics papers' quality assessment questions

Quality Assessment Questions	Score Criteria
1. Is it a theoretical/conceptual approach or research/study in AIED ethics?	Yes/No/Partial
2. Do the study's goals clearly relate to AIED ethics?	Yes/No/Partial
3. Does the theoretical/conceptual analysis or research meet those goals?	Yes/No/Partial
4. Are the relations between concepts, data, interpretation, and conclusions made clear?	Yes/No/Partial
5. Does the paper explicitly refer its contribution to think about/question the future of education?	Yes/No/Partial

Table 3 Reasons for the inclusion of additional papers

Papers Nos	Reasons for papers inclusion	Explanation
1, 2, 4, 7, 8, 10, 12, 16, 18, 19, 20, 21	Directly related to ethics and AI	These papers, directly related to ethics and AI, serve as the core material for examining ethical considerations within the field of education
3, 6, 11, 24	Relevance for RQ1	These papers offer insights into the technologies, targets, and applications in use, as well as the challenges they may pose
13, 15, 23	Relevance for RQ2	These papers offer insights into AIED, cultural differences and inclusion
9, 14, 17	Relevance for RQ3	These papers offer insights into AIED and emotions monitoring
5	Related to ethics and AI + author specialist in this field	This paper combines the connection to ethics and AI with the expertise of the author
22	Indirectly related to ethics and AI (children moral development)	This paper is indirectly related to ethics and AI, shedding light on children's moral development within the context of AI

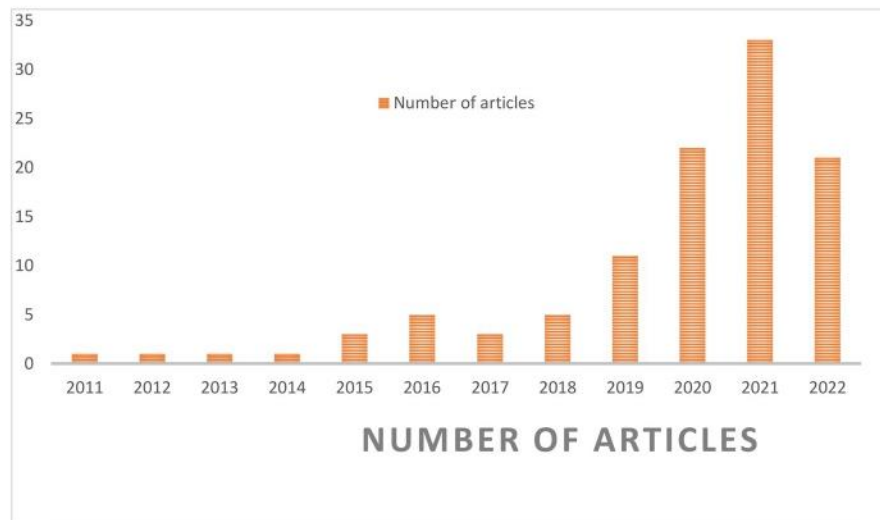


Fig. 2 Number of articles on AIED ethics between 2011–2022

Moreover, the spotlight on AI ethics in mainstream media and entertainment further fuelled academic research and discourse in the field of AIED.

MQ2. Who are the key authors in AIED ethics?

Identifying key contributors allows to recognise those who play a significant role in shaping the discourse within this domain. It provides valuable context to comprehend the methodologies employed in these studies and the perspectives from which they were elaborated. Furthermore, it opens the door to assisting researchers in tracking experts in this specific field of study. In this SLR, Holmes, W. has contributed the most publications with 4, followed by Tuomi, I. with 3. Additionally, there are some authors who have two publications each, including Cukurova, M., Dignum, V., Luckin, R., Mouta, A., Pinto Llorente, A. M., Shum, S. B., and Torrecilla Sánchez, E. M..

MQ3. Which countries exhibit the most significant productivity in terms of disseminating conceptual or theoretical approaches, frameworks, or interventions in the realm of AIED ethics?

The inclusion of this mapping question aligns with the need emphasised in the introduction to delve into how cultural differences are addressed within the context of AIED ethics. Figure 3 displays that 47% of the publications originate from the United Kingdom, with Finland contributing 20%. Spain and Sweden account for 13% of the papers collectively, while Australia contributes 7%. If the majority of papers in AIED during the specified research period can be attributed to the USA, it's noteworthy that, when the ethical dimension is introduced, a European

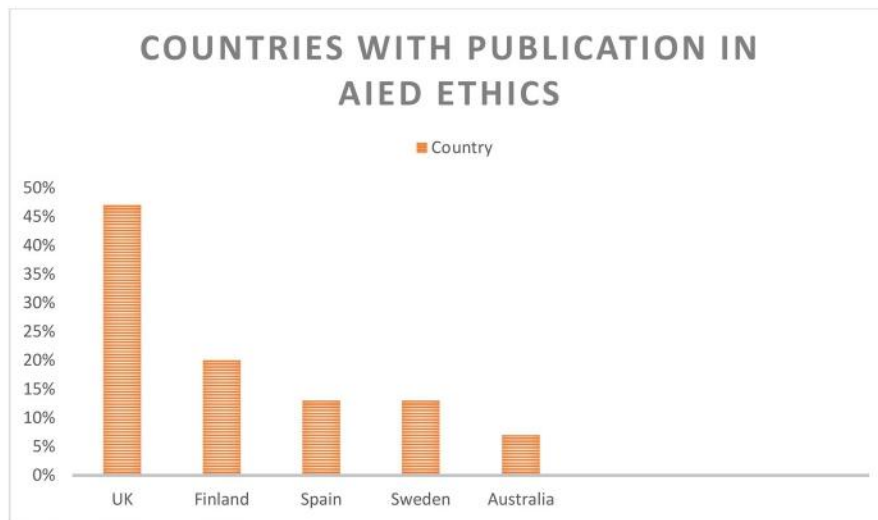


Fig. 3 Countries with publication in AIED ethics

discourse predominates. This can be understood due to Europe's long history of philosophical and ethical inquiry, ranging from the works of Enlightenment thinkers to contemporary ethical philosophy. These traditions often emphasise concepts such as human rights, individual autonomy, and the ethical responsibilities of technology developers and educators. These values align with the European Union's General Data Protection Regulation (GDPR), which places a strong emphasis on data privacy and ethics.

The concentration of research in AIED ethics within Europe suggests the presence of regional perspectives and priorities in addressing ethical considerations in educational technology. This also underscores the necessity for a global dialogue that encompasses diverse cultural viewpoints, as the ethical implications of AIED extend beyond geographical boundaries.

MQ4. From which subject areas are coming the main studies in AIED ethics?

The analysis of subject areas relied on the International Standard Classification of Education (ISCED), developed by UNESCO in 2012. This classification system encompasses 25 fields of education organised into 9 broad groups. The main areas that ethical discussions surrounding the use of AIED relate to are shown in Fig. 4: Science (e.g., computing, life sciences, mathematics) takes the lead, with over 50%. It is followed by Education where 32% of articles may be found. Engineering, Manufacturing, and Construction (e.g., architecture, engineering, manufacturing) follow with 5%. Health and Welfare (e.g., medicine, nursing, social work) and Humanities and Arts (e.g., languages, literature, history, philosophy, performing and visual arts, cultural studies) each cover 4% of the total number of articles. Finally, Social Sciences, Business and Law (e.g., economics, journalism,

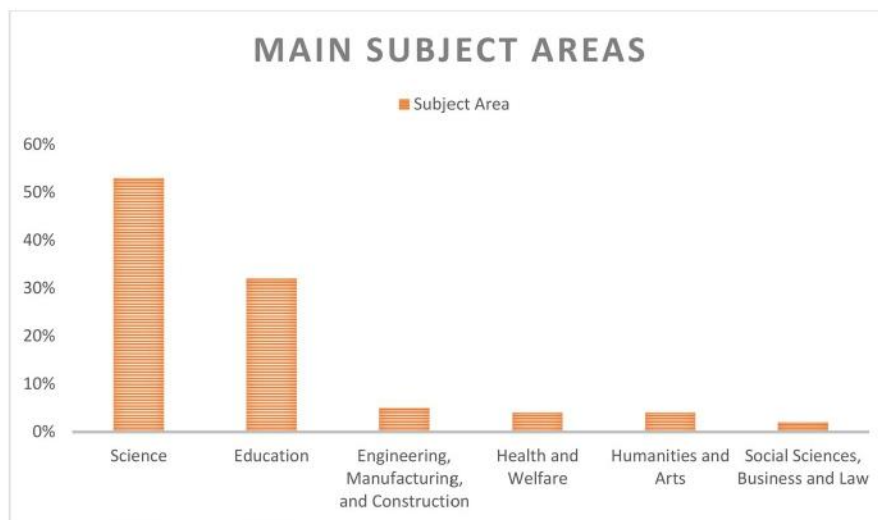


Fig. 4 Publications on AIED ethics main subject areas

political science, psychology, sociology) collectively have a total of 2% of the articles.

The diverse distribution of ethical discussions across academic domains in AIED underscores the need for a comprehensive ethical framework that addresses the unique challenges and concerns within each field. The dominant discourse in science, particularly computing, comprising over 50% of ethical discussions, highlights the central role of technology and data-driven approaches in AIED. While science, notably computing, dominates the discourse, it's crucial to broaden the conversation to ensure that other areas are not overlooked. The scarcity of ethical discussions in humanities and arts suggests a potential gap in recognising the human-centred dimension of learning and pedagogy.

SLR Research Analysis

The results obtained after analysing and studying the articles supporting this SLR are presented below according to the research questions.

RQ1. Which AIED technologies, targets and applications are being considered and what challenges do they pose in general?

AI technologies and the overall ethical challenges they pose have been the second most studied topic of research on the ethics of AIED between 2011 and 2022. Table 4 summarises 36% of the studies in this SLR, presenting AIED technologies, their targets, the primary educational purposes, and potential ethical challenges. In terms of targets, AIED technologies are being applied to a wide spectrum, including

Table 4 AIED technologies, applications, targets and the ethical challenges they pose

Studies	AIED Technologies	Targets	Purpose	Ethical Challenges
Bucea-Manea-Tomis et al., 2022; Chountia et al., 2022; Gal-lastegui et al., 2021; Holmes & Tuomi, 2022; Kitto & Knight, 2019; Latham & Goltz, 2019; Leonidis et al., 2012; Li et al., 2021; Miao et al., 2021; Pammer-Schindler & Rosé, 2022; Prinsloo & Slade, 2017; Qin et al., 2020; Serholt et al., 2017; Shum, 2018; Tuomi, 2018	Ambient classroom tools, linking the physical & digital worlds (pervasive & multimedia computing, sensor networks, robotics, agent-based software)	Schools (K-12) & Universities	Fees, attendance, retention and dropouts, school inspections, teachers' performance; monitor students' behaviour; provide related data to the classroom's services and applications	Utilitarianism & Deontology: data collected not only related with the learners, but with family; individuals' conflicting claims; interests of private organisations vs. public authorities, schools, universities; AIED instructional implications



Table 4 (continued)

Studies	AIED Technologies	Targets	Purpose	Ethical Challenges
Aboud et al., 2020; Adams et al., 2022; Belpaeme et al., 2018; Burr et al., 2020; Córdova & Vicari, 2022; Dobrosvetnova & Hannibal, 2020; Ghotbi & Ho, 2021; Gudíño Paredes et al., 2021; Holmes & Tuomi, 2022; Hood et al., 2015; Lamerás & Arnab, 2022; Li et al., 2021; Miao et al., 2021; Mohammed & Watson, 2019; Mouta et al., 2019; Pammer-Schindler & Rosé, 2022; Papa & Jackson, 2021; Renz & Vladova, 2021; Serholt et al., 2017; Smakman et al., 2021; Tundrea, 2020; Ungerer & Slade, 2022; Williams et al., 2018; Zhang & Aslan, 2021	ITS, AI-supported reading, smart & social robots, teachable agents, virtual and AR, AI-enabled collaborative learning, emotional facial recognition & regulation, sensing technology (gaze tracking)	Students (K-12, University, Master's and Ph.D.)	Personalised learning; meta-cognition & self-esteem improvement; learning disabilities support; improvement of writing before assessment; promotion of STEM subjects and computer programming; virtual school (impossibility to go); tutors (any part of the world) & peers (collaborative learning)	Human rights & Justice: surveillance causing anxiety; reduction of human interactions; human replacement; students' privacy (activity logs screen political views or sexual orientation), safety, trust, and fairness; sense of agency; self-efficacy; student freedom of choice (e.g., AI career guidance); children's moral development (e.g., social robots, talking dolls); posthumanist ethics of AIEDK-12
Adams et al., 2022; Belpaeme et al., 2018; Coghlan et al., 2021; Ghotbi & Ho, 2021; Holmes & Tuomi, 2022; Leonidis et al., 2012; Li et al., 2021; Miao et al., 2021; Mohammed & Watson, 2019; Tuomi, 2018	Adaptive assessment, automated writing evaluation, AI-driven discussion forum, AI-teaching assistants, "dual teacher"	Teachers	Assessing; helping learners stick to the curriculum; expertise feedback; asynchronous discussion; personalised content & activity; bureaucratic tasks	Justice & Human rights: data extrapolation; no creativity assessment; quantity preferred to relevance; inability to prevent "deep fake" assignments; incorrect correlations deriving from bias, posthumanist ethics of AIEDK-12

Table 4 (continued)

Studies	AIED Technologies	Targets	Purpose	Ethical Challenges
Holmes & Tuomi, 2022; Marras et al., 2022; Pammer-Schindler & Rosé, 2022; Pham & Wang, 2017; Westman et al., 2021	Multimodal mobile learning system for MOOCs	Lifelong Learning	Real time feedback; career guidance; speculative AI-assisted LLL assistant	Justice; learning opportunities imbalance among learners; inaccurate data; data ownership; data privacy and consent

school institutions (K-12) and universities, as well as individual students and teachers. This highlights the broad scope of AIED's influence within the educational ecosystem. AIED applications are designed to serve a multitude of educational purposes. The claims include personalised learning, meta-cognition improvement, support for students with learning disabilities, enhancement of writing skills, promotion of STEM subjects and computer programming, virtual schooling, and collaborative learning experiences. These applications are tailored to meet diverse educational needs and objectives.

The risks associated with the use of such technologies are frequently overlooked in the literature, revealing a substantial gap in understanding how automated measurement processes could potentially compromise well-being (Burr et al., 2020). For instance, while profiling higher education students may seem promising in terms of career advancement, it can also be viewed as manipulative, as it fails to acknowledge individuals' freedom of expression and choice (Tundrea, 2020). Sensitive information can also be inferred from seemingly innocuous data, as activity logs can be leveraged to deduce someone's political beliefs, ethnic identity, health status, or sexual orientation (Tundrea, 2020). Some authors draw attention to the potential conflict between the obligation of higher education institutions to fulfil their mandate of ensuring learning and their fiduciary duty to act ethically (Prinsloo & Slade, 2017).

While doubts persist among neuroscientists regarding the potential of certain technologies to improve learning, the risks of inaccurate results or unintended consequences linked to electroencephalography (EEG) remain significant. Despite these concerns, EEG sensors were already incorporated into headbands for detecting students' brain activity (Miao et al., 2021). Therefore, AI has faced criticism for being "dehumanising" in this context, as it often promotes prescriptive teaching with minimal interaction and automated pathways, ultimately diminishing students' agency (Miao et al., 2021).

Another worrisome issue pertains to children aged four to ten who perceive social robots as trustworthy. For example, a talking doll might influence children to reconsider their moral judgments (Williams et al., 2018). This raises further concerns about whether these robots are primarily profit-driven. Additionally, questions regarding privacy emerge as AI delves into children's socio-emotional and cognitive characteristics, as well as aspects of their home environment. Furthermore, transparency issues arise regarding who has access to this data (Mohammed & Watson, 2019).

The significance of robust assessment tools has been underscored by the COVID-19 pandemic. Gudiño Paredes et al. (2021) conducted a mixed-methods study to examine the impact of remote AI-powered proctored exams on the learning processes and academic integrity of online graduate students. This technology offers facial recognition, audio analysis, eye movement tracking, and object/face detection in the surroundings. While the results indicated a substantial reduction in dishonesty, students reported feeling compelled to cheat under observation, lacking internal motivation or a personal reflective process. They also expressed concerns about privacy and anxiety during assessments conducted in this manner.

Regarding teachers, the possibility of letting these AI technologies automate ineffective pedagogical practices (based in data incompleteness and bias, for example) is

also a reason for concern. In the long run, AI-based education may have the potential to disempower teachers (Miao et al., 2021). Another concern revolves around the development of learners' cyborg identities as they engage with AI, the changing relationship between humans and AI systems in a posthuman hybrid dynamic and how it may impact teachers and students (Adams et al., 2022). Finally, most studies assessing the effects of technology are typically carried out by the creators of that technology, often affiliated with commercial entities and involving a limited number of participants (Holmes & Tuomi, 2022).

These insights shed light on the transformative potential of AI in education, recognising how important it is to consider the ethical dimensions that come into play when implementing these technologies.

RQ2. How is AIED considering cultural differences and inclusion?

Although “the greatest good for the greatest number” is a concern when introducing these technologies, what can be considered good in an educational context is a challenging question. Equity, cultural and interpersonal differences, which are the leitmotiv of 17% of the papers under analysis, are an issue when it comes to massive technologies that may not easily adapt and respond to the specifics of the context and the people who are meant to use them, potentially undermining the fairness and fundamental rights of individuals. Table 5 demonstrates that out of the 18 articles included in the SLR, there are four distinct categories of ethical challenges pertaining to cultural and inclusion responsiveness. The “intercultural” challenge has to do with these systems' capacity to accommodate different cultural background and learners' values. Approximately 53% of the papers within this subject area highlight that AI solutions for educational purposes may not sufficiently consider cross-cultural variations. Another challenge has to do with “cultural realism” and the difficulty of representing the diversity of learners' particular physical characteristics. Some authors defend that AIED is best suited to Western, educated, industrialised, wealthy, and democratic nations (Nye, 2015; Ogan et al., 2015). For example, limited broadband access will leave people out of data sets and AI will be unintentionally biased against them (Miao et al., 2021). There's also the issue of reported misidentifications from facial recognition software in relation to darker skin tones. This imperatively requires the choice of suitable workaround solutions or the omission of facial recognition (Coghlan et al., 2021). Sanusi and Olaleye (2022) discuss the role of cultural competence and ethics in AI education, and how these factors influence students' learning of AI. It also touches upon the disparities between rural and urban students in terms of cultural and ethical competence. Holmes and Tuomi (2022) discuss AIED colonialism, which includes the adoption of single products in state education systems, language biases, and the imposition of specific pedagogical approaches. This colonialism varies in extent but often results in well-funded Global North AIED tools overshadowing locally sensitive alternatives.

The “inclusion” challenge refers to the effort of guaranteeing that AIED features recognise learners' neurological, physiological and psychological diversity and become accessible and accurate to all. Although it's a challenging and early-stage process, robots and empathic intelligent learning environments (ILEs) are already

Table 5 Cultural and inclusiveness ethical challenges of AIED and corresponding percentage of papers per topic

Studies	Ethical Challenges	Percentage of papers
Coghlan et al., 2021; Dieterle et al., 2022; Holmes & Tuomi, 2022; Miao et al., 2021; Nye, 2015; Ogan et al., 2015; Pinkwart, 2016; Roll & Wylie, 2016; Samusi and Olaleye (2022); Timms, 2016	Cross-cultural differences not being respected and tailored by AI	53%
Mohammed & Watson, 2019	Cultural realism not being incorporated into AI design and development	5%
Coghlan et al., 2021; Costa et al., 2018; Epp & Makos, 2013; Johnson & Lester, 2016; Pham & Wang, 2017; Radford et al., 2021; Rello et al., 2016	Inclusiveness not being taken into account (e.g., Special Educational Needs and Disabilities)	37%
Restrepo et al., 2019	Responsiveness to victims of violence in schools not being considered	5%

being designed to recognise emotional patterns and respond accordingly (Pham & Wang, 2017). This is especially beneficial for children with special needs. For instance, socially assistive robots proved more effective in reducing repetitive and stereotyped behaviours in autistic children compared to interactions with people (Costa et al., 2018). In the case of autism spectrum disorder, educational games are being used to enhance children's ability to distinguish emotions in a simulated learning environment, aiming to facilitate their transition between the virtual and real-world contexts (Epp & Makos, 2013). Another study showed that children engaging in a 30-min daily interaction with a caregiver and a social robot over one month improved their attention skills when the robot was not present (Johnson & Lester, 2016). The robot encouraged emotional storytelling, perspective-taking, and tailored the difficulty of activities based on past performance. These studies demonstrate that empathic robots can create more engaging and fearless learning experiences, especially for K-12 students (Johnson & Lester, 2016). While these advancements offer promising perspectives for inclusion in special education, challenges arise in the context of online proctoring. The technology may not be adequately prepared to interpret the behaviours of neurodiverse individuals or those with disabilities, potentially leading to false positives for cheating (Coghlan et al., 2021).

Finally, these systems may not be able to address issues of harassment, bullying, or discrimination that may occur in online educational environments and ensuring that AIED technologies have mechanisms in place to prevent or respond to such incidents. The European project ACACIA (Restrepo et al., 2019), funded in part by Erasmus+, features a chatbot named Artemisa designed to address sexual harassment and recruit volunteers to promote diversity and tolerance at the Peruvian National University of San Marcos. However, like other chatbots, it presents accessibility challenges for users, underscoring the need for training in accessibility, tolerance, and diversity acceptance to mitigate biases.

Despite the various difficulties mentioned earlier, there is a growing recognition of the importance of culturally inclusive research in the field of AIED (Nye, 2015). The data presented in Table 6 showcases a diversity of studies that underscore the need to develop AI systems that are socially and culturally aware. Some of the examples, include enculturated agents which can adapt their interactions and responses to align with the cultural background and preferences of the users,

Table 6 Cultural and inclusive features of AIED

Studies	Focus
Mohammed & Watson, 2019	Cultural roots of teaching and learning factored into ILEs
Mohammed & Watson, 2019	Enculturated conversational agents
Epp & Makos, 2013; Johnson & Lester, 2016; Pham & Wang, 2017; Radford et al., 2021; Rello et al., 2016	Empathic robots and ILEs for Special Education
Restrepo et al., 2019	Chatbots for victims of sexual harassment (e.g., European project ACACIA)

enhancing the effectiveness of communication and learning (Mohammed & Watson, 2019). Empathic robots and ILEs can also be tailored to meet the needs of students with special educational needs. These could include individuals with conditions such as autism, hearing or oral communication problems, and dyslexia (Epp & Makos, 2013; Johnson & Lester, 2016; Pham & Wang, 2017; Radford et al., 2021; Rello et al., 2016). Chatbots can also be used as a means of providing support to victims of sexual harassment, ensuring anonymity, accessibility, and responsiveness that can be particularly valuable in sensitive and emotionally challenging situations (Restrepo et al., 2019).

RQ3. How does AIED monitor emotions and what ethical challenges may be at stake?

Since AIED is primarily justified as a response to learner holistic needs, this section aims to understand how the affective dimension is addressed alongside the cognitive and performance dimensions. Regarding ethical considerations, only 4% of all papers analysed in this SLR address the particular ethical concerns that may arise from AIED's handling of emotions. The main concerns include negative feelings related to surveillance, direct correlations between behaviour and emotions, a lack of respect for affective privacy, and the induction of emotions (cf. Table 7). The biggest percentage of papers related to this topic, concentrate on the emotional outcomes of affective surveillance. Automated monitoring designed to enhance productivity and well-being has been associated with increased levels of stress and anxiety (Burr et al., 2020) and may result in nervousness during evaluation (Gudiño Paredes et al., 2021). Furthermore, despite advancements in AIED, there seems to be a missing context for understanding emotions and their meaning beyond the "signalling paradigm" of matching emotions with their corresponding behavioural signals (Dobrosovestnova & Hannibal, 2020). Furthermore, positive emotions, motivation, academic performance, and school achievements drive the integration of AI into formal education. Empirical studies suggest that the affective component in

Table 7 Ethical Challenges of AIED Emotion Control and corresponding percentage of papers per topic

Studies	Ethical Challenges	Percentage of papers
Burr et al., 2020; Gudiño Paredes et al., 2021	Human rights: emotional consequences of affective monitoring	40%
Dobrosovestnova & Hannibal, 2020	Utilitarianism and Justice: diversity of emotions not being recognised by AI	20%
Hudlicka, 2016	Human rights: affective privacy (the right to keep one's feelings to oneself)	20%
Hudlicka, 2016	Human rights: emotion induction (the deliberate process of changing someone's feelings)	20%

artificial agents enhances learning compared to experiences without emotional and social aspects. However, there is an excessive optimism surrounding these automated systems. While learners and teachers experience a range of emotions, from joy to frustration, the design of affective behaviour in educational robots has largely concentrated on conveying positive emotions. Given that complex and non-positive emotions are also relevant in designing social robots for educational purposes, it is critical to accurately model these ambivalent emotional traits (Dobrosrovestnova & Hannibal, 2020).

According to Hudlicka (2016), another potential area of concern pertains to how interactions with an agent can jeopardise the privacy of an individual's emotional experiences. Agents within the AIED context also possess the capability to induce or manipulate emotions, and virtual relationships with these agents may blur the lines between reality and fiction, potentially resulting in psychological challenges. These studies emphasise that the discussions regarding the management of emotions in educational technology play a crucial role within the larger ethical conversation in this domain.

RQ4. How is capacity building on AIED ethics being covered?

Table 8 presents the 41% of studies in this SLR covering AI ethics education and the targets, content, skills, and delivery methods of such education. Through a systematic policy review, Schiff (2021) found that education for AI includes training AI professionals (*e.g.*, computer scientists), preparing the workforce for AI, and broader public AI literacy. Many authors stress the need for ethics in Engineering education to bridge the gap between technology and society (Antoniou, 2021; Dignum, 2021; Dignum, 2020; Hoeschl, 2017; Park et al., 2021; Williams et al., 2020). Qualitative data reveals that US Information and Computer Science students often fail to consider the ethical implications of AI design for privacy and well-being without explicit guidance (McDonald & Pan, 2020). Another exploration in the US analysed 31 standalone AI ethics classes and 20 AI/Machine Learning technical courses, revealing both commendable practices and notable omissions, such as accessibility, diversity in the AI workforce, and sustainability (Garrett et al., 2020). In 2020, research conducted across 12 Australian universities indicated a lack of ethics education in Computer Science courses or a focus solely on micro-ethical concepts linked to professionalism and industry standards (Gorur et al., 2020). The University of North Carolina piloted an AI ethics course for Computer Science students, focusing on explicit ethical agents, and suggested the value of prototyping/hands-on approaches and challenging students to employ diverse ethical approaches and summarise their implications (Green, 2021). Similarly, the University of Oulu in Finland introduced an AI ethics pilot course, covering a range of AI applications, legislative and ethics aspects, and the pros and cons of AI applications (Tuovinen & Rohunen, 2021). Future implementations were recommended to incorporate unexpected ethical issues, including methods like case study analyses and role reversal between defenders and opponents.

Studies in various fields of health, including anatomy, psychiatry, and clinical psychology, highlight the growing importance of ethics and education in the

Table 8 Capacity building on AI/ED ethics

Studies	Targets	Training Content and Skills	Delivery Methods
Antoniou, 2021; Bogina et al., 2021; Dignum, 2020; Dignum, 2021; Garrett et al., 2020; Goldsmith et al., 2020; Gorur et al., 2020; Green, 2021; Herzog et al., 2022; Lucic et al. (2022); McDonald & Pan, 2020; Miao et al., 2021; Schiff, 2021; Sjöden, 2020, Tuovinen & Rohunen, 2021; Williams et al., 2020	Future AI professionals, especially Engineering	Distributed nature of AI applications (human-agent interaction); distributed learning entities (autonomous & self-organising AI); inclusion/diversity in design and in AI workforce; governance, sociotechnical, legal and economic models; prediction models to understand AI impact on every profession; multiple competencies	Seminars, summer schools, challenge-based learning engineering ethics courses, embedded lectures; hands-on approaches and case-studies
Dignum, 2021; Miao et al., 2021; Reisch, 2021; Xu, 2020	General Public (Including vulnerable and older people)	AI literacy; imagination and humanities; opportunities for on-the-job training; diverse and inclusive AI; training of ethicists among citizens; capacity to adapt to rapid change	AI online media courses
Ali et al., 2019; Bates, 2011; Bleas et al., 2021; Chiu et al., 2021; Charisi et al., 2020; Dignum, 2021; Gauld et al., 2021; Javed et al., 2022; Karaca et al., 2021; Katznelson & Gerke, 2021; Khurana, M., 2020; Kong et al., 2023; Lazarus et al., 2022; Lee, 2014; Lee et al., 2021; Loftus & Madden, 2020; Miao et al., 2021; Ng et al. (2023); Ottenbreit-Leftwich et al., 2023; Shih et al., 2021; Srinivasan & Uchino, 2021; Zhang et al., 2023	Students (k-12, TVET & Higher Education)	STEM, humanities, arts and social sciences together; medical courses – philosophy of ethics, cognitive neuroscience, computational psychiatry, clinical practice; informed choices about mental health through AI apps; perceived AI readiness (e.g., MAIRS-MS); ethical dilemmas concerning the use of AI in medical practice; agency; meaningful content connecting to prior knowledge; technical concepts and processes, ethical and societal implications, and career futures in the AI era	Mandatory modules, lectures or elective courses; constructivist hands/minds-on experiences w/robots; project-based pedagogy; cognitive levels, course content, and disciplines (BAG model)

Table 8 (continued)

Studies	Targets	Training Content and Skills	Delivery Methods
Bogina et al., 2021; Chounta et al., 2022; Gary, 2019; Holmes et al., 2021; Luo, 2019; Miao et al., 2021; Mouta et al., 2021; Ottenbreit-Leftwich et al., 2023; Sjöden, 2020; Williams & Breazeal, 2020; Yau et al., 2022	Teachers, School Deans, Researchers (k-12 & Higher Education)	Equitable AI research; personal emotional support; AIED history, philosophy of technology use and adoption; usefulness of AI in education; application of AI-powered EMIS/LMS; computational thinking; human interaction, higher-order thinking, and human values; content that teachers feel confident to teach	Non-traditional educational approaches; participatory methods

context of AI adoption. While AI holds promise for improving diagnoses and prognoses, it must align with medical epistemological frameworks to address emerging ethical and clinical concerns (Gauld et al., 2021). Research with clinical psychology Master's students at the University of Basel has shown that despite some familiarity with AI/ML tools, they require education to assist patients in making informed choices regarding mental health AI/ML applications, taking into account issues like privacy, equality, and discrimination (Blease et al., 2021). Other authors suggest that when preparing curricula for AI education, it is crucial to consult with students to understand their needs and their perceived readiness for AI-related topics. To facilitate this, Karaca et al. (2021) have developed the MAIRS-MS, a reliable tool for assessing students' readiness for AI and its applications in medicine.

Schiff's (2021) study found that ethical training for the appropriate implementation of AI in the education sector was almost absent. Ng et al. (2023) conducted a SLR that includes thematic and content analysis of 49 publications from 2000 to 2020. The review highlights that AI teaching and learning primarily focused on computer science education at the university level before 2021. The findings presented in this review emphasise the significance of educating individuals in AI literacy and AI ethics. So, there are not as many programmes reporting training on ethical issues around AI, and even fewer on best practices for ethical integration of AI resources. Additionally, there are no reported tools to assess pedagogical practices using AI. Even in training projects conducted in schools, there is a noticeable absence of emphasis on the ethical considerations associated with AIED. Furthermore, this training still addresses shortcomings in countries where AI is already a reality in the classroom, such as China. Its unsystematic nature, non-intentionality and the lousy quality of its supporting materials show there is still a long way to go (Gong et al., 2021). Furthermore, when not absent, it appears that training specifically for the use of AIED is only being provided to teachers, school administrators, and researchers. Loftus and Madden (2020) emphasise the importance of ensuring that students also understand the datafication of their own lives and learning processes, and advocate for placing students at the heart of the construction of AI-powered models, which again draws attention to the importance of sense of agency.

According to Dignum (2021), the digital age is no longer compatible with the separation of STEM from humanities, arts, or social sciences. In fact, arts seem to be a great platform to promote AI education, considering it an expression medium and its role fostering empathy, diversity, and inclusion in the AI pipeline (Srinivasan & Uchino, 2021). Xu (2020) also suggests that all ethical challenges in introducing AIED must be considered from a humanistic educational perspective. Following this principle, some examples of AI ethics training in schools were conceived. Research from Ottenbreit-Leftwich et al. (2023) focuses on introducing AI education to K-12 students and explores the potential starting point for teachers to teach AI and computer science concepts. It suggests that AI ethics can be a compelling entry point. Teachers showed more confidence in discussing AI ethics with their students, which led to engaged discussions. The research aims to lay the groundwork for elementary AI education by considering students' ideas, experiences, and ethics as essential components for curriculum design in K-12 education.

Nevertheless, another topic of concern has been exposed. Although 42 studies were dedicated to the topic, only four of them covered the impact assessment phase of these training programmes. A study by Lee et al. (2021) considered the impact of a summer workshop on AI literature on middle school students, mainly from under-represented groups in STEM. Certain benefits included a notable improvement in students' grasp of AI and its potential biases, enhanced adaptability to future AI-related employment, a better understanding of the consequences of their actions on others, improved capacity to discuss ethical AI issues with their families, and the ability to leverage their family's resources for self-improvement. Another study (Kong et al., 2023) highlights the potential of AI literacy education for senior secondary students, emphasising that programming knowledge is not a prerequisite for understanding AI concepts. The results suggest that with sufficient learning time and project-based pedagogy, senior secondary students can develop AI literacy, although there may be challenges in comprehending complex AI ethical principles, which require further guidance and time. The work of Lucic et al. (2022), presents a course at the University of Amsterdam aiming at providing students with a comprehensive understanding of FACT-AI topics and algorithmic harm through lectures, paper discussions, and a reproducibility project. Students engage with the open-source and research communities, creating a public code repository. The course emphasises the importance of reproducibility and received positive feedback from students who appreciate the critical perspectives gained and insights into AI research. The course successfully motivates students and promotes critical thinking in AI. Finally, using a phenomenographic approach, Yau et al. (2022) studied 28 in-service teachers from 17 secondary schools in Hong Kong after implementing an AI curriculum. They identified six categories of teacher conceptions related to teaching AI, including technology bridging, knowledge delivery, interest stimulation, ethics establishment, capability cultivation, and intellectual development. The study presents a hierarchical outcome space that illustrates the range of surface to deep conceptions held by teachers. It offers insights into cultivating both technical and non-technical teachers' competence in AI education, aiding teacher educators and policymakers in enhancing AI education for K-12 students.

RQ5. What principles, regulations and frameworks are there for AIED?

The concerns described in the previous RQs have justified the adoption of guidelines for the use of AI in education. The first attempts to adopt general ethical usage guidelines were unsuccessful, as it was quickly realised that the specifics of the education sector required a specific framework. So, educational actors are now requested to produce workable ethical frameworks to tackle AI potential risks, improving educational institutions and student outcomes (Weber, 2020). However, it is crucial to develop consistent terminology and scope in formal standardisation efforts, especially in the context of Information Technology for Learning, Education, and Training, within and across standardisation bodies (Mason et al., 2020). Furthermore, it's worth noting that specific discussions on AIED in education policies are taking place in countries like China, India, Italy, Kenya, Malta, Singapore, South Korea, Spain and the United States. Although these nine countries are discussing

some version of AI for education, only four or five are discussing it in depth beyond superficial mentions (Schiff, 2021).

As indicated in Table 9, which pertains to the 8% of articles discussing this subject, some ethical frameworks show tensions and gaps concerning the ethical advancement and application of learning analytics, while others are not appropriate for use in the field of education. For example, the GDPR is too complex when used in education (Kitto & Knight, 2019) and it can be difficult to apply these guidelines in sensitive cases where the primary concern is the safety of people (Al-Omran et al., 2019). As of 2019, there were no specific policies or regulations regarding AIED, despite efforts to ensure trustworthy AI (Holmes et al., 2018). In response to this absence, guidelines such as the "Beijing Consensus on Artificial Intelligence and Education" (UNESCO, 2019b) and "The Ethical Framework for AI in Education" (The Institute for Ethical AI in Education, 2021) have been introduced. The UNESCO Consensus on AI in education was developed through the collaboration of multiple stakeholders, including government ministers, international representatives, and experts. It outlines seven key principles for UNESCO's member states: prioritise AI in education policies to achieve SDG 4 goals, support AIED-enhanced pedagogies when benefits outweigh risks, promote AI tools for lifelong and personalised learning, base policies on evidence, provide comprehensive AI training for teachers, cultivate critical skills for the AI era, and encourage equitable, transparent, and ethical use of AIED, with a focus on gender equality (UNESCO, 2019b). "The Ethical Framework for AI in Education" (The Institute for Ethical AI in Education, 2021) appears to be the most sophisticated and up-to-date tool to closely monitor AI technologies used in the various stages of AI adoption in education, from pre-procurement to implementation and impact evaluation. Launched in 2018 by the University of Buckingham, this framework became publicly available in 2021. It resulted from a two-and-a-half-year effort, which included collecting interviews from policymakers, academics, philosophers, ethicists, industry experts, and educators to establish a consensus regarding the integration of AI into the education sector. The framework addresses ethical design, privacy, equity, transparency, and accountability concerns, considering the sector's specific needs. For example, in relation to privacy, it states that even if an institution is required to continuously assess students, it must also establish safe spaces where learners are not assessed. When it comes to autonomy and agency, the framework draws attention to the actions that should be taken when the AI system predicts an unfavourable outcome. It also underscores the importance of ensuring that AI systems are designed to benefit learning without leading to addictive behaviours. This framework is the first step in addressing the ethical challenges of AIED and ensuring it is educationally sound from the ground up.

Finally, Miller and Tuomi (2022) emphasise the importance of sense-making when envisioning the future. They highlight the significance of employing various forms of anticipation to perceive the world from different perspectives. This approach is rooted in a shift in the field of futures studies towards an ontological perspective, which allows to reconceptualise the future as a point of origin. It positions anticipation and anticipatory processes about the future as integral aspects of the present. This challenge could be integrated into the ongoing enhancement of

Table 9 Ethical AI and AIED Guidelines and Frameworks

Studies	Guidelines/ Frameworks	Purpose
Al-Omran et al., 2019	Software Engineering Code of Ethics and Professional Practice; IEEE Code of conduct; ACM Code of Ethics	Preparing stakeholders on tech ethical considerations
Kitto & Knight, 2019	General Data Protection Regulation (GDPR)	Safeguarding individuals' rights when producing personal data in daily activities
UNESCO, 2019a	UNESCO's Rights, Openness, Access and Multi-stakeholder Governance)	Guaranteeing human rights and emerging ethical issues in the application of AIED
Islam et al., 2019	Ethics guidelines for trustworthy AI—High-Level Expert Group on AI	Assuring human agency, accountability and transparency; technical safety; diversity, non-discrimination and fairness; societal and environmental well-being
Dignum, 2021	ART principles for responsible AI	Anticipating AI societal impact; accountability, responsibility, and transparency
UNESCO, 2019b	Beijing Consensus on Artificial Intelligence and Education	Achieving SDG 4; beneficial AIED-empowered pedagogies; ubiquitous lifelong and personalised learning; evidence-based policies; auditable uses of AIED
The Institute for Ethical AI in Education, 2021	The Ethical Framework for AI in Education	Ensuring AIED provides the greatest possible advantage to the learner

frameworks related to AIED and educational ethics, fostering broader, more creative, and innovative processes that aren't limited by predefined objectives.

Discussion

The following sections delve into various aspects concerning the implications of the findings from this SLR. They encompass the integration of a theoretical model for analysing pedagogical implications based on the collected data, the identification of research gaps, recommendations for teacher training in AIED ethics, and a critical examination of potential limitations in this research.

Implications of this SLR for rethinking the ethics of education: using a theoretical model

The introduction of a new technology in the realm of education provides an opportunity not only to reconsider the specific ethical consequences it introduces but also to contemplate the broader ethics associated with education itself. The ik.model (Mouta et al., 2015) represents an extended framework building upon the TPACK model, encompassing the human element within educational technology integration. This framework comprises four pedagogical dimensions: "technologies and resources", "content knowledge", "learning processes and strategies", and "educational actors and their relationships". Subsequently, this theoretical framework is applied to assess the potential consequences of the aforementioned results.

Technologies and Resources Dimension

RQ1 and RQ3 emphasise the critical importance of creating inclusive AI systems in education, underscoring the significance of diversity in decision-making. Yet, ethical discussions on AI in education are primarily Western-centric and STEM-dominated, as indicated by MQ2, MQ3, and MQ4, potentially resulting in biases and hindering inclusive AI system development. RQ5 stress the importance of collaboration in evaluating ethical frameworks for AIED, highlighting the need to balance student interests, meaningful innovation, and overcoming resistance to change in education. Furthermore, it points out the challenges of applying these frameworks across diverse educational contexts and levels. In light of these findings, three key outcomes emerge: (1) funding for research and educational programmes including underrepresented areas (e.g., arts, health, humanities), populations and groups, exemplified by AIED DEIA mentoring fellowships ("Call for Fellow Nominations", 2023); (2) systematic collaboration among AI researchers, developers, and practitioners, drawing from school fieldwork to create technically robust and culturally sensitive AI systems. This would be an opportunity to build on teachers' own ethical development, by creating opportunities for them to exercise their skills as judging actors (Arendt, 1958). (3) It is recommended to develop an assessment tool that can adapt to different contexts and monitor the positive impact of these technologies,

involving a wide range of stakeholders. These frameworks should also ensure seamless integration with educational standards as students progress through their educational journey and into their jobs.

Content Knowledge Dimension

The results from RQ4 indicate that while there are training programmes addressing ethical challenges in AIED for teachers, school administrators, and researchers, students usually receive limited exposure to the ethical use of AI. Their education primarily focuses on deontological and on-the-job ethics. This approach overlooks the valuable perspective that students can bring to the development of AIED environments, considering that they are the primary beneficiaries of these technologies. To bridge this gap, curriculum infusion is proposed as a suitable strategy. It allows for the inclusion of emerging topics in a meaningful way, while accommodating the busy academic calendar. By incorporating AI concepts across various subjects and applying them to tasks utilising AIED resources, educators can facilitate discussions about AI's functions, impacts, and ethics from diverse subject perspectives. Ultimately, this approach enhances students' acquisition of content knowledge, application to practical tasks, and cultivates a comprehensive understanding of AI systems and their effects on individuals and groups.

Learning Processes and Strategies Dimension

The results of RQ1, RQ2, RQ3, and RQ5 underscore the importance of preserving students' diversity and sense of agency in AI-based education. For personalisation in education to be effective, it needs to strike a balance between tailoring content to individual needs and ensuring that students have the opportunity to explore diverse perspectives and develop a holistic self-understanding. Exploring one's own interests, competencies, and values – three critical dimensions of identity – is essential for informed decision-making throughout life and fostering a strong attachment to the learning experience. Thus, there is a need for epistemological reflection as an essential aspect of contemplating the significance of pedagogical innovations in the contemporary world (Trindade & Cosme, 2010). This goal can bring teachers and students together in project-based learning, encouraging them to explore diverse sources of knowledge and engage with them using various processes and literacies. This is even more important given the access to language models like ChatGPT. The challenges it entails present an opportunity to place questioning and critical thinking at the core of education. Approaches like the flipped classroom can help students engage meaningfully with AI technologies, fostering thoughtful exploration. Independent learning methods, followed by teacher-led discussions focused on problem-solving, offer a dual benefit: they allow students to deepen their understanding independently and encourage dialogic developmental processes when they share and explore their learning outcomes with their peers. Developing these skills empowers students as citizens, while assisting societies in navigating the evolving technology landscape.

Moreover, adapting evaluation criteria to accommodate diverse student expressions is critical when AI offers customised learning experiences. This process entails analysing AI-generated data and presenting it to students, teachers, and families for input and a systemic understanding. Effective use of AIED assessment tools can encourage self-assessment as a valuable learning tool, allowing students to critically evaluate technology, promoting self-reflection, self-regulation, and citizenship. In the era of dataism, educators face the challenge of exploring dataism's onto-epistemic grammar with their students, including its anthropocentric perspective, the drive for ontological security, and the thirst for absolute knowability (Andreotti et al., 2015; Lados et al., 2022; and Stein et al., 2017).

Only 4% of the papers in this field address the ethical aspects of AIED's emotional management (cf. RQ3), indicating a predominant focus on performance as the primary rationale for using AIED. However, this emphasis on performance neglects the fact that it results from a complex interplay of both cognitive and emotional factors. Emotional well-being and social skills are key for overall achievement. Currently, these AI technologies often align with an educational paradigm that prioritises performance and global rankings. Adhering to established and conventional assumptions about knowledge, teacher and learner roles, educational goals, and learning methods when incorporating AI systems may exacerbate existing issues, making distances wider, instead of addressing broader educational and developmental needs. Such an approach also hinders the potential of these technologies to foster a decentralised and self-directed approach to education, as advocated by thinkers like Illich (1971). By solely concentrating on the short-term advantages, educators and policymakers might overlook or underestimate the broader impacts and ethical considerations associated with AI in education, hindering its beneficial potential.

Educational Actors and their Relationships Dimension

RQ1 and RQ2 shed light on the importance of recognising the significance of shared values, both explicit and implicit educational agreements, and the involvement of diverse stakeholders in decision-making concerning the use of AIED. MQ3 has also suggested that achieving geographical diversity in research on AIED ethics remains a persistent challenge. This recognition should be coupled with an awareness of disparities in access and variations in pedagogical quality. According to Christakopoulou et al. (2001), a school is a multifaceted entity encompassing aspects of a social, economic, and political community. Additionally, it serves as a personal environment where attachments are formed, and memories are created. The introduction of AI-based education challenges all these dimensions, prompting a critical examination of the roles, relationships, and power dynamics of various stakeholders involved in the educational ecosystem. It calls for a reevaluation of how education is not only about imparting knowledge but also about nurturing a sense of belonging, empowerment, and active participation within a rapidly evolving educational landscape.

Uncovering research blind spots on AIED ethics

To tackle the challenge of rethinking education ethics in the age of AIED, it is crucial to carefully consider both the insights gained from this SLR and the aspects that remain concealed or overlooked. One of those missing aspects has to do with the lack of incorporation of AIED ethics in philosophical or psychological paradigms of moral or ethical development. This vagueness in discussions may limit the definition of criteria to thoroughly evaluate the impacts of AIED use. Moreover, there is a notable oversight when it comes to discussions about the broader concept of transhumanism introduced by AI systems. There was only one paper that raised awareness of the question of whether there should be limits on using technology to extend or enhance cognitive abilities, as well as the dynamics of relationships in hybrid environments. How are issues of equity addressed when only a privileged few will have access to competitive AI features, granting them a significant advantage? Additionally, discussions often lack critical perspectives on the ethical boundaries of the concept of intelligence as defined by AIED research and design. Overreliance on behaviourist and cognitive approaches may sideline aesthetics, emotions, morality, and social development, potentially reinforcing reductionist views of human intelligence. Prioritising optimisation may undermine the value of reflective and contemplative thinking, which is crucial to develop strategies for solving complex problems. Another perspective that can be gained from this study is that by adopting a participatory approach involving educators and students in research and inquiry-based learning, the full potential of AIED can be harnessed while encouraging a reflexive and critical attitude that supports comprehensive ethical growth.

Recommendations for educators on AIED ethics

It becomes clear that the ethical implementation of AIED requires a comprehensive approach. The results emphasise the importance of participatory methods and dialogic ethics within this context. The following recommendations on AIED ethics are based on the insights derived from this study and will be followed on the continuation of this project research: (1) engaging educators through focus groups – respecting the insights from this research, it is recommended arranging focus group sessions with educators to provide them with a platform to engage in meaningful discussions about the ethical use of AIED in educational institutions; (2) developing a pilot training programme – data generated from these focus group sessions will form the basis for creating a pilot training programme aimed at promoting ethical considerations and practices in the integration of AIED within educational settings; (3) exploring content and delivery methods accordingly – during the programme's development, it is critical to align with research findings, adhering to a socio-constructivist framework that emphasises active participation and engagement when designing content, delivery methods, duration, and activities; (4) training in comprehensive topics – develop the training curriculum to cover a wide range of topics, including the ones that were overlooked in previous work (AIED technologies,

potentialities, and challenges; ethical considerations specific to education; interdisciplinary perspectives on AIED; stakeholders in AIED, power dynamics, interests, and needs; effects of AIED ethical challenges on student agency, self-determination, and emotional well-being; AIED's role in pedagogy and innovative learning processes; learning analytics, data collection, analysis, and interpretation in the context of AIED; evaluation of AIED effectiveness using informed criteria; promotion of communities of practice in AIED for knowledge sharing and addressing challenges); (5) role-playing with students – incorporate challenging and hands-on engaging tasks that encourage reflection on how AIED-specific applications could impact students' lives, which can then be further discussed in training; (6) evaluating through self and peer assessment – align with the training's defined criteria, facilitating an evaluation of the created resources' responsiveness and effectiveness; (7) promoting inclusive participation – invite a diverse group of teachers as trainees, including individuals from typically underrepresented countries and representing various fields; (8) assessing attitudes towards AIED ethics – include quantitative data collection moments at the beginning and end of the training programme to facilitate the measurement of its impact and inform future improvements; (9) gathering teacher perspectives – involve teachers in a final semi-structured interview to gain insights into their perspectives regarding key criteria that should be prioritised in a continuing professional development programme on AIED ethics, with a specific focus on their training experiences.

Potential limitations of this research

In conclusion, it is important to acknowledge that, despite the inclusion of over 100 papers in this study, there are certain limitations to consider. Since our research covers papers published until the end of 2022, studies currently being published are not taken into account. The most recent research can provide updated insights as the field is growing both conceptually and in school practice. Another limitation of this SLR can be some articles left behind. In fact, the general search term “artificial intelligence” was chosen because it would be difficult to cover all its technologies, such as “robots”, “educational chatbots”, “machine learning”, “intelligent tutoring systems”, “exploratory learning environments”, “teachable agents”, “dual-teacher model”, “speech/image recognition”, “autonomous agents”, “affect detection”, and so on. The idea was that papers covering a specific AIED technology would mention the word “artificial intelligence” at least once. Finally, 24 papers were added as first or second references in the later stage of the review. There is a possibility that other papers with similar characteristics were overlooked during the initial screening process, which relied on titles and abstracts for assessment. The implication of potentially missing such papers is that the review may not have captured the full spectrum of relevant literature, and valuable insights or perspectives on AIED and ethics could have been omitted from the analysis. To mitigate this limitation, multiple rounds of screening and comprehensive search strategies were employed. However, despite best efforts, it is possible that some relevant papers may still have been inadvertently omitted.

Conclusion

Research interest in the ethics of designing, developing, and implementing AIED has been steadily increasing since 2018. This interest has experienced a surge after 2020, as a result of the COVID-19 pandemic and the rise of AIED technologies, leading to a significant increase in AIED ethics research. The majority of papers analysed in this SLR presented concerns about dimensions such as fairness, inclusion, autonomy, and agency. The publications analysed also showed efforts to implement AIED ethical programmes in K-12 and higher education. The main conclusions drawn from this SLR highlight the significance of participatory processes in researching and implementing AIED. To achieve this, it is essential to engage a range of educational stakeholders, including students, teachers, school administrators, parents, and researchers. Although regulatory ethical frameworks for AIED have been introduced, they have come quite late and do not account for the specific needs of learners and pedagogy. Generic AI ethics principles do not adequately address the part of agency responsible for learning, which no technology with promises of better personalisation or guidance should struggle against. Therefore, ethical frameworks for the design and use of AIED should be developed through participatory processes that recognise the specific needs and tasks of its main actors (students and teachers), and respect each learning community's heritage, as well as the will and capacity for innovation.

Furthermore, this SLR uncovered the importance of providing teachers with support to effectively utilise AIED technologies while preserving students' sense of agency and promoting lifelong learning potential. Additionally, rethinking and aligning evaluation parameters is crucial to ensure that ethical concerns are taken into account and that data is incorporated from multiple feedback sources. The SLR also revealed a failure to incorporate AIED ethics into a philosophical or psychological paradigm of moral or ethical development. The lack of a defined conceptual corpus that fits the educational landscape makes it difficult to understand the extent to which AIED meaningfully addresses ethical aspects of individual, community, and organisational development. Moreover, using AI technologies can facilitate self-reflection, reflection on the actions of other agents (persons and machines), and balance personal standards and value systems while considering interpersonal connections and mutual obligations.

In conclusion, this systematic literature review sheds light on the key topics that should be included in teacher training on education ethics when using AIED. It also suggests that while AIED research has been addressing ethical considerations, there is still significant room for growth in terms of analysing its unique heuristics within the context of education ethics.

Acknowledgements This research is being developed with a grant from Fundação para a Ciência e Tecnologia (No.2020.06258.BD).

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data availability The database supporting this SLR is publicly available at: <https://docs.google.com/spreadsheets/d/1u6ArXbQ5w4bxhHbKAcRUCZLUSACK0AoG/edit?pli=1#gid=1005218435>.

Declarations

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abboud, R., Arya, A., & Pandi, M. (2020). Redefining the Digital Divide in the age of AI: the harvest of the 25th anniversary. In L. Gómez Chova, A. López Martínez & I. Candel Torres (Eds.), *INTED2020 Proceedings* (pp. 4483–4492). IATED Academy. <https://doi.org/10.21125/inted.2020.1241>
- Adams, C., Pente, P., Lermeyer, G., Turville, J., & Rockwell, G. (2022). Artificial Intelligence and Teachers' New Ethical Obligations. *The International Review of Information Ethics*, 31(1). <https://doi.org/10.29173/irrie483>
- Ali, S., Williams, R., Payne B., Park H., & Breazeal, C. (2019). Constructionism, ethics, and creativity: Developing primary and middle school artificial intelligence education. *Proceedings of the International Workshop on Education in Artificial Intelligence K-12*. Retrieved January 19, 2022, from <https://www.media.mit.edu/publications/constructionism-ethics-and-creativity/>
- Al-Omran, G., Al-Abdulhadi, S., & Jan, M.R. (2019). Ethics in artificial intelligence. *Proceedings of the International Conference on Industrial Engineering and Operations Management*, 940–949. Retrieved January 19, 2022, from <https://www.ieomsociety.org/gcc2019/papers/337.pdf>
- Andreotti, V., Stein, S., Ahenakew, C., & Hunt, D. (2015). Mapping interpretations of decolonization in the context of higher education. *Decolonization: Indigeneity, Education & Society*, 4(1), 21–40.
- Antoniou, J. (2021). Dealing with emerging AI technologies: Teaching and learning ethics for AI. In J. Antoniou (Ed.), *EAI/Springer Innovations in Communication and Computing* (pp. 79–93). Springer. https://doi.org/10.1007/978-3-030-52559-0_6
- Arendt, H. (1958). *The human condition*. University of Chicago Press.
- Bates, R. A. (2011, June). *AI & SciFi: Teaching Writing, history, Technology, Literature, and Ethics* [Paper presentation]. ASEE Annual Conference & Exposition, Vancouver, BC. <https://doi.org/10.18260/1-2-17433>
- Belpaeme, T., Kennedy, J., Ramachandran, A., Scassellati, B., & Tanaka, F. (2018). Social robots for education: A review. *Science Robotics*, 3(21). <https://doi.org/10.1126/scirobotics.aat5954>
- Blease, C., Kharko, A., Annoni, M., Gaab, J., & Locher, C. (2021). Machine Learning in Clinical Psychology and Psychotherapy Education: A Mixed Methods Pilot Survey of Postgraduate Students at a Swiss University. *Frontiers in Public Health*, 9, 623088. <https://doi.org/10.3389/fpubh.2021.623088>
- Bogina, V., Hartman, A., Kuflik, T., & Shulner-Tal, A. (2021). Educating Software and AI Stakeholders About Algorithmic Fairness, Accountability, Transparency and Ethics. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00248-0>
- Bozkurt, A., Karadeniz, A., Bañeres, D., Guerrero-Roldán, A., & Rodríguez, M. E. (2021). Artificial Intelligence and Reflections from Educational Landscape: A Review of AI Studies in Half a Century. *Sustainability*, 13, 800. <https://doi.org/10.3390/SU13020800>
- Bucea-Manea-Tonis, R., Kuleto, V., Gudei, S. C. D., Lianu, C., Lianu, C., Ilic, M. P., & Paun, D. (2022). Artificial Intelligence Potential in Higher Education Institutions Enhanced Learning Environment in Romania and Serbia. *Sustainability*, 14(10), 5842. <https://doi.org/10.3390/su14105842>

- Burr, C., Taddeo, M., & Floridi, L. (2020). The Ethics of Digital Well-Being: A Thematic Review. *Science and Engineering Ethics*, 26, 2313–2343. <https://doi.org/10.1007/s11948-020-00175-8>
- Call for Fellow Nominations (2023, 31 march). *AIED 2023*. Retrieved April 28, 2023, from https://www.aied2023.org/c_mentoring_fellowship.html
- Cave, S. (2020). The Problem with Intelligence: Its Value-Laden History and the Future of AI. In A. Markham, J. Powles, T. Walsh, A.L. Washington (Eds.), *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)* (pp. 29–35). Association for Computing Machinery. <https://doi.org/10.1145/3375627.3375813>
- Charisi, V., Malinverni, L., Rubegni, E. & Schaper, M.M. (2020). Empowering Children's Critical Reflections on AI, Robotics and Other Intelligent Technologies. In I. Šmorgun & G. Berget (Eds.), *Proceedings of the 11th Nordic Conference on Human-Computer Interaction: Shaping Experiences, Shaping Society* (pp. 1–4). Association for Computing Machinery. <https://doi.org/10.1145/3419249.3420090>
- Chiu, T. K. F., Meng, H., Chai, C., King, I., Wong, S., & Yam, Y. (2021). Creation and Evaluation of a Pre-tertiary Artificial Intelligence (AI) Curriculum. *IEEE Transactions on Education*, 65(1), 30–39. <https://doi.org/10.1109/TE.2021.3085878>
- Chounta, I. A., Bardone, E., Raudsep, A., & Pedaste, M. (2022). Exploring Teachers' Perceptions of Artificial Intelligence as a Tool to Support their Practice in Estonian K-12 Education. *International Journal of Artificial Intelligence in Education*, 32, 725–755. <https://doi.org/10.1007/S40593-021-00243-5>
- Christakopoulou, S., Dawson, J., & Gari, A. (2001). The Community Well-Being Questionnaire: Theoretical Context and Initial Assessment of Its Reliability and Validity. *Social Indicators Research*, 56, 319–349. <https://doi.org/10.1023/A:1012478207457>
- Coghlan, S., Miller, T., & Paterson, J. (2021). Good Proctor or “Big Brother”? Ethics of Online Exam Supervision Technologies. *Philosophy & Technology*, 34, 1581–1606. <https://doi.org/10.1007/s13347-021-00476-1>
- Córdova, P.R., & Vicari, R.M. (2022). Practical Ethical Issues for Artificial Intelligence in Education. In A. Reis, J. Barroso, P. Martins, A. Jimoyiannis, R.YM. Huang, & R. Henriques (Eds.), *Technology and Innovation in Learning, Teaching and Education. TECH-EDU 2022. Communications in Computer and Information Science* (pp. 437–445). Springer, Cham. https://doi.org/10.1007/978-3-031-22918-3_34
- Costa, A.P., Charpiot, L., Lera, F.J., Ziafati, P., Nazarihorram, A., van der Torre, L., & Steffgen, G. (2018). More Attention and Less Repetitive and Stereotyped Behaviors using a Robot with Children with Autism. In J.J. Cabibihan, F. Mastrogiovanni, A.K. Pandey, S. Rossi, M. Staffa (Eds.), *27th IEEE International Symposium on Robot and Human Interactive Communication (RO-MAN)* (pp. 534–539). e IEEE Xplore Digital Library. <https://doi.org/10.1109/ROMAN.2018.8525747>
- Dieterle, E., Dede, C., & Walker, M. (2022). The cyclical ethical effects of using artificial intelligence in education. *AI & society*, 1–11. Advance online publication. <https://doi.org/10.1007/s00146-022-01497-w>
- Dignum, V. (2020). AI is multidisciplinary. *AI Matters*, 5(4), 18–21. <https://doi.org/10.1145/3375637.3375644>
- Dignum, V. (2021). ‘The role and challenges of education for responsible AI’. *London Review of Education*, 19(1), 1, 1–11. <https://doi.org/10.14324/LRE.19.1.01>
- Dobrosrovestnova, A. & Hannibal, G. (2020). Teachers' Disappointment: Theoretical Perspective on the Inclusion of Ambivalent Emotions in Human-Robot Interactions in Education. In T. Belpaeme, J. Young, H. Gunes & L. Riek (Eds.), *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 471–480. Association for Computing Machinery. <https://doi.org/10.1145/3319502.3374816>
- Epp, C. D., & Makos, A. (2013). Using simulated learners and simulated learning environments within a special education context. *CEUR Workshop Proceedings*, 1009, 1–10. Retrieved January 19, 2022, from <https://ceur-ws.org/Vol-1009/0401.pdf>
- Gallastegui, G., Miguel, L., & Forradellas, F. R. (2021). Business Methodology for the Application in University Environments of Predictive Machine Learning Models Based on an Ethical Taxonomy of the Student's Digital Twin. *Administrative Sciences*, 11(4), 118. <https://doi.org/10.3390/admsci11040118>
- García-Peñalvo, F.J., Corell, A., Abella-García, V. & Grande-De-Prado, M. (2021). Recommendations for Mandatory Online Assessment in Higher Education During the COVID-19 Pandemic. In D. Burgos, A. Tlili, & A. Tabacco (Eds.), *Radical Solutions for Education in a Crisis Context. COVID-19*

- as an Opportunity for Global Learning (pp. 85–89). Springer Nature. https://doi.org/10.1007/978-981-15-7869-4_6
- Garrett, N., Beard, N., & Fiesler, C. (2020). More Than "If Time Allows": The Role of Ethics in AI Education. In A. Markham, J. Powles, T. Walsh, A.L. Washington (Eds.), *Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society (AIES '20)* (pp. 272–278). Association for Computing Machinery. <https://doi.org/10.1145/3375627.3375868>
- Gary, E. (2019). Ethics to prepare educators for professional service robots in classrooms. In E. Miyoshi, E. G. Kowch, J. C. Liu, Q. Jin, Z. Li, H. Zhang (Eds.), *International Joint Conference on Information, Media, and Engineering Proceedings* (pp. 478–484). INSTITUTE OF ELECTRICAL AND ELECTRONICS ENGINEERS. <https://doi.org/10.1109/IJCIME49369.2019.00102>
- Gauld, C., Micoulaud-Franchi, J., & Dumas, G. (2021). Comment on Starke et al.: 'Computing schizophrenia: Ethical challenges for machine learning in psychiatry': From machine learning to student learning: Pedagogical challenges for psychiatry. *Psychological Medicine*, 51(14), 2509–2511. <https://doi.org/10.1017/S0033291720003906>
- Ghotbi, N., & Ho, T. (2021). Moral Awareness of College Students Regarding Artificial Intelligence. *Asian Bioethics Review*, 13(4). <https://doi.org/10.1007/s41649-021-00182-2>
- Goldsmith, J., Burton, E., Dueber, D., Goldstein, B., Sampson, S., & Toland, M. (2020). Assessing Ethical Thinking about AI. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(09), 13525–13528. <https://doi.org/10.1609/aaai.v34i09.7075>
- Gong, X., Tang, Y., Liu, X., Jing, S., Cui, W., Liang, J., & Wang, F. Y. (2021). K-9 Artificial Intelligence Education in Qingdao: Issues, Challenges and Suggestions. *2020 IEEE International Conference on Networking, Sensing and Control*, 1–6. <https://doi.org/10.1109/ICNSC48988.2020.9238087>
- Gorur, R., Hoon, L., & Kowal, E. (2020). Computer Science Ethics Education in Australia – A Work in Progress. In H. Mitsuhashi, Y. Goda, Y. Ohashi, M. M. T. Rodrigo, J. Shen, N. Venkatarayalu, G. Wong, M. Yamada, C.U. Lei (Eds.), *2020 IEEE International Conference on Teaching, Assessment, and Learning for Engineering (TALE)*, 945–947. IEEE. <https://doi.org/10.1109/TALE48869.2020.9368375>
- Green, N. L. (2021). An AI Ethics Course Highlighting Explicit Ethical Agents. In M. Fourcade, B. Kuipers, S. Lazar, D. Mulligan (Eds.), *Proceedings of the 2021 AAAI/ACM Conference on AI, Ethics, and Society*. Association for Computing Machinery (pp. 519–524). <https://doi.org/10.1145/3461702.3462552>
- Gudiño Paredes, S., Jasso Peña, F. D., & de La Fuente Alcazar, J. M. (2021). Remote proctored exams: Integrity assurance in online education?. *Distance Education*, 42, 200–218. <https://doi.org/10.1080/01587919.2021.1910495>
- Herzog, C., Leinweber, N. A., Engelhard, S. A., & Engelhard, L. H. (2022). Autonomous Ferries and Cargo Ships: Discovering Ethical Issues via a Challenge-Based Learning Approach in Higher Education. *IEEE International Symposium on Technology and Society (ISTAS)*, 2022, 1–6. <https://doi.org/10.1109/ISTAS55053.2022.10227124>
- Holmes, W., & Tuomi, I. (2022). State of the art and practice in AI in education. *European Journal of Education*, 57, 542–570. <https://doi.org/10.1111/ejed.12533>
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S. B., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2021). Ethics of AI in Education: Towards a Community-Wide Framework. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00239-1>
- Holmes, W., Bektik, D., Whitelock, D., & Woolf, B. (2018). *Ethics in AIED: Who cares?* [Workshop]. *19th International Conference on Artificial Intelligence in Education (AIED'18)*. London. Retrieved January 19, 2022, from https://oro.open.ac.uk/53443/1/AIED_2018_paper_14%20%283%29.pdf
- Hood, D., Lemaignan, S., & Dillenbourg, P. (2015). When Children Teach a Robot to Write: An Autonomous Teachable Humanoid Which Uses Simulated Handwriting. In J. Adams, W. Smart, B. Mutlu, L. Takayama (Eds.), *ACM/IEEE International Conference on Human-Robot Interaction* (pp. 83–90). Association for Computing Machinery. <https://doi.org/10.1145/2696454.2696479>
- Hudlicka, E. (2016). Virtual affective agents and therapeutic games. In D. D. Luxton (Ed.), *Artificial Intelligence in Behavioral and Mental Health Care* (pp. 81–115). Elsevier.
- Illich, I. (1971). *Deschooling society*. Harper & Row.
- Islam, Z. U., Theodorou, A., Nieves, J. C., & Dignum, V. (2019). Ethical guidelines for trustworthy AI systems. High-level expert group on artificial intelligence. *European Commission*. Retrieved January 19, 2022, from <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

- Javed, R. T., Nasir, O., Borit, M., Vanhee, L., Zea, E., Gupta, S., Vinuesa, R., & Qadir, J. (2022). Get out of the BAG! Silos in AI ethics education: unsupervised topic modeling analysis of global AI Curricula. *Journal of Artificial Intelligence Research*, 73, 933–965. <https://doi.org/10.1613/jair.1.13550>
- Johnson, W. L., & Lester, J. C. (2016). Face-to-Face Interaction with Pedagogical Agents, Twenty Years Later. *International Journal of Artificial Intelligence in Education*, 26, 25–36. <https://doi.org/10.1007/s40593-015-0065-9>
- Karaca, O., Çalışkan, S., & Demir, K. (2021). Medical artificial intelligence readiness scale for medical students (MAIRS-MS) – development, validity and reliability study. *BMC Medical Education*, 21. <https://doi.org/10.1186/s12909-021-02546-6>
- Katznelson, G., & Gerke, S. (2021). The need for health AI ethics in medical school education. *Advances in Health Sciences Education*, 26(4), 1447–1458. <https://doi.org/10.1007/s10459-021-10040-3>
- Khurana, M. P. (2020). Keeping Pace: The Need for Digital Health Education in Medical Schools. *Academic Medicine*. *Academic Medicine*, 1629–1630. <https://doi.org/10.1097/acm.00000000000003672>
- Kitto, K., & Knight, S. (2019). Practical ethics for building learning analytics. *British Journal of Educational Technology*, 50(6), 2855–2870. <https://doi.org/10.1111/bjet.12868>
- Kong, S.-C., Cheung, W.M.-Y., & Zhang, G. (2023). Evaluating an artificial intelligence literacy programme for developing university students' conceptual understanding, literacy, empowerment and ethical awareness. *Educational Technology & Society*, 26(1), 16–30.
- Lados, G., Hegedus, G., & Kovacs, Z. (2022). The Role of Identity and Ontological (In)Security in Return Migration: An Empirical Perspective from Hungary. *Journal of International Migration and Integration*, 24(2), 1–20. <https://doi.org/10.1007/s12134-022-00964-z>
- Lameras, P., & Arnab, S. (2022). Power to the Teachers: An Exploratory Review on Artificial Intelligence in Education. *Information*, 13(1), 14. <https://doi.org/10.3390/info13010014>
- Latham, A., & Goltz, S. (2019). A Survey of the General Public's Views on the Ethics of Using AI in Education. In S. Isotani et al. (Eds.), *Artificial Intelligence in Education. 20th International Conference, AIED 2019 Proceedings, Part I* (pp. 194–206). Springer. https://doi.org/10.1007/978-3-030-23204-7_17
- Lazarus, M. D., Truong, M., Douglas, P., & Selwyn, N. (2022). Artificial intelligence and clinical anatomical education: Promises and perils. *Anatomical Sciences Education*. <https://doi.org/10.1002/ase.2221>. Advanceonlinepublication.10.1002/ase.2221
- Lee, N. (2014). Digital Da Vinci: Computers in the arts and sciences. Springer. <https://doi.org/10.1007/978-1-4939-0965-0>
- Lee, I., Ali, S., Zhang, H., DiPaola, D., & Breazeal, C. (2021, March). *Developing Middle School Students' AI Literacy* [Paper presentation]. SIGCSE '21, Virtual Event, USA. <https://doi.org/10.1145/3408877.3432513>
- Leonidis, A., Korozi, M., Margetis, G., Ntoa, S., Papagiannakis, H., Antona, M., & Stephanidis, C. (2012). A Glimpse into the Ambient Classroom. *Bulletin of the IEEE Technical Committee on Learning Technology*, 14, 3–6.
- Li, Y., Li, S., & Wang, L. (2021). The Integration Development of Artificial Intelligence and Education. *Proceedings of the 16th International Conference on Computer Science & Education*, 994–997. <https://doi.org/10.1109/ICCSE51940.2021.9569551>
- Lofus, M., & Madden, M. G. (2020). A pedagogy of data and Artificial Intelligence for student subjectification. *Teaching in Higher Education*, 25(4), 456–475. <https://doi.org/10.1080/13562517.2020.1748593>
- Lucic, A., Bleeker, M., Jullien, S., Bhargav, S., & de Rijke, M. (2022). Reproducibility as a Mechanism for Teaching Fairness, Accountability, Confidentiality, and Transparency in Artificial Intelligence. *Proceedings of the AAAI Conference on Artificial Intelligence*, 36(11), 12792–12800. <https://doi.org/10.1609/aaai.v36i11.21558>
- Luo, S. (2019). Research on the change of educational management in the era of artificial intelligence. In Z. Hou (Ed.), *Proceedings of the 12th International Conference on Intelligent Computation Technology and Automation*, 442–445. IEEE Computer Society. <https://doi.org/10.1109/ICICTA49267.2019.00101>
- Marras, M., Boratto, L., Ramos, G., & Fenu, G. (2022). Equality of Learning Opportunity via Individual Fairness in Personalized Recommendations. *International Journal of Artificial Intelligence in Education*, 32, 636–684. <https://doi.org/10.1007/s40593-021-00271-1>

- Mason, J., Peoples, B. E. & Lee, J. (2020). Questioning the scope of AI standardization in learning, education, and training. *Journal of ICT Standardization*, 8(2), 107–122. <https://doi.org/10.13052/jicts2245-800X.822>
- McDonald, N., & Pan, S. (2020). Intersectional AI: A Study of How Information Science Students Think about Ethics and Their Impact. *Proceedings of the ACM on Human-Computer Interaction*, 4(2), 1–19. <https://doi.org/10.1145/3415218>
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). AI and education – Guidance for policy-makers. *United Nations Educational, Scientific and Cultural Organization*. Retrieved January 19, 2022, from <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- Miller, R., & Tuomi, I. (2022). Making the futures of AI in education: Why and how imagining the future matters. *European Journal of Education*, 57, 537–541. <https://doi.org/10.1111/ejed.12529>
- Mohammed, P. S., & Watson, E. (2019). Towards Inclusive Education in the Age of Artificial Intelligence: Perspectives, Challenges, and Opportunities. In J. Knox, Y. Wang, & M. Gallagher. (Eds.), *Artificial Intelligence and Inclusive Education Speculative Futures and Emerging Practices* (pp. 17–37). Springer. https://doi.org/10.1007/978-981-13-8161-4_2
- Mouta, A., Torrecilla Sánchez, E., & Pinto-Llorente, A. (2020). Sense of Agency in Times of Automation: A Teachers' Professional Development Proposal on the Ethical Challenges of AI Applied to Education. *Artificial Intelligence in Education*, 12164, 405–408. https://doi.org/10.1007/978-3-030-52240-7_73
- Mouta, A., Paulino, A., Ferreira, J., & Couto, F. (2015). Pedagogy as a Techné: Meaningful ICT Integration in formal Learning Scenarios. In G. Roman, J. Komuves & R. Hahn (Eds.), *Proceedings of the European Distance and E-Learning Network 2015 Annual Conference* (64–72). EDEN. <https://doi.org/10.38069/edenconf-2015-ac-0009>
- Mouta, A., Torrecilla Sánchez, E., & Pinto-Llorente, A. (2019). Blending machines, learning, and ethics. In M. González, F. Sedano, C. Llamas, F. García-Peñalvo (Eds.), *Proceedings of the Seventh International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 993–998). Association for Computing Machinery. <https://doi.org/10.1145/3362789.3362909>
- Mouta, A., Pinto-Llorente, A., & Torrecilla Sánchez, E. (2021). Blending machines, learning, sense of agency, and ethics: Designing an in-depth framework with Experts using the Delphi Method approach. In M. Alier, D. Fonseca (Eds.), *Proceedings of the Ninth International Conference on Technological Ecosystems for Enhancing Multiculturality* (pp. 665–670). Association for Computing Machinery. <https://doi.org/10.1145/3486011.3486545>
- Ng, D. T. K., Lee, M., Tan, R. J. Y., Hu, X., Downie, J. S., & Chu, S. K. W. (2023). published online 2022). A review of AI teaching and learning from 2000 to 2020. *Education and Information Technologies*, 28, 8445–8501. <https://doi.org/10.1007/s10639-022-11491-w>
- Nizhnikov, S. (2018). Foundations of ethics and education in the philosophy of transcendentalism. *Advances in Social Science, Education and Humanities Research*, 232, 604–607. <https://doi.org/10.2991/icadce-18.2018.127>
- Nye, B. D. (2015). Intelligent tutoring systems by and for the developing world: A review of trends and approaches for educational technology in a global context. *International Journal of Artificial Intelligence in Education*, 25(2), 177–203. <https://doi.org/10.1007/s40593-014-0028-6>
- Ogan, A., Yarzebinski, E., Fernández, P., & Casas, I. (2015). Cognitive tutor use in Chile: understanding classroom and lab culture. In C. Conati, N. Heffernan, A. Mitrovic & M. F. Verdejo (Eds.), *Proceedings of the 17th international conference on artificial intelligence in education (AIED), Lecture notes in artificial intelligence* (pp. 318–327). Springer International Publishing. https://doi.org/10.1007/978-3-319-19773-9_32
- Ottensbreit-Leftwich, A., Glazewski, K., Jeon, M., Jantaraweragul, K., Hmelo-Silver, C., Scribner, A., Lee, S., Mott, B., & Lester, J. (2023). published online 2022). Lessons Learned for AI Education with Elementary Students and Teachers. *International Journal of Artificial Intelligence in Education*, 33, 267–289. <https://doi.org/10.1007/s40593-022-00304-3>
- Page, M., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lahu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S. Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ*, 372 (71). <https://doi.org/10.1136/bmj.n71>
- Pammer-Schindler, V., & Rosé, C. (2022). Data-Related Ethics Issues in Technologies for Informal Professional Learning. *International Journal of Artificial Intelligence in Education*, 32, 609–635. <https://doi.org/10.1007/s40593-021-00259-x>

- Pantelimon, F.-V., Bologna, R., Toma, A., & Posedaru, B.-S. (2021). The Evolution of AI-Driven Educational Systems during the COVID-19 Pandemic. *Sustainability*, *13*, 13501. <https://doi.org/10.3390/su132313501>
- Papa, R., & Jackson, K. M. (2021). Enduring Questions, Innovative Technologies: Educational Theories Interface with AI. In K. Arai (Ed.), *Intelligent Computing - Proceedings of the 2021 Computing Conference* (pp. 725–742). Springer. https://doi.org/10.1007/978-3-030-80126-7_51
- Petticrew, M., & Roberts, H. (2008). *Systematic reviews in the social sciences: A practical guide*. John Wiley & Sons.
- Pham, P., & Wang, J. (2017). AttentiveLearner²: A Multimodal Approach for Improving MOOC Learning on Mobile Devices. In E. André, R. Baker, X. Hu, M., Rodrigo, B., du Boulay (Eds.), *Artificial Intelligence in Education. AIED 2017. Lecture Notes in Computer Science* (pp. 561–564). Springer. https://doi.org/10.1007/978-3-319-61425-0_64
- Pinkwart, N. (2016). Another 25 years of AIED? Challenges and opportunities for intelligent educational technologies of the future. *International Journal of Artificial Intelligence in Education*, *26*, 771–783. <https://doi.org/10.1007/s40593-016-0099-7>
- Prinsloo, P., & Slade, S. (2017). An elephant in the learning analytics room: the obligation to act. In M. Hatala (Ed.), *Proceedings of the Seventh International Learning Analytics & Knowledge Conference* (pp. 46–55). Association for Computing Machinery. <https://doi.org/10.1145/3027385.3027406>
- Qin, F., Li, K., & Yan, J. (2020). Understanding user trust in artificial intelligence-based educational systems: Evidence from China. *British Journal of Educational Technology*, *51*(5), 1693–1710. <https://doi.org/10.1111/bjjet.12994>
- Radford, J., Richard, G., Richard, H., & Serrurier, M. (2021). Detecting Dyslexia from Audio Records: An AI Approach. In C. Gehin, B. Wacogne, A. Fred, H. Gamboa (Eds.), *Proceedings of the 14th International Joint Conference on Biomedical Engineering Systems and Technologies – HEALTHINF* (pp. 58–66). SciTePress. <https://doi.org/10.5220/0010196000580066>
- Ramírez-Montoya, M., & García-Peñalvo, F. (2018). Co-creation and open innovation: Systematic literature review. [Co-creación e innovación abierta: Revisión sistemática de literatura]. *Comunicar*, *54*, 9–18. <https://doi.org/10.3916/C54-2018-01>
- Reisach, U. (2021). The responsibility of social media in times of societal and political manipulation. *European Journal of Operational Research*, *291*(3), 906–917. <https://doi.org/10.1016/j.ejor.2020.09.020>
- Rello, L., Williams, K., Ali, A., White, N. C., & Bigham, J.P. (2016). Dyctective: towards detecting dyslexia across languages using an online game. In L. Rello, V. Conway (Eds.), *Proceedings of the 13th Web for All Conference* (pp. 1–4). Association for Computing Machinery. <https://doi.org/10.1145/2899475.2899491>
- Renz, A., & Vladova, G. (2021). Reinvigorating the Discourse on Human-Centered Artificial Intelligence in Educational Technologies. *Technology Innovation Management Review*, *11*, 5–16. <https://doi.org/10.22215/TIMREVIEW/1438>
- Restrepo, E.G., Baldassarre, M., & G. Boticario, J. (2019). Accessibility, biases and ethics in chatbots and intelligent agents for education. In L. G. Chova, A. L. Martínez, I. C. Torres (Eds.), *EDULEARN19 Proceedings* (pp. 8824–8833). IATED Academy. <https://doi.org/10.21125/EDULEARN.2019.2196>
- Roll, I., & Wylie, R. (2016). Evolution and Revolution in Artificial Intelligence in Education. *International Journal of Artificial Intelligence in Education*, *26*, 582–599. <https://doi.org/10.1007/s40593-016-0110-3>
- Sanusi, I. T., & Olaley, S. A. (2022). An Insight into Cultural Competence and Ethics in K-12 Artificial Intelligence Education. *IEEE Global Engineering Education Conference (EDUCON)*, *2022*, 790–794. <https://doi.org/10.1109/EDUCON52537.2022.9766818>
- Schiff, D. (2021). Education for AI, not AI for Education: The Role of Education and Ethics in National AI Policy Strategies. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00270-2>
- Serholt, S., Barendregt, W., Vasalou, A., Alves-Oliveira, P., Jones, A., Petisca, S., & Paiva, A. (2017). The case of classroom robots: Teachers' deliberations on the ethical tensions. *AI & Society*, *32*(4), 613–631. <https://doi.org/10.1007/s00146-016-0667-2>
- Shih, P.-K., Lin, C.-H., Wu, L., & Yu, C.-C. (2021). Learning Ethics in AI—Teaching Non-Engineering Undergraduates through Situated Learning. *Sustainability*, *13*(7), 3718. <https://doi.org/10.3390/su13073718>

- Shum, S. B. (2018). Transitioning education's knowledge infrastructure: Shaping design or shouting from the touchline? In J. Kay, R. Luckin (Eds.), *Proceedings of the 13th International Conference of the Learning Sciences* (pp. 5). International Society of the Learning Sciences. Retrieved January 19, 2022, from <https://hdl.handle.net/10453/133232>
- Sjödén, B. (2020). When Lying, Hiding and Deceiving Promotes Learning - A Case for Augmented Intelligence with Augmented Ethics. I. I. Bittencourt, M. Cukurova, K. Muldner, R. Luckin, E. Millán (Eds.), *Artificial Intelligence in Education*, (pp. 291–295). Springer. https://doi.org/10.1007/978-3-030-52240-7_53
- Smakman, M. H. J., Konijn, E. A., Vogt, P., & Pankowska, P. (2021). Attitudes towards Social Robots in Education: Enthusiast, Practical, Troubled, Sceptic, and Mindfully Positive. *Robotics*, 10, 24. <https://doi.org/10.3390/robotics10010024>
- Srinivasan, R., & Uchino, K. (2021). The role of arts in shaping AI ethics. In A. Fokoue, C. Agunwa, K. Lee, L. T. Quigley, S. Hobson (Eds.), *CEUR Workshop Proceedings* (pp. 1–6). CEUR. Retrieved January 19, 2022, from https://ceur-ws.org/Vol-2812/RDAI-2021_paper_3.pdf
- Stein, Z., Engberg, J., & Brown, T. (2017). The seductions of quantification: Measuring human rights, gender violence, and sex trafficking. *Signs: Journal of Women in Culture and Society*, 42(2), 397–422. <https://doi.org/10.1086/690937>
- The Institute for Ethical AI in Education (2021). *The ethical framework for AI in education report*. Retrieved September 7, 2023, from <https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>
- Timms, M. J. (2016). Letting artificial intelligence in education out of the box: Educational cobots and smart classrooms. *International Journal of Artificial Intelligence in Education*, 26, 701–712. <https://doi.org/10.1007/s40593-016-0095-y>
- Trindade, Rui, & Cosme, Ariana. (2010). *Educar e aprender na Escola: Questões, perspectivas e respostas pedagógicas*. Fundação Manuel Leão.
- Tundra, E. (2020). Artificial Intelligence in Higher Education: challenges and opportunities, In L.G. Chova, A.L. Martínez, I.C. Torres (Eds.), *INTED2020 Proceedings* (pp.2041–2049). IATED Academy. <https://doi.org/10.21125/inted.2020.0644>
- Tuomi, I. (2018). *The Impact of Artificial Intelligence on Learning, Teaching, and Education. Policies for the Future*. In M. Cabrera, R. Vuorikari, Y. Punie (Eds.), Joint Research Centre (European Commission). <https://doi.org/10.2760/12297>
- Tuovinen, L., & Rohunen, A. (2021). Teaching AI ethics to engineering students: Reflections on syllabus design and teaching methods. In J. Koskinen, M.M. Rantanen, A-M. Tuikka, S. Knaapi-Junnila (Eds.), *Proceedings of the Conference on Technology Ethics* (pp. 19–33). CEUR. Retrieved January 19, 2022, from https://ceur-ws.org/Vol-3069/FP_02.pdf
- UNESCO. (2012). International standard classification of education (ISCED) 2011. *United Nations Educational, Scientific and Cultural Organization*. UNESCO. Retrieved January 19, 2022, from <https://uis.unesco.org/sites/default/files/documents/international-standard-classification-of-education-isced-2011-en.pdf>
- UNESCO. (2019a). Steering AI and advanced ICTs for knowledge societies a rights, openness, access, and multistakeholder perspective. UNESCO. Retrieved January 19, 2022, from <https://unesdoc.unesco.org/ark:/48223/pf0000372132>
- UNESCO. (2019b). Beijing consensus on artificial intelligence and education. UNESCO. Retrieved January 19, 2022, from <https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- Ungerer, L., & Slade, S. (2022). *Ethical Considerations of Artificial Intelligence in Learning Analytics in Distance Education Contexts*. https://doi.org/10.1007/978-981-19-0786-9_8
- Weber, A. (2020). Ethics concerns in artificial intelligence use in education. In L. Gómez Chova, A. López Martínez & I. Candel Torres (Eds.), *INTED2020 Proceedings* (pp. 4539–4544). IATED Academy. <https://doi.org/10.21125/inted.2020.1262>
- Weiner, N. (1960). Some moral and technical consequences of automation. science. *American Association for the Advancement of Science*, 131, 1355–1358. <https://doi.org/10.1126/science.131.3410.1355>
- Westman, S., Kauttonen, J., Klemetti, A., Korhonen, N., Manninen, M., Mononen, O., Niittymäki, S., & Paananen, H. (2021). Artificial Intelligence for Career Guidance – Current Requirements and Prospects for the Future. *IAFOR Journal of Education: Technology in Education*, 9(4), 43–62. <https://doi.org/10.22492/ije.9.4.03>

- Williams, T., Zhu, Q., & Grollman, D. (2020). An Experimental Ethics Approach to Robot Ethics Education. *Proceedings of the AAAI Conference on Artificial Intelligence*, 34(09), 13428–13435. <https://doi.org/10.1609/aaai.v34i09.7067>
- Williams, R., & Breazeal, C. (2020, July). How to train your robot: A middle school ai and ethics curriculum [session]. *International Workshop on Education in Artificial Intelligence K-12 (EDUAI '20)*, Virtual Event. Retrieved January 19, 2022, from https://dam-prod.media.mit.edu/x/2020/06/19/EduAI_2020.pdf
- Williams, R., Machado, C.V., Druga, S., Breazeal, C., & Maes, P. (2018). “My Doll Says It’s OK”: A Study of Children’s Conformity to a Talking Doll. In M. Høiseth, G.E. Baykal, A. Peter, B. Farschian (Eds.), *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 625–631). Association for Computing Machinery. <https://doi.org/10.1145/3202185.3210788>
- Xu, L. (2020). The Dilemma and Countermeasures of AI in Educational Application. In Y. Tan (Ed.), *Proceedings of 2020 4th International Conference on Computer Science and Artificial Intelligence* (pp. 289–294). <https://doi.org/10.1145/3445815.3445863>
- Yau, K. W., Chai, C. S., Chiu, T. K., Meng, H. M., King, I., & Yam, Y. (2022). A phenomenographic approach on teacher conceptions of teaching Artificial Intelligence (AI) in K-12 schools. *Education and Information Technologies*, 28, 1041–1064. <https://doi.org/10.1007/s10639-022-11161-x>
- Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education: Artificial Intelligence*, 2, 100025. <https://doi.org/10.1016/J.CAEAI.2021.100025>
- Zhang, H., Lee, I., Ali, S., DiPaola, D., Cheng, Y., & Breazeal, C. (2023). published online 2022). Integrating Ethics and Career Futures with Technical Learning to Promote AI Literacy for Middle School Students: An Exploratory Study. *International Journal of Artificial Intelligence in Education*, 33, 290–324. <https://doi.org/10.1007/s40593-022-00293-3>

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

APPENDIX B

PUBLISHED MANUSCRIPT IN EDUCATION AND INFORMATION TECHNOLOGIES

Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2024). Design of a future scenarios toolkit for an ethical implementation of Artificial Intelligence in education. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12229-y>

Journal Quality Indicators:

JCR – Q1. Journal Impact Factor (2024): 5.4. JIF Rank: 22/756. Percentile: 97.2%

SJR – Q1. SJR (2024): 1.654. H-Index: 97



Design of a future scenarios toolkit for an ethical implementation of artificial intelligence in education

Ana Mouta¹ · Eva María Torrecilla-Sánchez² · Ana María Pinto-Llorente²

Received: 6 April 2023 / Accepted: 21 September 2023
© The Author(s) 2023

Abstract

In the 1970s, research on artificial intelligence in education emerged with the aim of acknowledging and accommodating the psychological aspects of the learning process. Since then, its applications have evolved and it is now used for student learning and assessment, teachers' pedagogical practice, management of educational institutions, and lifelong learning. Nevertheless, the ethical challenges of educational programmes using these systems have not been thoroughly studied. Anchored on the theoretical frame of dialogic ethics, this paper presents a section of a participatory futures research project. The goal of the research is to develop a toolkit that educators can use to ensure a smooth and ethical transition to artificial intelligence-based education while preserving the interests of educational development. This paper emphasises the need for an informed and participatory process that involves all stakeholders and begins with an expert consultation through the Delphi method, the results of which allowed the construction of eight hypothetical futures scenarios. These scenarios provide evidence that examining the ethics of using artificial intelligence systems presents an opportunity to reflect on the ethics of education as a whole. They highlight the challenge of balancing the benefits and drawbacks of such systems, especially concerning educational goals and the interplay between diverse educational actors and personal development in educational settings. The study outcomes are intended to encourage discussions on the integration of ethical artificial intelligence in education and facilitate the continuing professional development of teachers by equipping them with scenarios that can be used as a resource for training purposes.

Keywords Artificial intelligence in education · Ethics · Delphi method · Toolkit

✉ Ana Mouta
mouta.ana@usal.es

¹ Faculty of Education, University of Salamanca, P. de Canalejas, 169, Salamanca 37008, España

² Faculty of Education, University of Salamanca, P. de Canalejas, 169, Salamanca 37008, España

Published online: 05 October 2023

Springer

1 Introduction

Although artificial intelligence in education (AIED) research emerged in the 1970s, it has evolved slowly, and it is only in the last decade that the use of these systems has seen a boom in Western countries. A recent literature review of AIED technologies from 1993 to 2020 found evidence of usage for a variety of systems targeting school management, students, teachers and lifelong learning (Zhang & Aslan, 2021). These systems seem to hold promise for education as they are able to make decisions in complex situations, update behaviour in response to environmental changes, and coexist with other systems and people in physical environments (Dignum, 2021). According to Luckin et al. (2016), for teaching and learning artificial intelligence (AI) offer is wide-ranging and encompasses equipping teachers with AI teaching assistants, the provision of personalised support for each learner and individual tutors for learners in every subject. AI can analyse vast amounts of data about each student, including their learning styles, strengths, weaknesses, and preferences. This data-driven approach allows AI systems to create personalised learning paths and recommend specific learning materials, resources, and activities that align with each student's unique requirements and pace of learning (Miao et al., 2021). AIED may also help educators understand how learners are acquiring a wide range of skills. This is made possible through embedded assessments within the learning process, timely evaluations, the ability to adapt to learners' aptitude and knowledge levels, refreshed insights into learning progress, and the identification of factors influencing learning (U.S. Department of Education, Office of Educational Technology, 2023). This AI-driven adaptive assessments offer a more dynamic and responsive process to each student's performance. Adaptability ensures that students are neither overwhelmed nor bored by the assessment, as the difficulty level matches their current proficiency. By monitoring language patterns and interactions, AI systems can also gauge a student's level of engagement and emotional state during the learning process (Luckin et al., 2016). This information can help teachers adjust their pedagogical methods to better respond to this sentiment analysis. This way educators may help students to better understand and improve emotional preparedness for the educational process.

Intelligent tutoring support can identify areas where a student is struggling and offer targeted explanations, hints, and feedback to help the student grasp the material (Wang et al., 2023). The tutoring support can take various forms, such as interactive simulations, virtual dialogues, or step-by-step problem-solving walkthroughs. Moreover, intelligent tutoring systems can learn from the interactions with students and improve their effectiveness over time. The more data the AI system collects, the better it becomes at predicting the most effective approaches for individual learners, making the tutoring experience more efficient and relevant (Khosravi et al., 2022). By analysing historical data of student performance, AI systems can make informed predictions about future learning outcomes. This enables educators to identify students who might be at risk of falling behind or excelling, allowing for timely intervention and support.

By moving beyond the traditional 'stop-and-test' approach, AI-based education has the ability to address achievement gaps, enhance teacher proficiency, mitigate teacher turnover, and alleviate areas with significant teacher shortages (Luckin et

al., 2016). AIED is also capable of offering more intelligent and timely professional development tools, while also supporting parents in their efforts to assist their child's learning. Furthermore, AI systems can continuously update learning content based on emerging trends, new research, or changing educational standards. This ensures that students have access to the most current and relevant information, providing a more up-to-date and accurate learning experience. AI-based school management systems can automate administrative tasks such as resource allocation, scheduling, admissions, timetabling, attendance and homework monitoring, and school inspections (Miao et al., 2021). They can help optimise school operations, enhance efficiency, and improve communication between stakeholders. Finally, lifelong learning companions will be available to advise, recommend, and track learning, and more flexible learning environments will allow learners to study at their preferred time and location (Luckin et al., 2016).

But at what cost to personal and social development do these systems operate? Richards and Dignum (2019) affirm the need for a systematic examination of the values and ethics that justify the use of these technologies in education, considering the pedagogical approaches they can foster and their societal impact.

1.1 Challenges of using AI in learning environments

Targeting students and embedded in these systems, pedagogical agents were conceived two decades ago to attractively simulate human-like interactions between learners and content. Positive emotions, which benefit the learning experience and academic performance, are also incorporated as affective components in artificial agents (Dobrosovetsnova & Hannibal, 2020). Nonetheless, some concerns related to affective privacy, emotion induction, and virtual relationships between a human and an agent may arise from these interactions (Hudlicka, 2016). Social robots and talking dolls also have the potential to bring about changes on children's moral development (Williams et al., 2018). The Council of Europe (2022) has reflected on some other challenges arising from the use of AI-based systems in education: there is insufficient evidence for their effectiveness, their impact on cognition is still unknown, and these technologies seem to limit not only students' but also teachers' agency. There are concerns regarding systems made for teachers, which could result in automating ineffective pedagogical practices and disempowering teachers and parents alike. Similarly, implementing AI in school administration faces several challenges that need to be addressed. Using AI to learn about students through learning analytics raises privacy concerns (mood analysis and activity logs to hit political views, ethnic identity, health or sexual orientation), safety, trust and fairness (Tundrea, 2020). Utilitarianism and deontology are challenged since data collected is not only related with the learners, but with their colleagues and even family. Another aspect to consider is the different assumptions on AIED ethics among private organisations, developers, government agencies, research centres, universities and schools (Popenici & Kerr, 2017).

1.2 Coverage of AIED ethics by research, policy and training

Given the above challenges, the present study aims to develop a toolkit of scenarios to reflect on the ethics of education in the advent of AIED that can be used in teachers' continuing professional development. Despite the considerable attention given to the ethics of general AI through numerous studies, principles, and regulations (Jobin et al., 2019; Nguyen et al., 2022), there has been no research conducted on creating tools to enhance teachers' ability to effectively utilise this technology. Furthermore, systematic educational policies for AIED are still a mirage. In fact, countries like China, India, Italy, Kenya, Malta, Singapore, South Korea, Spain and the United States are debating AIED in their policies, but only five of them include it in the context of their AI policies (Schiff, 2021). In what concerns the preparation for the use of AI in schools, capacity building focuses primarily on its technical component and it is almost only aimed at secondary and tertiary education in computer science courses. In fact, there is still a lack of teacher and parent education, and limited training opportunities for the general public (Miao et al., 2021).

Previous ethical approaches to using AIED seem to have failed to address a crucial pedagogical concern in light of the challenges posed by AI. Once AIED relates to the application of AI technologies in learning environments, its ethics must plainly consider the ethics of education. This means it should encompass teacher expectations' ethics, resource and expertise allocation, gender and ethnic biases, behaviour and discipline, accuracy and validity of assessments, knowledge quality, teacher roles, power relations between teachers and students, and particular approaches to pedagogy, such as constructivism (Holmes et al., 2021). Additionally, in this process of AI integration into classrooms, it is expected that education will continue to serve as a space for the democratic formation of public thought, language, and concepts related to social, economic, political, cultural, ethical, and caring aspects of life (Lynch, 2022).

1.3 Arguments for a participatory approach to AIED use ethics

Recognising the significance of stakeholder involvement in the design of large societal projects (Bahadorestani et al., 2020), this study strives to give educational actors a voice in establishing a secure and valuable environment when using AIED technologies. The novelty of this research topic and the limited number of project impact assessments make it particularly relevant to involve educational actors in discussions. Zuboff (2019) highlights the exceptional nature of these technological advancements that cannot be captured by current frameworks, reinforcing the importance of using participatory methods and futures approaches. Stiegler (1998) characterises technology as the science that accompanies the creation of technical objects, and he believes that its greatest strength is precisely its unpredictability. Unlike what the "robot myth" suggests, technology's dynamics are not controlled by automation, but rather by the object itself that is prone to unpredictability, making it difficult to predict its future development. However, by using "technology maieutic" experts can contribute to a constructive evaluation of the future of these systems and their applications. Therefore, it is crucial to incorporate a variety of expertise, including local and tra-

ditional knowledge and practices, in policy design, implementation, and evaluation of AIED. This approach also increases transparency, accountability, and legitimacy of decision-making (ICAT, 2020). Furthermore, regarding technologies, few research articles have examined how stakeholder engagement is considered by research teams to evaluate key characteristics of the technologies to be developed (Nygaard et al., 2021). So, this is the proposal of the research presented throughout this paper: to merge ethics with participatory, deliberative and stakeholder approaches, based on the assumption that the public can make an ethically informed assessment of a new technology and that the moral insights of various individuals involved in the creation of these systems can enhance ethical evaluations (Brey, 2017). In fact, from a pragmatic point of view, a participatory study can lead to a more diverse and comprehensive analysis; and from an ethical perspective, this study recognises the human right to be part of public decision-making processes that affect people's lives. This idea aligns with the democratic and emancipatory ideals of the modernist Enlightenment (Santos, 2012). So, the implementation of the ethics of discussion is justified because it is deemed more suitable than the Kantian dialectic in identifying practical solutions to real-world ethical dilemmas. The ethics of discussion is based on a dialogical concept of reason inspired by the "linguistic turn" of analytical philosophy. The tasks of deontological ethics must be carried out by communicative reason, embodied in an open discussion with the plurality of members of an ideal community of argumentation. This would allow for a closer connection between ethical argumentation, thought, and practical action, insofar as the social agents themselves, as ethical subjects, participate in argumentative activity and introduce various ethical contents into the discussion in their materiality (Santos, 2012).

Therefore, this study intends to answer to the following research question: How do diverse stakeholders' ethical viewpoints regarding the integration of AIED impact the shaping of potential future scenarios? How can these imagined scenarios be effectively utilised to craft a continuing professional development toolkit that supports the ongoing growth of educators, specifically in addressing the ethical dimensions tied to the utilisation of AIED? Slaughter (2020) suggests that concepts regarding the future should be incorporated into the curriculum, teacher training, and educational systems. Thus, this study will further enable conversations with educators concerning the primary themes of a training curriculum, which will aid in the facilitation of the toolkit's use. The aim of this training programme is to guarantee that AIED is used in a meaningful and ethical manner that prioritises educational and pedagogical objectives. This toolkit can also drive both risk and impact assessments that support educational policy design and implementation.

As AIED theory is still in its early years of growth, this research integrates futures studies methodologies to anticipate, monitor, and address the ethical challenges that these technologies may pose (Gidley, 2017). For more than 60 years, futures studies have evolved from a method that makes predictions to a method that questions the possible, probable, or preferred transformations and impacts of an existing object as it moves to the future (Hines, 2020). The fourth approach for conducting these investigations involves participatory action learning/research, which centres around stakeholders taking an active role in shaping their own future, drawing on their beliefs about what the future holds and what is most important to them (Inayatul-

lah, 2007). Mapping alternative futures is also a way of agency looking into the future. So, the option of creating techno-ethical scenarios is justified by the fact that they have proven to be appropriate to study moral change. They allow for an ethical analysis based on the expected future moral values of the stakeholders involved (Brey, 2017). In particular for this research, it was chosen a participatory scenario planning approach (van Notten, 2006) and most of the content is based on the Delphi consultation (Bond et al., 2021; Dinges et al., 2020; Nuwan et al., 2021). Delphi is a useful futures studies method that enables idea generation on unexplored or controversial topics by bringing together anonymous specialists from diverse regions and disciplines, allowing freedom of expression and change of opinion (Green, 2014).

2 Methodology: the Delphi method

The main objective of this study is to investigate the ethical challenges related to the integration of AIED from the perspective of multiple educational stakeholders using the Delphi technique. The study aims to develop an informed toolkit that can be utilised in continuing professional development for educators in different regions. The expertise, perspectives, and viewpoints of experts were sought to gather insights on various aspects, such as AIED technologies, applications, purposes, contexts, educational actors, subjective experience, impact on subjectification, socialisation, and qualification, as well as usage drivers, ethical concerns, and existing regulations. The ultimate goal is to provide educators with the necessary sensitivity, knowledge and resources, empowering them to participate in constructive discussions and make informed and meaningful decisions concerning the ethical integration of AIED in various educational environments.

2.1 Expert group constitution

The research coordination group consisted of three researchers and the expert group was purposefully formed through a criterion sampling method (Patton, 1990). This implied the selection of participants based on predefined criteria that focused on their substantive knowledge of the problem under study (Ogbeifun et al., 2016). Furthermore, there is controversy over the use of the term “expert” and how to appropriately identify a professional as such (Hasson et al., 2000). Therefore, in the context of this research and given the novelty of the debate on the ethics of AIED use (in research, policy and training), there was an urgent need to define consensual criteria for what an “expert” can be. Based on other Delphi studies (Arteaga-Martínez et al., 2021; García et al., 2019), proven knowledge, extensive professional experience in the field of study, and sensitivity to scientific research (grounded on previous collaborations) were selected as preferred criteria. Furthermore, the participants’ professional diversity was appreciated since it brings varying perspectives from individuals in distinct fields (Renzi & Freitas, 2015). To manage the impact of non-acceptances, a larger number of individuals was initially invited, surpassing the preferred group size (Ogbeifun et al., 2016). So, during the initial stage, 30 AIED experts were chosen from different regions, and eventually, 18 of them consented to take part. Out of the

selected participants, five individuals did not respond to the invitation, and seven faced difficulties in fully engaging with the process due to their professional commitments and ultimately declined the invitation.

The eligibility criteria selected were the following: 1. work experience in the field of technology for education (EdTech) as (a) government advocate, opinion-maker, or supplier; (b) researcher; (c) specialist in implementation and evaluation of technologies in education; and (d) specialist in EdTech development; 2. professional experience (PE) in the field of over 10 years; 3. previous collaboration with academic research (PCR); 4. perception of the self as a specialist in education with technologies or EdTech. If any of the four criteria were not met, the potential participant was deemed ineligible. For the first criterion, participants were considered eligible if they had profiles (a), (b), (c), or (d), but it was not expected for all four profiles to be present in the same participant. Criteria three and four were implemented to ensure highly qualified panel members with a high level of expertise. With this purpose, the coefficient of expert confidence ($K = \frac{1}{2} (Kc + Ka)$) was added (Almenara & Osuna, 2013; Sanromà-Giménez et al., 2021). Kc is understood as the self-assessment knowledge coefficient on the topic (on a scale of 0 to 10) multiplied by 0.1. The argumentation coefficient (Ka) was determined based on the participant's involvement in previous research (criterion #3) and their years of professional experience: 1 for more than 30 years; 0.8 for 20 to 30 years; 0.5 for 10 to 20 years. The coefficient of expert competence has been 0.73. To reduce observer bias, the data recorded in each round was analysed by multiple observers from the lead research team: the three researchers attempted to ensure interrater reliability of the collected data.

Responses to the first round came from 18 participants (100% participation rate) whose sociodemographic and occupational profiles are presented in Tables 1 and 2. The mean age of the participants is 44.5 years (SD 7.42) and 16.6% identify as women and 83.3% as men. They are employed in various continents, encompassing countries like Portugal (PT), Timor, United Arab Emirates (UAE), and United States of America (USA). All have professional experience in the field of education and hold either a Master's or PhD degree. The experts' occupational field can be grouped as follows: corporate and business (61%), academic (22%), government (11%), non-profit and community-based (6%).

Table 1 Experts' sociodemographic profile per generic professional category

Delphi Experts' Profiles								
Generic Professional Profile	No. of Experts	Age Range			Gender		Working Country	
1. EdTech government advocate, opinion-maker, supplier	5	31–40	41–50	51–60	Men	Women	PT	UAE
		2	3	0	4	1	4	1
2. Researcher on EdTech	5	31–40	41–50	51–60	Men	Women	PT	Timor
		1	2	2	4	1	4	1
3. EdTech Specialist (implementation & evaluation)	3	31–40	41–50	51–60	Men	Women	PT	USA
		1	1	1	2	1	2	1
4. Specialist in EdTech development	5	31–40	41–50	51–60	Men	Women	PT	USA
		2	2	1	5	0	3	2

Table 2 Experts' professional profile

Expert No.	Generic Professional Category	Qualification	Professional category	Occupational Field	PCR	PE	K
1	EdTech Advocate	Master in Management	EdTech broker	Corporate and Business	Yes	20–30	0,7
2	EdTech Specialist (implementation & evaluation)	Doctor in Education	Government EdTech advisor/ School Dean	Government	Yes	30–40	1
3	Researcher on EdTech	Doctor in Psychology	Researcher/ Education Lecturer	Academic	Yes	20–30	0,85
4	Specialist in EdTech development	Doctor in Psychology	CEO in Data Business Analytics	Corporate and Business	Yes	40–50	0,85
5	EdTech Specialist (implementation & evaluation)	Master in Computer Science/ Engineering	Joint venture analyst in EdTech	Corporate and Business	Yes	20–30	0,8
6	EdTech Specialist (implementation & evaluation)	Doctor in Sociology	Education programmes consultant and CPD trainer	Non-profit and community-based	Yes	10–20	0,65
7	EdTech Advocate	Master in Education	Government EdTech advisor	Government	Yes	20–30	0,8
8	Researcher on EdTech	Doctor in Anthropology	Researcher/ Anthropology Lecturer	Academic	Yes	20–30	0,75
9	EdTech Advocate	Master in Education	EdTech broker	Corporate and Business	Yes	20–30	0,75
10	Specialist in EdTech development	Master in Electrical Engineering	Software developer	Corporate and Business	Yes	10–20	0,6
11	Researcher on EdTech	Doctor in Chemistry	Researcher/ Computer Science Lecturer	Academic	Yes	30–40	1
12	Specialist in EdTech development	Master in Computer Science/ Engineering	Software developer	Corporate and Business	Yes	10–20	0,55
13	EdTech Advocate	Master in Management	EdTech broker	Corporate and Business	Yes	20–30	0,7
14	Specialist in EdTech development	Master in Civil Engineering	Software developer	Corporate and Business	Yes	20–30	0,8
15	Specialist in EdTech development	Master in Electrical Engineering	Systems engineer	Corporate and Business	Yes	20–30	0,75
16	Researcher on EdTech	Doctor in Philosophy	Researcher/ Media and Communication Lecturer	Academic	Yes	30–40	1
17	EdTech Advocate	Master in Computer Science/ Engineering	Chair of the Board at EdTech company	Corporate and Business	Yes	10–20	0,6
18	Researcher on EdTech	Doctor in Education	Pedagogy R&D in business and CPD trainer	Corporate and Business	Yes	10–20	0,7

2.2 Rounds implementation

The implementation of the Delphi method involved three iteration loops, with a synthesis facilitated by the researcher's regular feedback and the comparison of the results with informed literature (Green, 2014). The various rounds included (1) answering a questionnaire; (2) reviewing first answers and select the most important critical points for each criterion; (3) vote on the new ideas to define a final list of criteria – this one would provide the inputs for constructing hypothetical scenarios that reflect the ethical challenges AIED poses; (4) discuss the plausibility of the scenarios, rewrite them and select those that better portray the ethical challenges of AIED.

2.2.1 Iteration 1

Participants were given an 8-item questionnaire (cf. <https://forms.gle/2CqBDsyy3p2n1jpE8>) to share their knowledge, vision, and opinion on the intersection of AIED and ethics. In the process of designing the questionnaire, the research team ensured its validity by drawing insights from various sources and grounding the questionnaire items on relevant literature, specifically codes for the responsible use of AI. The Artificial Intelligence's Ethics guidelines for trustworthy AI from the European Commission (2019) and Nesta's (2019) Map of the global AI governance landscape were instrumental in shaping the questionnaire, particularly questions 4, 5, and 6. These documents shared similar principles and recommendations, encompassing AI creation, function, and outcome stages. While incorporating these recommendations, it was noted that some guidelines were broad and lacked specific guidance for practical implementation in educational settings. To address this, questions 1, 2, 3, and 7 were derived from Holmes et al.'s seminal work (2019) on the promises and implications for teaching and learning of AIED. To ensure reliability, the questionnaire underwent a pilot test with a small group of individuals similar to the target participants, including an Edtech developer, an educational researcher on ICT, and an Edtech purchaser. Their thorough review of the questions helped identify any ambiguities or misunderstandings, leading to necessary adjustments to enhance clarity. Only one minor change related to language clarity was made in the last question. Furthermore, the team employed the test-retest reliability approach to assess the stability of responses over time. Participants were asked to freely answer the questions and then answer them again after three months to evaluate the consistency of responses on separate occasions before the questionnaire was sent to the experts. After collecting the experts' 18 responses, the coordination team condensed each meaning unit to identify broader categories, and descriptive statistics were used to determine the frequency of each category.

2.2.2 Iteration 2

In the second round, the 18 participants were presented with the results and asked to rank the importance of each category based on their personal views. Of the participants, 12 (67% participation rate) submitted their responses, which is still within the recommended range of 5 to 20 experts for qualitative research on a new topic

(Landeta, 1999). While the participants were aware of all study phases, the decrease in participation rate can be attributed to the demanding nature of this research phase, which occurred during the sudden pandemic-related restrictions and uncertainty in 2020. Content that was classified as medium-high and high was incorporated at this stage (representing challenging or very challenging issues), and categories with a sum of frequencies equal to or exceeding eight (more than half of the participants' votes) were retained for the third iteration.

2.2.3 Iteration 3

A. Content Relational Analysis.

Following the tradition of merging Delphi data with current literature, the third and final round combined the collected data with “The Ethical Framework for AI in Education” (The Institute for Ethical AI in Education, 2021), created to guide the design, procurement and application of AI on behalf of learners. The goals of the ethical framework and the opportunities and challenges of AIED found by the experts were consistent and therefore interrelated by two elements of the coordination team and reviewed by two others. The cognitive mapping presented in Table 3 was used for the experts to construct hypothetical scenarios based on these ethical categories and the possible outcomes of AIED implementation in different scenarios and from the perspective of diverse educational actors.

B. Hypothetical Scenarios Construction.

Scenarios can be either normative or exploratory. Normative scenarios show ways to achieve desirable outcomes, while exploratory scenarios explore potential developments, regardless of whether they are desirable (Kosow & Gaßner, 2008). In this study, we followed the basic steps of exploratory scenario planning proposed by Dean (2019). The first step was the (1) scoping phase, which involved defining the exercise's thematic coverage, stakeholders, and timeline. The (2) information-gathering phase analysed various data sources, including updated key reports like “The Ethical Framework for AI in Education”. The (3) trend and uncertainty analysis involved analysing possible future situations in terms of their likely impact and level of uncertainty. The principal investigators of this study performed this analysis, followed by the Delphi experts in the third iteration, as further described below. In the (4) scenario-building phase, the coordination team created eight hypothetical scenarios based on the experts' input. These scenarios were designed as short exploratory vignettes that presented a difficult-to-solve dilemma, following the orthogonal construction (Wright et al., 2014) and portraying one of four situations (the horizontal axis representing the degree of impact and the vertical axis representing the degree of uncertainty). They describe potential risks that may arise while striving to achieve eight out of the nine goals outlined in “The Ethical Framework for AI in Education”. The objective of managing administration and workload was excluded from this analysis as the experts' insights regarding opportunities and challenges did not

align with this category. Table 3 provides an example of the approach taken during this phase.

All of these scenarios were thoroughly examined by the experts. This exercise, which involves using key criteria to assess scenario quality, has a long tradition (Greeuw et al., 2000; Kreibich, 2007). Although scenarios are of a hypothetical nature, they are by no means arbitrary and must be evaluated according to criteria such as plausibility, consistency, comprehensibility and traceability, distinctness, transparency, degree of integration, and quality of reception (Kosow & Gaßner, 2008). In this study, participants were asked to give feedback on each scenario based on five criteria: (1) plausibility – whether it seems possible, (2) consistency – whether it makes sense logically, (3) comprehensibility – whether it is easy to understand, (4) relevance – whether it is relevant, and (5) distinctiveness – whether it is different from the others.

3 Results per round

3.1 Iteration 1

For each category created upon the participants' responses, frequencies were determined by descriptive statistics (cf. Table 4).

3.2 Iteration 2

Regarding the future of AIED, some of the initially proposed categories have been excluded. This happened with the items “Employment” and “Relational and Societal Factors”, which look at how education relates to impacts on other specific social layers. This appears to be true for both the negative perspectives – “Loss of interaction & detachment” – and the positive ones – “More humanistic causes, leisure and culture” (both from the category “Relational and Societal Factors”) or “Better matching people-education-jobs” (category: “Employment”). The most extreme views concentrate a smaller number of votes: for example, “Education dissolution” (3 votes for medium-high and none for high) or “Schools dissolution” (5 votes combining medium-high and high). The highest rated answers correspond to questions more directly related to academic achievement and the improvement of didactic resources: “Knowledge Management & Share”, “Learning Processes” and “Skills enhancement”. Looking at the critical positive/negative spectrum, four dimensions stand out positively, meaning they can have a noteworthy impact: “Instant data uploads on any topic” and “Personalised/Smart Learning Platforms and MOOC's” (both from the category “Knowledge Management & Share”), “Real time engagement and performance assessment and feedback” (category: “Learning Processes”), “Enhanced high-level processing” (category: “Skills enhancement”). All of them related to improving academic performance. On the negative or more challenging side, four dimensions stood out: “Education as a business”, “Information as commodity”, “Mainstream thinking and standardised behaviours”, and “Larger learning divide”. All of these points belong to the “Broader Implications” category, showing that these experts

Table 3 Delphi third round: tables with hypothetical scenarios sent to experts

SCENARIO 1. Objective from “The Ethical Framework for AI in Education”: Achieving Educational Goals
AI should be used to achieve well-defined educational goals based on strong societal, educational or scientific evidence that this is for the benefit of learner.

Risks	AIED Targets	AIED Function	AIED Outcomes	Schools support
Future with AIED	School Administrators and Government	Content Production	Education Quality	Product Procurement
Broader Implications	Responsibility for frameworks and assessment	Accountability for content selection and curation	Accelerated Individuation	Framework definition for procurement
Information as commodity			Reproduction of trainable skills	
Mainstream thinking and standardised behaviours			Human-human Interaction	
Knowledge Management & Share			Challenge in terms of emotions and deep meaning	
Instant data uploads on any topic				
Personalised/Smart Learning Platforms and MOOC's				
Skills enhancement				
Enhanced high-level processing				

Table 3 (continued)

SCENARIO 1. Objective from "The Ethical Framework for AI in Education": Achieving Educational Goals <i>-AI should be used to achieve well-defined educational goals based on strong societal, educational or scientific evidence that this is for the benefit of learner.</i>	AIED Targets	AIED Function	AIED Outcomes	Schools support
Risks				
Future with AIED				
Hypothetical Scenario: May, 2033. Chloe is wakened by her robot friend Roy. Chloe checks her personalised learning chronogram for the week and starts her learning day in her garden. She activates her mindfulness chip and breathes in deeply. The program suggested a Steve Jobs' hologram on entrepreneurship for the morning, followed by a virtual meeting with John and Mitchell to discuss best practices to run a sustainable out-of-the-earth business ten years in the future. When they meet, they quickly benchmark what has already been done in the field, using their extended brains (what once was Internet is now connected to neural processes through implanted chips). Her backpack was prepared by Roy for the afternoon, with lots of sensors and some acrylic paints. Chloe creates a handmade painting that is able to provide multiple experiences through sensors: for example, if a viewer were to touch the sun, they might feel heat. All her academic performances are automatically graded on 10 parameters, including engagement, creativity, critical thinking, and empathy. She will only have one brick-and-mortar learning experience this week with ten students – five of them will meet on that day for the first time. The capacity to build learning opportunities with new colleagues is a highly appreciated skill set. She was given the names of the new colleagues and after receiving her teachers' feedback along with the automated grading, she is ready for her favourite part of the day. She automatically screens all the new colleagues' social media profiles through her favourite filters: dating history (longer and shorter relationships), trips to other countries, preferred clothing styles, and personal physical enhancements already completed. Each one immediately receives a score that matches Chloe's already recorded preferences. She celebrates with Roy the fact that she found a 17-year-old boy who had a particularly high matching score. If nothing else, at least academic compatibility and empathy will be certain. She can hardly wait for that face-to-face learning day!				

seem to agree on more global negative impacts of using AIED, namely in terms of politics and asymmetries related to (the quality of) access to education, its instrumentalization for profit and the emergence of dominant standardised attitudes.

3.3 Iteration 3

To assess and redesign the eight scenarios, the experts' feedback was analysed and broken down into meaningful parts. For instance, one expert raised a consistency question about the 2nd scenario related to evaluation: "[isn't there] more fear of correcting errors immediately or wanting to respond to the review? Do you think that could happen? The immediacy can trigger action and in some cases trigger fear" (Expert No.3). These inputs were considered and the scenarios updated accordingly. Five scenarios were evaluated based on how easy they were to understand. On the positive side, six scenarios were considered highly plausible, and five were deemed relevant. However, in terms of plausibility, five scenarios were less convincing (scenarios 1, 3, 4, 6, and 8). Parts of five out of the eight scenarios were neither believable (scenarios 1, 2, 4, 5, and 8) nor consistent (scenarios 1, 2, 3, 5, and 7). The fifth scenario, which concerns privacy, raised concerns about potential opposition during teacher training. The experts recommended clarifying the distinctions between scenarios 3 and 5 and proposed a dialogue between scenarios 1 and 8. Scenarios 1 and 5 underwent significant changes. The last three scenarios (6, 7, and 8) received fewer comments, either because they were more consistent or because the experts were less likely to provide feedback after initial involvement. In fact, the literature suggests that the optimal number of scenarios should be between two and four or five for ease of manageability (Dean, 2019).

In summary, the experts' recommendations were focused on several areas, including: (1) increasing the trade-offs between good and bad outcomes in the scenarios; (2) adding more biographical details to the characters; (3) setting some of the dates further in the future; and (4) creating clearer distinctions between scenarios that examine fairness and the preservation of a privileged status by private schools. Additionally, two participants predicted that some scenarios may be difficult for teachers to discuss due to their futuristic nature or because they portray teachers as passive. Regarding comprehensibility, the experts' understanding of future outcomes and trends resulting from AIED growth is depicted in Fig. 1. Certain factors are emerging as trends (low uncertainty and high impact) in the economic (E), political (P), social (S), and technological (T) domains. Other aspects reflect critical uncertainties (high uncertainty and high impact), primarily in the social (S) domain. Examples of these critical uncertainties include the dominance of classification and labelling systems, the gamification of human experience, AI's influence over emotional expression, the power of corporations as a civilisational threat, the importance of human relationships, the possibility of parental rejection of AI without proper ethical oversight, and the need to preserve students' aesthetic development in the highly technical global context.

After analysing all the data, a final list of future scenarios was generated (cf. https://drive.google.com/file/d/1o6ToayuZ80Knj4R6QsWBbL7NW0aAm7/view?usp=share_link). This list will serve as a comprehensive resource for continuing professional development with teachers.

Education and Information Technologies

Table 4 Delphi first round results: frequencies per category

	FREQ
1. If you were asked to forecast the future (academic/social/emotional data mapping for employability, online learning, impact and ROI...) of Education with Artificial Intelligence what would you highlight?	
Broader Implications	
Education as business	1
Education dissolution	1
Information as commodity	1
Mainstream thinking and standardised behaviours	2
Larger learning divide	1
School dissolution	1
Trade-off: Privacy	1
Employment	
Better matching people-education-jobs	4
Knowledge Management & Share	
Instant data uploads on any topic	2
Personalised/Smart Learning Platforms and MOOC's	3
Learning Processes	
Better teacher-student relationships	1
Learning personalisation	1
Learning practices impact & predictability	2
Real time engagement and performance assessment and feedback	2
Teacher 's routine tasks facilitation	2
Relational and Societal Factors	
Loss of Interaction & Detachment	2
More humanistic causes, leisure and culture	1
Skills enhancement	
Enhanced high-level processing	4
Strengthened role of emotional intelligence and social skills	2
2. From the applications of AI in formal Education (automated assessment or feedback, curated content, learning personalisation, virtual teaching assistance...), which of them should encompass a particular ethically sound design?	FREQ
Automated Assessment/Feedback	
Integration of cognitive, social, emotional and psychological inputs into automated assessment	7
Curated Content	
Consider bias, limits, and manipulation	2
Learning Personalisation	
Constraint of personal freedom and capacity of thought	2
Misplacement of personal singularities	2
Privacy (publishers, vendors, schools, students)	2
Virtual Teaching Assistance	
Respect for students' self-determination and autonomy in virtual teaching	2
3. Regarding the learner, the educator or the school institution, which AI target would be of more ethical concern to you?	FREQ
Learner	
Education systems not prepared to respond to kids will to know more about AI	1
Performance-oriented solutions - avoidance of kids' "bad" emotions (frustration, boredom...) and refusal of its developmental insight	2
Students' exposure to AI conception shortcomings	2
Students' vulnerability to abusive power	6

Table 4 (continued)

Schools	
Data processing, encryption and storage	2
Educational actors training	2
Protection of school as a pre-technological space for thinking and imagining	1
School Administrators and Government	
Responsibility for frameworks and assessment	1
Teachers	
Preparation and training	3
Risk of replacement	1
4. When thinking of AI applied to Education, if you had accountability for AI creation which issues (accountability, biases, diversity representation, explainability, fairness...) would you be more ethically apprehensive about?	FREQ
Accountability	
Aligning industry's goals with education systems and individuals	5
Reinforcing accountability first to enable AI development	1
Diversity representation	
Inclusiveness, equity and fairness	8
Representation of local strengths and needs	1
Technological and ontological diversity	1
Explainability	
Fostering critical users	1
Enhancing transparent processes despite interests	2
5. When thinking of AI applied to Education, if you had accountability for AI function which issues (cyber security, decision making processes, hacking, privacy control...) would you be more ethically apprehensive about?	FREQ
Cyber Security	
Hacking and exploitation	5
Privacy control	6
Unauthorised use of personal information	3
Content Production	
Accountability for content selection and curation	4
6. When thinking of AI applied to Education, if you had accountability for AI outcome which issues (automation, citizenship, human-human interaction, labour changes...) would you be more ethically apprehensive about?	FREQ
Automation	
Blindfolded processes	4
Education Quality	
Accelerated Individuation	1
Reproduction of trainable skills	1
Human-human Interaction	
Challenge in terms of emotions and deep meaning	3
Danger of virtual interaction replication	1
Distance between people	1
Interaction among community members	1
Political Citizenship	
Vulnerability for those who have lower self-determination	3
Training and labour changes	
Demanding labour market	2
Students' academic and career development	1

Education and Information Technologies

Table 4 (continued)

Vocational Training threatened	1
7. In terms of school support (product procurement, students' privacy, teachers preparation...), which areas would you consider of priority care?	FREQ
Educational actors	
Communities' awareness and preparation	1
Involvement of all educational actors in AI rules definition	1
Product Procurement	
Framework definition for procurement	2
Sorted by internal school stakeholders	2
Teachers Preparation	
Capable of guiding students in tech usefulness	1
Greater AI literacy, awareness, and questioning	10
8. Considering the topics under study, please add any note or commentary, if you wish to.	FREQ
AI and its applications cannot be understood as all the same	1
AI cannot replace Education as human relation	1
AI outcomes are political matters	1
Philosophical conundrum of ethics: utilitarian position ("the most good for the most people")	1
VS. Kantian universal ethics	
Schools should not be wired	1
Tech innovation should be an act of imagination exerted by all	1

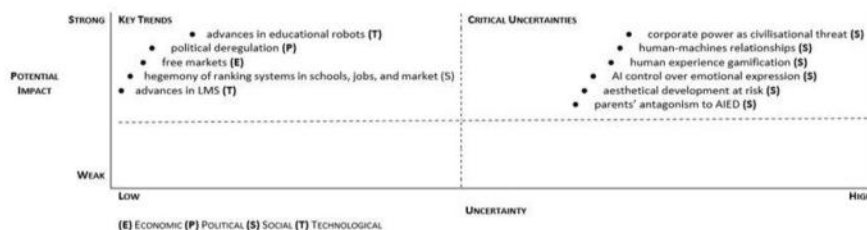


Fig. 1 Impact and Uncertainty grid for AIED futures scenarios

4 Discussion

As Holmes et al. (2021) suggest, the ethics of AIED is expected to call our attention to the ethics of education in the first place. Each new technology entering the realm of education is, in fact, an opportunity to rethink education ethics and how particular technological features may hinder specific aspects of pedagogy. Such challenging systems present an opportunity for schools to discuss and define their ethical common ground and to develop strategies to overcome any obstacles that may arise. Collaborating in this way can be helpful to identify and solve unexpected problems that may emerge while using AIED, problems that were not taken into account during the design process. So, which aspects of pedagogy need to be safeguarded by ethical considerations? By using the “ik.model” (Mouta et al., 2015), which is designed to assess how technology is integrated into education while prioritising educational goals, it is possible to grasp the potential risks and benefits of

having such an autonomous agent incorporated into pedagogy. First of all, it can be argued that the majority of ethical preoccupations of using AI in education were found in the “relational dimension” (the social domain of the impact and uncertainty grid), highlighting the ethical importance of how people interact with one another in the advent of AIED. The designed scenarios acknowledge the importance of shared values, including both explicit and implicit educational agreements. They also prioritise stakeholder involvement in decision-making processes, while being mindful of disparities in technological access and pedagogical quality. The first scenario centred around achieving educational goals and all participants agreed that the use of AI in education must be backed by robust evidence that demonstrate its beneficial effect on learners, what was previously highlighted by research (Richards & Dignum, 2019). The third scenario placed an emphasis on equity and highlighted the risk of a wider learning gap, where the disparities between public and private schools could become more prominent. This disparity extends to significant gaps between developed and developing nations, socioeconomic groups within countries, and those who have AI-enhanced jobs versus those who are susceptible to being replaced by them (Miao et al., 2021). The seventh scenario addresses informed engagement and recommends that students and other education actors should possess an adequate understanding of AI and its implications. The experts suggest that individuals with AIED knowledge and the ability to question should participate in establishing AI policies at the school level. Levinas (1969) contended that ethics must primarily acknowledge the importance of the interpersonal dimension, when in search of its existential ground. And this comes before any consideration of concepts like utility, virtue, or duty. Thus, it is crucial that the design, implementation (education with AIED), and evaluation (educational results) of activities with AIED be collaborative and shaped by people who have the capacity to consider the individual and societal benefits and drawbacks of its adoption and governance.

When designing these systems (“technological dimension”), it is important to be cautious about how the technologies are built, considering the type of connections people may form with machines. The sixth scenario emphasises the need for transparency and accountability in overseeing the operation of AI systems. In scenario eight, the importance of involving individuals who understand the potential consequences of AI on individuals and society in the design of these technologies is highlighted. This situation is exemplified by social robots that interact with humans. In fact, a review of the literature on the use of educational robots has evaluated their impact on four main dimensions, which are expected to be carefully scrutinised: (1) privacy; (2) human replacement; (3) impact on students; and (4) accountability (Serholt et al., 2017). To meet this need, it is encouraged to use a Trustworthy AI Ethics Guide in both creating and utilising AI technologies (European Commission, Directorate-General for Communications Networks, Content and Technology, 2019) and also to promote a participatory design of these technologies, informed by a variety of educational stakeholders and research fields.

Considering “content knowledge” and implementation through the lens of Levinas, education can be viewed as an ethical practice that aims to create spaces where individuals can engage with one another in a caring manner. The fourth scenario demonstrates how AI systems can undermine student autonomy, disregarding even

the most capable and perceptive students. Dependence on automated decisions and AI-driven personalisation can limit opportunities for student interaction and focus on knowledge that is easier to automate, hindering their development of resourcefulness, self-efficacy, self-regulation (Miao et al., 2021), and the recognition of themselves as the citizens they already are. Moral deskilling can also affect educators, who increasingly rely on AI machines to make decisions and become less critical and morally engaged (Tundrea, 2020). The fifth scenario focuses on privacy and the use of personal data to achieve educational goals. Experts warned of the possibility of education becoming a business with AIED being used, with many opportunities to enhance teacher training but at the expense of privacy and with the main goal of providing a specific service (Pammer-Schindler & Rosé, 2021).

In the dataism era, another ethical concern of AI-based education relates to the possibility of turning individuals into measurable and controllable entities through digital experiences. According to Han's (2014) argument that dataism could reduce self-tracking to mere self-surveillance, it's crucial to foster collaboration between teachers and students to envision and establish desirable futures with this unprecedented level of access to data. This is an invitation to reflect on what it means to be an individual in a group, and to foster mutual growth through reciprocal interactions. Educators also have the responsibility to unpack with their students the onto-epistemic grammar of dataism. This ethical undertaking involves exploring the anthropocentric perspective (Andreotti et al., 2015) underlying this desire, as well as the drive for ontological security (Lados et al., 2022) and the thirst for absolute knowability (Stein et al., 2017). This also provides a chance to use pedagogical strategies for a deeply purposeful and ethical learning experience. Project-based learning and curriculum infusion can be powerful strategies for achieving integrated goals. By incorporating other ethical, societal, and political concerns from different fields' perspectives, these approaches can address the challenges posed by systems that can grow in agency through our own inputs, while still meeting curriculum standards. Educators can use a variety of subjects and make students apply them to the task, while also availing themselves of AI resources. Engaging in discussions about AI functions from the perspectives of different subjects such as Mathematics, Science, History, and Languages (or any other) can serve as a means of strengthening newly acquired knowledge in these areas, applying it to practical and analytical tasks, and simultaneously building a more varied and intricate understanding of the AI systems in question. By doing so, education can move toward a more ethical exercise of freedom, even in the face of digital pressure.

Considering evaluation and "learning processes" dimension, AIED can provide just-in-time assessments, as well as new insights into how learning is progressing. But before recognising the potential benefits of incorporating AI-based assessment into learning environments, it is necessary to address ethical concerns related to educational assessment. While it is true that obtaining high-quality knowledge is extremely valued and that AI can improve the processes of encoding, storage, and retrieval by offering personalised pathways, discussions with education experts indicate that this can only be achieved if there is mutual agreement and respect between individuals and what to expect from their interaction with autonomous systems. This is even more crucial now, as the pandemic caused by Covid-19 has given new impe-

tus to technology (García-Peñalvo et al., 2021). As exam proctoring in some regions was a response to the problem of not being able to test students in physical situations, AI was identified as a possible solution to a large number of educational challenges. These ethical concerns were directed towards the second scenario that focused on forms of assessment. The experts believed that automated assessment and feedback on cognitive, social and emotional performance could become a reality in the near future and that this could present challenges and potential risks. Earlier research suggested that supervision is effective in decreasing deceitful actions. However, students may only behave honestly because they know they are being watched and not because of any intrinsic drive or self-reflection. This can lead to feelings of discomfort, such as a lack of privacy and anxiety, during the assessment process (Gudiño Paredes et al., 2021).

Drawing from Hannah Arendt's ideas, Coulter and Wiens (2002) suggest that in order to make sound educational evaluations it is essential to establish a connection between the teacher (actor) and the researcher (spectator). It is critical to challenge teachers to become accurate judges and actors themselves, which involves creating opportunities for them to appear. In fact, this is the goal and major ethical responsibility of this research: to engage teachers in the development of the curriculum for a course on AIED as part of their continuing professional development, using this scenarios toolkit as a basis for discussion. This represents an effort to urge teachers to become judging actors, which constitutes both a moral-political and an educational issue. These teachers are expected to engage with each individual child in complex communities, balancing guidance and agency and encouraging children to make informed judgments about the actions of others, and reflecting on their own actions and choices. In order to expect teachers to foster these skills, they must receive training in these very principles. Furthermore, the process of using AI systems to evaluate performance presents a challenge: everyone involved should be asked to participate in understanding the feedback given. When personal perceptions of performance differ from the classifications provided by the AI system, it can be both a valuable opportunity for personal growth (through insight) and understanding of how AIED works. This approach will allow for constructive criticism and questioning, forming the foundation for critical engagement with the world.

It's important to recognise that AI is more than just a neutral tool; it's an agent that learns, interacts, and can impact outcomes, which can create conflicts between students, teachers, and the educational system in terms of agency. While AI-powered chatbots and virtual assistants can provide students with 24/7 support and resources, thereby increasing their autonomy, there are also risks associated with AIED that could undermine this. For instance, if AI is used to make decisions such as determining which courses learners should take or what career path they should pursue, it could limit their options and opportunities for self-determination. This, in turn, may restrict their ability to explore their identity (which is crucial for psychosocial development) and form a sense of self. Similarly, if AI is used to monitor student behaviour or performance, it could lead to a surveillance culture that restricts students' ability to take risks and make mistakes, which are deemed essential for growth. Furthermore, the Vygotskian notion of 'scaffolding,' which involves a skilled mentor providing guidance and encouragement for action, may be interpreted differently in the context

of AI. Since AI may not be able to offer the same level of support and encouragement as a human mentor, it could alter the perception of the teacher's role, who is expected to provide challenging emotional experiences that are crucial for confidently engaging with the world.

To conclude, this toolkit aims to bring intervention by providing educators with a comprehensive set of resources and guidelines, enabling them to effectively address pedagogical challenges, including the integration of AIED as supportive tools when appropriate, considering ethical aspects and potential challenges. Designed for use in educator continuing professional development, the toolkit will consist of training modules and workshops covering fundamental AI concepts and their applications in education, how AI can address current educational needs, the benefits and challenges of AI implementation in the classroom, and the contextual integration of AI in various settings. It will also foster participatory and collective agency and decision-making among educational stakeholders to define ethical and pedagogical aspects of AIED implementation that better suit their educational contexts and interests. Additionally, this toolkit can be effectively employed as a scenario-based learning tool for students in project or inquiry-based learning, encouraging exploration of real-life situations and challenges that arise when using AI in the classroom, thereby empowering students and enhancing agency within the school environment. Finally, it can serve as a valuable resource for developers, providers, and educational decision-makers by offering guidance on ethical considerations related to AI usage in education. Within the scope of the present study, the subsequent stage will entail the concrete development and execution of a continuing professional development curriculum designed for educators. In this phase, a close partnership with teachers will be established to initiate the pilot testing of the toolkit within different educational environments. This methodological approach aims to elicit relevant insights, refine the toolkit's operational effectiveness, and systematically evaluate its influence on educators' pedagogical approaches and students' ethical learning experiences pertaining to AIED.

Some limitations of this study should now be highlighted. While this study's methodology may seem uncertain and speculative, it does not rely on predictive analysis, but rather on plausible or possible futures (Brey, 2017). One strength of the study is the use of short-term future narratives and providing information on the potential and dependencies of emerging technologies, which helps to bolster the decisions made throughout the work (Brey, 2017). Nevertheless, the study faces certain constraints, including the difficulty of conducting research grounded in objective moral reasoning, ensuring fairness, mitigating unequal power dynamics, and fostering equal participation (Hagendijk & Irwin, 2006). To mitigate these challenges, the Delphi method was used. Another limitation that should be acknowledged has to do with the fact this study involved a relatively small number of expert participants ($n=18$), which may raise concerns about the representativeness of the insights gathered. There is a risk that some perspectives or expertise relevant to the topic might be underrepresented, leading to conclusions that might not fully capture the complexity and nuances of the ethical challenges related to the integration of AIED in education. To address the limitations of the sample size, the research coordination team made a deliberate attempt to include a diverse group of experts from different regions, including Europe, Southeast Asia, Middle East, and North America. These experts possess var-

ied backgrounds and expertise in the field of AIED. There was also a concern to ensure that the eligibility criteria encompassed a range of professional profiles, such as government advocates, researchers, specialists in implementation, and specialists in EdTech development. Additionally, the Delphi method involves multiple iteration loops and expert feedback, meaning that data collection continues until a point of saturation is reached, where new insights or themes are no longer emerging from the panel. With 18 participants, it seemed possibly to efficiently reach this point, allowing for an in-depth exploration of the research questions. In the continuation of this research, qualitative research methods will be integrated, including focus groups with educators. These methods will complement the Delphi method, providing deeper insights into participants' perspectives and experiences.

5 Conclusions

This paper reports on a study that analysed an expert consultation on AIED. The goal of the study was to foster debate on ethical AI integration in education and support teachers' continuing professional development through scenarios that will serve as a toolkit for discussing training syllabi. The scenarios created feature a combination of current AIED technologies and some dystopian elements. They highlight how these systems may significantly impact our daily lives, interactions, thoughts, and emotions, being reasonable to expect that there may be significant challenges that arise at a societal or even civilisational level. Therefore, it's important for educators to be mindful of the potential risks and benefits of using AIED, particularly with regard to emotion recognition and social choice, and collaboratively establish purpose for its use. This means knowing fundamental characteristics, potentialities, and challenges of AI, including its general functions, and being transparent about how AI is being used. It also means involving students in decisions about how these technologies are being implemented, how its inputs are incorporated into the learning experience, acknowledging that AI should be used to support student agency.

The scenario toolkit created will serve as the foundation for conducting focus group discussions with educators, with the aim of anticipating the challenges and aligning educational objectives and practices with the context in which these AI technologies are intended to be employed. In this phase it will be determined how to integrate the data gathered into training programmes that promote the ethical use of AIED, while taking into account the diverse access, availability, and implementation scenarios across various regions. During these discussions, experts advised exercising caution, as teachers may feel uncomfortable discussing unfamiliar topics, causing the conversation to steer towards familiar territory. In the medium term, the goal of this research will be to equip educators with the appropriate resources to participate in such discussions, preventing resistance and fostering constructive dialogue that enhances the overall discourse.

The decision to employ a participatory method was taken to obtain a more comprehensive perspective on the ethical challenges of AIED implementation. Being the initial phase of a research project that will subsequently involve educators, this first step aimed to stimulate increased involvement from pertinent stakeholders who could

influence policymaking. The experts' unique considerations may already enable them to contribute to unprecedented critical evaluation of AI technologies' impact in education as responsible actors in the field. In addition, the methodology adopted in this study aims to conform to ethical principles, active participation and agency, which are the exactly same criteria proposed for the assessment of AIED. Not only does this methodology aim to ensure ethical research practices, but it also seeks to instil and promote the values it supports in the individuals involved, thereby guaranteeing their application in the way learners are encouraged to develop in the presence of AI-based systems.

Acknowledgements A portion of the data presented in this article was first introduced in a previous publication, "Blending machines, learning, sense of agency, and ethics: Designing an in-depth framework with experts using the Delphi Method approach" (<https://doi.org/10.1145/3486011.3486545>). This earlier paper provided a broad overview of the first author's PhD thesis, covering only the initial two iterations of the Delphi process and did not include the third iteration, which forms the core output of this current research. The present article offers a toolkit featuring eight AIED ethics scenarios, along with a more suitable introduction to the field of research, and a thoroughly description of the methodology used. In addition, it includes comprehensive sections for discussing and drawing conclusions from the research. In contrast, the previous version was primarily intended for academic purposes and formed a part of a PhD consortium, not providing a complete account of the research work.

Authors' contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ana Mouta, Eva María Torrecilla-Sánchez and Ana María Pinto-Llorente. The first draft of the manuscript was written by Ana Mouta and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding The research leading to these results received funding from Fundação para a Ciência e Tecnologia under Grant Agreement No 2020.06258.BD. Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature.

Data Availability The datasets generated during and/or analysed during the current study are not publicly available to protect study participant privacy, but are available from the corresponding author on reasonable request.

Declarations

Competing interests All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Almenara, J. C., & Osuna, J. B. (2013). La utilización del juicio de experto para la evaluación de TIC: el coeficiente de competencia experta. *Bordón*, 65(2), 25–38. <https://doi.org/10.13042/brp.2013.65202>.
- Andreotti, V., Stein, S., Ahenakew, C., & Hunt, D. (2015). Mapping interpretations of decolonization in the context of higher education. *Decolonization: Indigeneity Education & Society*, 4(1), 21–40.
- Arteaga-Martínez, B., Macías-Sánchez, J., Pla-Castells, M., & Ramírez-García, M. (2021). Design and validation of an observation instrument for mathematics classes in secondary education: Expert panel and Delphi process. *RELIEVE*, 27(2), art3. <https://doi.org/10.30827/relieve.v27i2.21812>.
- Bahadorestani, A., Naderpajouh, N., & Sadiq, R. (2020). Planning for sustainable stakeholder engagement based on the assessment of conflicting interests in projects. *Journal of Cleaner Production*, 242(46), <https://doi.org/10.1016/j.jclepro.2019.118402>.
- Bond, M., Marín, V. I., & Bedenlier, S. (2021). International collaboration in the field of Educational Research: A Delphi Study. *Journal of new Approaches in Education Research*, 10(2), 190–213. <https://doi.org/10.7821/naer.2021.7.614>.
- Brey, P. (2017). Ethics of Emerging Technologies. In S. O. Hansson (Ed.), *Methods for the Ethics of Technology*. Rowman and Littlefield International.
- Coulter, D., & Wiens, J. R. (2002). Educational Judgment: Linking the actor and the Spectator. *Educational Researcher*, 31(4), 15–25. <https://doi.org/10.3102/0013189X031004015>.
- Council of Europe (2022). Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law. <https://rm.coe.int/prems-092922-gbr-2517-ai-and-education-txt-16x24-web/1680a956e3>.
- Dean, M. (2019). Scenario Planning: A Literature Review. <https://doi.org/10.13140/RG.2.2.12629.24802>.
- U.S. Department of Education, Office of Educational Technology (2023). *Artificial Intelligence and Future of Teaching and Learning: Insights and Recommendations*. Washington, DC.
- Dignum, V. (2021). The role and challenges of education for responsible AI. *London Review of Education*, 19(1), 1, 1–11. <https://doi.org/10.14324/LRE.19.1.01>.
- Dinges, M., Wang, A., & Schuch, K. (2020). Using the Delphi method in evaluations: Incorporating a future oriented perspective in evaluations. *Fteval Journal for Research and Technology Policy Evaluation*, 50, 44–50. <https://doi.org/10.22163/fteval.2020.469>.
- Dobrosovetsnova, A., & Hannibal, G. (2020). Teachers' Disappointment: Theoretical Perspective on the Inclusion of Ambivalent Emotions in Human-Robot Interactions in Education. In T. Belpaeme & J. Young (Eds.), *Proceedings of the 2020 ACM/IEEE International Conference on Human-Robot Interaction* (pp. 471–480). ACM. <https://doi.org/10.1145/3319502.3374816>.
- European, & Commission, Directorate-General for Communications Networks, Content and Technology, (2019). Ethics guidelines for trustworthy AI. <https://data.europa.eu/doi/10.2759/346720>.
- García, I. A., Moreno, M. F., Prado, G. R., Sáez, B., Puertas, M. A., & Gómez, A. A. (2019). Validación de un cuestionario sobre actitudes y práctica de actividad física y otros hábitos saludables mediante el método Delphi. *Revista Española De Salud Pública*, 93. <https://scielo.isciii.es/pdf/resp/v93/1135-5727-resp-93-e201909081.pdf>.
- García-Peñalvo, F. J., Corell, A., Abella-García, V., & Grande-De-Prado, M. (2021). Recommendations for Mandatory Online Assessment in Higher Education During the COVID-19 Pandemic. In D. Burgos, A. Tlili, & A. Tabacco (Eds.), *Radical Solutions for Education in a Crisis Context. COVID-19 as an Opportunity for Global Learning* (pp. 85–89). Springer Nature. https://doi.org/10.1007/978-981-15-7869-4_6.
- Gidley, J. M. (2017). *The future: A very short introduction*. Oxford University Press.
- Green, R. (2014). The Delphi technique in Educational Research. *SAGE Open*, 4(2), 1–8. <https://doi.org/10.1177/2158244014529773>.
- Greeuw, S. C. H., Asselt, M. B. A., Grosskurth, J., Storms, C., Rijkens-Klomp, N., Rothman, D. S., & Rotmans, J. (2000). *Cloudy Crystal Balls: An assessment of recent european and global scenario studies and models*. European Environment Agency. https://www.eea.europa.eu/publications/Environmental_issues_series_17.
- Gudiño Paredes, S., Jasso Peña, F. D., & de La Alcazar, F., J.M (2021). Remote proctored exams: Integrity assurance in online education? *Distance Education*, 42(2), 200–218. <https://doi.org/10.1080/01587919.2021.1910495>.
- Hagendijk, R., & Irwin, I. (2006). Public deliberation and governance: Engaging with Science and Technology in Contemporary Europe. *Minerva*, 44(2), 167–184. <https://doi.org/10.1007/s11024-006-0012-x>.

Education and Information Technologies

- Han, B. C. (2014). In M. S. Pereira (Ed.), *Psicopolítica*. Trans.). Relógio D'Água. (Original work published 2014).
- Hasson, F., Keeney, S., & McKenna, H. (2000). Research guidelines for the Delphi Survey technique. *Journal of Advanced Nursing*, 32, 1008–1015. <https://doi.org/10.1046/j.1365-2648.2000.t01-1-01567.x>.
- Hines, A. (2020). When did it start? Origin of the foresight field. *World Futures Review*, 12(1), 4–11. <https://doi.org/10.1177/1946756719889053>.
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial Intelligence in Education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Shum, S., Santos, O., Rodrigo, M., Cukurova, M., Bittencourt, I., & Koedinger, K. (2021). Ethics of AI in education: Towards a community-wide Framework. *International Journal of Artificial Intelligence in Education*, 32. <https://doi.org/10.1007/s40593-021-00239-1>.
- Hudlicka, E. (2016). Virtual affective agents and therapeutic games. In D. D. Luxton (Ed.), *Artificial Intelligence in behavioral and Mental Health Care* (pp. 81–115). Elsevier.
- ICAT (Initiative for Climate Action Transparency) (2020). Stakeholder Participation Guide: Supporting Stakeholder Participation in Design, Implementation and Assessment of Policies and Actions. Climate, Community & Biodiversity Alliance and Verra. <https://climateactiontransparency.org/wp-content/uploads/2020/04/Stakeholder-Participation-Assessment-Guide.pdf>.
- Inayatullah, S. (2007). *Questioning the future: Methods and tools for Organizational and Societal Transformation*. Tamkang University.
- Jobin, A., Ienca, M., & Vayena, E. (2019). Artificial intelligence: The global landscape of ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399. <https://doi.org/10.1038/s42256-019-0088-2>.
- Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>.
- Kosow, H., & Gaßner, R. (2008). *Methods of future and scenario analysis: Overview, assessment, and selection criteria*. German Development Institute. <https://www.idos-research.de/en/studies/article/methods-of-future-and-scenario-analysis-overview-assessment-and-selection-criteria/>.
- Kreibich, R. (2007). Wissenschaftsverständnis und Methodik der Zukunftsforschung. *Zeitschrift für Semiotik*, 29(2–3), 177–198.
- Lados, G., Hegedus, G., & Kovacs, Z. (2022). The role of identity and ontological (in)security in Return Migration: An empirical perspective from Hungary. *Journal of International Migration and Integration*, 24(2), 1–20. <https://doi.org/10.1007/s12134-022-00964-z>.
- Landeta, J. (1999). *El método Delphi. Una técnica de previsión para la incertidumbre*. Ariel.
- Levinas, E. (1969). *Totality and infinity: An essay on exteriority*. Duquesne University Press.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.
- Lynch, K. (2022). *Care and capitalism*. Polity Press.
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). AI and education – Guidance for policy-makers. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>.
- Mouta, A., Paulino, A., Ferreira, J., & Couto, F. (2015). Pedagogy as a Techné: Meaningful ICT Integration in formal Learning Scenarios. In G. Roman, J. Komuves & R. Hahn (Eds.), *Proceedings of the European Distance and E-Learning Network 2015 Annual Conference* (64–72). EDEN. <https://doi.org/10.38069/edenconf-2015-ac-0009>.
- Nesta (2019, January 22). Mapping global approaches to AI governance. <https://www.nesta.org.uk/blog/mapping-global-approaches-ai-governance/>.
- Nguyen, A., Ngo, H. N., Hong, Y., Dang, B., & Nguyen, B. P. T. (2022). Ethical principles for artificial intelligence in education. *Educational Technology Research and Development*, 1–25. <https://doi.org/10.1007/s10639-022-11316-w>.
- Nuwan, P. M. M. C., Perera, B. A. K. S., & Dewagoda, K. G. (2021). Development of Core Competencies of Construction Managers: The Effect of Training and Education. *Tech Know Learn*, 26, 945–984. <https://doi.org/10.1007/s10758-020-09474-2>.
- Nygaard, K., Graversgaard, M., Dalgaard, T., Jacobsen, B. H., & Schaper, S. (2021). The role of Stakeholder Engagement in developing New Technologies and Innovation for Nitrogen Reduction in Waters: A longitudinal study. *Water*, 13(22), 3313. <https://doi.org/10.3390/w13223313>. MDPIAG.

- Ogbeifun, E., Agwa-Ejon, J., Mbohwa, C., & Pretorius, J. H. (2016, March). The Delphi technique: A credible research methodology [Paper presentation]. 2016 International Conference on Industrial Engineering and Operations Management. Kuala Lumpur, Malaysia.
- Pammer-Schindler, V., & Rosé, C. (2021). Data-Related Ethics Issues in Technologies for Informal Professional Learning. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00259-x>.
- Patton, M. (1990). *Qualitative evaluation and research methods*. Newbury Park.
- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22. <https://doi.org/10.1186/s41039-017-0062-8>.
- Renzi, A., & Freitas, S. (2015). The Delphi Method for future scenarios construction. *Procedia Manufacturing*, 3, 5785–5791. <https://doi.org/10.1016/j.promfg.2015.07.826>.
- Richards, D., & Dignum, V. (2019). Supporting and challenging learners through pedagogical agents: Addressing ethical issues through designing for values. *British Journal of Educational Technology*, 50(6), 2885–2901. <https://doi.org/10.1111/bjet.12863>.
- Sanromà-Giménez, M., Lázaro Cantabrana, J. L., Usart Rodríguez, M., & Gisbert-Cervera, M. (2021). Design and validation of an Assessment Tool for Educational Mobile Applications used with autistic learners. *Journal of New Approaches in Educational Research*, 10(1), 101–121. <https://doi.org/10.7821/naer.2021.1.574>.
- Santos, J. M. (2012). *Introdução à Ética Sistema Solar*.
- Schiff, D. (2021). Education for AI, not AI for education: The role of Education and Ethics in National AI policy strategies. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-021-00270-2>.
- Serholt, S., Barendregt, W., Vasalou, A., Alves-Oliveira, P., Jones, A., Petisca, S., & Paiva, A. (2017). The case of classroom robots: Teachers' deliberations on the ethical tensions. *AI & Society*, 32(4), 613–631. <https://doi.org/10.1007/s00146-016-0667-2>.
- Slaughter, R. A. (2020). Futures studies as a Quest for meaning. *World Futures Review*, 12(1), 26–39. <https://doi.org/10.1177/1946756719870277>.
- Stein, Z., Engberg, J., & Brown, T. (2017). The seductions of quantification: Measuring human rights, gender violence, and sex trafficking. *Signs: Journal of Women in Culture and Society*, 42(2), 397–422. <https://doi.org/10.1086/690937>.
- Stiegler, B. (1998). *Technics and Time, 1. The fault of Epimetheus*. Stanford University Press.
- The Institute for Ethical AI in Education (2021). The ethical Framework for AI in Education. <https://fb77c667c4d6e21c1e06b-cdn.net/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>.
- Tundrea, E. (2020). Artificial Intelligence in Higher Education: challenges and opportunities. INTED2020 Proceedings, 2041–2049. <https://doi.org/10.21125/inted.2020.0644>.
- van Notten, P. (2006). Scenario Development: A typology of approaches. In OECD (Ed.), *Think scenarios, Rethink Education* (pp. 69–92). OECD Publishing.
- Wang, H., Thili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X., & Fei, C. (2023). Examining the applications of intelligent tutoring systems in real educational contexts: A systematic literature review from the social experiment perspective. *Education and information technologies*, 1–36. Advance online publication. <https://doi.org/10.1007/s10639-022-11555-x>.
- Williams, R., Machado, C. V., Druga, S., Breazeal, C., & Maes, P. (2018). "My Doll Says It's OK": A Study of Children's Conformity to a Talking Doll. In M. Hoiseth, G.E. Baykal, A. Peter, B. Farschian (Eds.), *Proceedings of the 17th ACM Conference on Interaction Design and Children* (pp. 625–631). Association for Computing Machinery. <https://doi.org/10.1145/3202185.3210788>.
- Wright, D., Finn, R., Gellert, R., Gutwirth, S., Schütz, P., Friedewald, M., Venier, S., & Mordini, E. (2014). Ethical dilemma scenarios and emerging technologies. *Technological Forecasting and Social Change*, 87, 325–336. <https://doi.org/10.1016/j.techfore.2013.12.008>.
- Zhang, K., & Aslan, A. B. (2021). AI technologies for education: Recent research & future directions. *Computers and Education Artificial Intelligence*, 2, 100025. <https://doi.org/10.1016/J.CAEAI.2021.100025>.
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the New Frontier of Power*. Profile Books.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

APPENDIX C

PUBLISHED MANUSCRIPT IN DIGITAL SOCIETY

Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2025). “Where is agency moving to?”: Exploring the interplay between AI technologies in education and human agency. *Digital Society*, 4(49). <https://doi.org/10.1007/s44206-025-00203-9>

Acknowledgment Note:

This publication resulted from a collaboration within the research network “AI in Education: Ethical and Epistemic Perspectives”, coordinated by Dr. Vlasta Sikimić of the Philosophy & Ethics group at Eindhoven University of Technology. As part of this collaboration, the author was invited to contribute to the Topical Collection titled “AI in Education: Ethical and Epistemic Perspectives” in the Springer journal “Digital Society”.

“Digital Society” does not have a JIF yet because it is a new journal published by Springer, dedicated to Ethics, Socio-Legal and Governance of Digital Technology.

Digital Society (2025) 4:49
<https://doi.org/10.1007/s44206-025-00203-9>

ORIGINAL PAPER



“Where is Agency Moving to?”: Exploring the Interplay between AI Technologies in Education and Human Agency

Ana Mouta¹ · Ana María Pinto-Llorente² · Eva María Torrecilla-Sánchez²

Received: 1 November 2024 / Accepted: 12 May 2025
© The Author(s) 2025

Abstract

This study explores the relationship between human agency and the use of sociotechnical AI technologies in educational contexts. While AI applications in education promise benefits such as personalised learning experiences and enhanced efficiency, empirical evidence supporting these claims remains limited, particularly regarding their implications for agency across different levels. Reflecting on these issues, this research critically evaluates how these technologies impact agency dynamics, namely on the subjective, intersubjective, and collective dimensions of agency as perceived by educators. The current study forms part of a broader Educational Design Research project, structured across multiple stages (including a Systematic Literature Review, a Delphi study, and focus groups with teachers). Drawing on this triangulated methodological architecture, the paper specifically focuses on the qualitative analysis of discussions about how AI systems may transform human agency across different levels. This methodology facilitates a comprehensive examination of both the opportunities and risks associated with the increasing integration of AI in educational settings. Insights gathered reveal that educators present a layered understanding of their students’ subjective agency but may overlook the broader intersubjective impacts of AI systems, potentially leading to narratives that envision a future where the role of teachers is diminished. While the immediate focus is on understanding educators’ perceptions of agency transformations through AI integration, the study also contributes to a wider objective informing the design of a professional development framework for teachers. In particular, the research aims to provide critical foundations for the development of pedagogical interventions that intentionally address agency within educational contexts where AI systems are increasingly present. education, providing meaningful substance for debates among educators, policymakers, and researchers.

Keywords Teacher agency · Artificial intelligence in education · Professional development · Training · Human agency · Participatory research

Extended author information available on the last page of the article

Published online: 11 June 2025

Springer

1 Introduction

In an age of increasing automation and technological reliance, the dynamics of agency are evolving. Interactions with sociotechnical systems, especially Artificial Intelligence (AI), present new challenges and opportunities for individual and collective agency. In response, there is a pressing need to foster forms of agency that are not only personal but also relational and collectively distributed, grounded in individuals' capacity for internal conversation, which enables them to reflect on and shape their engagements (Archer, 2003), and in their interdependence with others and with the natural world in which we live and learn (UNESCO, 2021).

More recently, the interplay between human and non-human entities has challenged the idea of agency as an individual property. These non-human entities can include machines, algorithms, technologies, animals, and even objects or environmental factors. Thus, debates have been exploring the extent to which agency is distributed across a network of actors, both human and non-human (Deleuze & Guattari, 1980; Haraway, 1985, 2016; Latour, 2005, 2018). Latour (2018), in particular, explores how the alignment (or misalignment) between macro-level planetary concerns, meso-level institutions, and micro-level everyday actions is entangled, calling for multiple levels of agency. These discussions hold particular significance in helping remember and actively engage with our shared agency within this trading zone. As Latour (2014, 2018) suggests, there are two main values to this approach: first, it allows us to identify potential actants before they become actors; second, it fosters a connection between citizens and their concerns, linking them to their surroundings and shared environment, gradually shaping a common world to be collectively built.

Building on these broader debates on distributed and relational agency, this paper turns to the field of education as a critical space for exploring how agency is experienced, challenged, and shaped in the age of AI. Education is not only where individual autonomy is cultivated, but also where collective responsibilities are negotiated.

In light of these concerns, this study forms part of a broader Educational Design Research (EDR) project, structured across multiple stages. The research has progressed through a Systematic Literature Review (SLR), a Delphi study employing Futures Studies methodology, and focus groups involving teachers and trainers from K-12 and higher education sectors. In this way, the study builds on a triangulated methodological architecture, drawing insights from critical mapping, scenario exploration, and empirical inquiry. Together, these phases aim to collaboratively inform the design of a professional development course focused on agency and AI integration. While the present paper primarily focuses on the findings from the focus group study on the theme of agency, it is situated within this wider EDR framework, which also engages with the regulatory and policy implications of AI in education. In particular, the findings highlight the absence of clear agency frameworks to guide the ethical adoption of AI technologies, which is an issue that will be further explored through the subsequent design, implementation, and evaluation stages of the training project.

To explore these questions in greater depth, the next subsections of the introduction elaborate on two core foundations: first, by tracing how the concept of human agency has evolved across philosophical, psychological, and sociological traditions, and second, by exploring how agency has been framed and fostered within educa-

tional contexts. These discussions then lead into a framework for examining how AI affects agency on three interconnected levels (subjective, intersubjective, and collective), which structure both the theoretical framing and the empirical analysis presented in the paper.

1.1 The Concept of Human Agency

The concept of human agency, which can be largely defined by the capacity to act independently and make choices that influence one's environment, holds profound significance for human and societal development (Bandura, 2001). As a temporally embedded process, shaped by the past, oriented towards future possibilities, and responsive to present contingencies, it allows individuals to explore and reconfigure their relationship to social structures, which are themselves temporal and relational. Social actors continuously shift their orientations within overlapping temporal contexts, enabling them to engage in flexible, situated action (Emirbayer & Mische, 1998). This notion extends beyond mere individual decision-making, encapsulating social, cultural, and technological dimensions that shape collective experiences and behaviours. In a world characterised by rapid technological advancements and evolving societal norms, understanding agency has become increasingly critical, particularly within the realms of education. The ability to assert agency is not only foundational for personal development but also key for collectives, fostering responsible citizenship, engaged practices, and contextual-relevant frameworks for action within societies (UNESCO, 2021).

Debates surrounding agency have long focused on the nature of free will, moral responsibility, and the tension between individual autonomy and social structures. During the Enlightenment, thinkers framed human freedom as a response to deterministic views rooted in religious morality (Biesta & Tedder, 2006). People were seen as wilful beings, capable of independent judgment and autonomous action, embodying the Enlightenment ideal of rational individuals with moral agency. Philosophical perspectives on agency evolved to include both normative, non-rational action (influenced by Kantian ethics) and rational, instrumental action, later developed by American pragmatism and Continental phenomenology. The former reflects a moral will, where actors (individually or collectively) engage in moral action (Kant, 2005), while the latter frames agency as rational, goal-oriented behaviour to achieve personal interests or fulfil material needs (Marx & Engels, 1976).

Throughout the early modern period, questions of personal identity and consciousness began intersecting with concepts of agency, with existentialists like Sartre highlighting the importance of individual responsibility in one's choices (Sartre, 1943). By the 20th century, sociological thought had expanded to explore how agency is deeply influenced by historical and social contexts, investigating its embeddedness within socio-economic structures and power dynamics (Foucault, 1975). This exploration led to the development of theories aimed at bridging the structure/agency dualism. Giddens (1991) proposed that social structures emerge from repeated patterns of human action and are simultaneously shaped by, and constraining of, that activity, thus placing structure as both a product and a constraint of agency. Bourdieu (1990) similarly argued that habitual patterns, or practices, become structure over

time, guiding future actions. However, reactions to Giddens's theory emphasised that agency is not merely the repetition of patterned action; it also encompasses the ability to break with established rules, reflecting thought, reflexivity, and creativity (Emirbayer & Mische, 1998).

The rise of phenomenology introduced discussions about self-ownership and the pre-reflective experiences of causing actions, deepening the understanding of how individuals experience agency. Within the framework of social cognitive theory, Bandura (2006) offers one of the most comprehensive accounts of human agency and remains a central reference in educational and developmental psychology. His theory acknowledges the contribution of genetic heritage in providing the neural structures and mechanisms necessary for distinctively human traits like generative thinking, symbolic communication, intentionality, foresight, self-regulation, and reflective self-awareness. It rejects the dualism between individual agency and disembodied social structures, positing that humans intentionally influence their functioning and circumstances. Bandura (2001, 2006) identifies three forms of agency: individual, proxy, and collective, each playing a distinct role in how individuals and groups interact with the world and shape their existence. Individual agency is characterised by a person's ability to exercise control over their own actions. For Bandura, self-reflectiveness, the metacognitive ability to observe and evaluate one's thoughts and actions, adjusting strategies as needed to accomplish goals, is the most distinctive feature of human agency. This capacity for reflection enables individuals to align their actions with moral standards, making agency not just a matter of execution but also of ethical and moral self-regulation.

Beyond individual agency, this author is also one of the few major theorists to develop the concept of proxy agency, which occurs when a person depends on others who have the necessary resources, skills, or authority to act on their behalf. While this form of agency can be crucial for achieving certain outcomes, it also creates a reliance on external agents, potentially leading to a diffusion of personal responsibility and affecting moral accountability. This is particularly pertinent in the context of AI and education. Contemporary uses of AI systems, whether through adaptive learning platforms, generative tools, or automated feedback mechanisms, frequently involve students and teachers delegating cognitive or pedagogical tasks to technological agents. Bandura's conceptualisation helps unpack the psychological implications of such delegation, including how it affects motivation, ownership, and accountability. This dimension of proxy agency is still underdeveloped in many contemporary theories of agency, which tend to focus either on individual autonomy or distributed, collective enactments without addressing the intermediary or surrogate roles played by technologies.

On the other hand, collective agency arises when individuals engage together to achieve shared goals. This form of agency is distinguished by not merely being the sum of individual contributions but emerging from the synergistic interactions among group members, where each plays a key role. It is especially significant in social and educational settings, where a sense of belonging and shared responsibility fosters meaningful endeavours. Bandura argues that these three forms of agency often intermingle.

1.2 Human Agency and Education

The 1960s and 1970s saw a resurgence of the New School movement, originating at the start of the century, which advocated for greater democratic involvement in education. This shift was, in part, a reaction to the impacts of industrialisation, which had turned traditional schooling into a standardised and conformity-driven system. Influenced by socialist and anarchist ideals, New Schools aimed to create spaces where students could develop freely and critically engage with societal norms and structures. However, it wasn't until the 1990s that education fully pivoted toward fostering learner autonomy. This shift focused on students' responsibility for their own learning (Jones, 2007) and was influenced by Carl Rogers's humanistic approach (Rogers, 1983). In terms of learning, agency is seen as both a process and an outcome. According to Bandura (2001), learners with agency can bring about things intentionally and be responsible for their self-development, adaptation, and self-renewal. At the individual level, a robust sense of agency enhances motivation, enabling to respond to challenges with resilience. Research indicates that high self-efficacy correlates with sustained engagement and achievement across various contexts (Vera et al., 2014). Initiatives prioritising the cultivation of agency encourage students to embrace autonomy, build confidence, and engage in self-directed learning (Davis-Keen, 2005).

While crucial for individual achievement, the field of education extends far beyond the outcomes of personal agency. Educational environments are privileged spaces for cultivating the political foundation of democracies (Dewey, 1916) through collective agency, enabling students not only to take ownership of their learning but also to engage in dialogic practices that reflect the interests, dynamics, and unique experiences of their institutions. This interplay between individual and collective agency is also demonstrated by Zumeta et al. (2016), who found that shared emotional experiences and group identification in collective activities enhance team achievement, highlighting the importance of community in nurturing agency. Collaborative learning experiences in educational settings foster not only academic achievement but also social connectedness and shared responsibility – critical components for building cohesive societies (OECD, 2012). Thus, experiencing opportunities for agency within educational contexts enables students to recognise their role within a larger community, fostering a sense of belonging and collective participation. In this regard, Dewey's (1916) advocacy for an education based on principles of experimental democracy gains renewed significance.

1.3 Education and AI

Contemporary research highlights the potential of Artificial Intelligence in Education (AIED) to personalise learning experiences, streamline administrative processes, and offer tailored pedagogical support (Wang et al., 2023; Mousavinasab et al., 2018). Intelligent tutoring systems (ITS) and adaptive learning platforms can now analyse large datasets on students' strengths, weaknesses, preferences, and learning styles, creating customised learning pathways that adjust dynamically to individual needs (Miao et al., 2021; Luckin et al., 2016). More recently, greater attention has been given to the fact that students are neurodiverse, bringing with them a range of assets

derived from their experiences at home, within their communities, and across their cultural backgrounds. In response to this diversity, researchers have recognised that AI systems, with their capacity to detect patterns in complex datasets and to generate customised content, have the potential to extend beyond the most common learning trajectories and to offer personalised educational experiences that would be difficult for individuals to produce unaided (U.S. Department of Education, Office of Educational Technology [USDOE-OET], 2023). Moreover, embedded adaptive assessments facilitate real-time feedback and allow for continuous monitoring of learners' progress, engagement, and emotional states, thereby supporting more responsive teaching practices (USDOE-OET, 2023; Luckin et al., 2016). AI-driven tutoring support, through interactive simulations and virtual dialogues, can further enhance comprehension by delivering targeted feedback and gradually refining instructional strategies through ongoing learner interaction (Wang et al., 2023; Khosravi et al., 2022).

As AIED systems continue to collect and analyse performance data, they increasingly enable early identification of students at risk of falling behind, thus allowing for timely pedagogical interventions. Beyond supporting individual students, the use of AI in school administration is often seen as a way to strengthen collective agency, enabling institutions to make more informed decisions, better allocate resources, and respond more thoughtfully to the needs of their communities.

Although these technological advancements are often celebrated for their efficiency and their potential to enhance learning and education quality, they also carry the risk of diminishing individuals' sense of control and autonomy (Moore, 2016). Studies indicate that the design of human-machine interfaces significantly impacts users' perceived agency, with greater agency associated with human-like interactions (Berberian et al., 2012; Obhi & Hall, 2011). This necessitates a critical examination of how we perceive and integrate these sociotechnical systems into educational practices, ensuring that they serve as tools that provoke our collective engagement towards agency. Furthermore, the ethical implications of agency extend beyond personal development to influence societal norms and values. A diminished sense of agency can lead to moral detachment and detached behaviour, affecting how individuals attribute responsibility for their actions (Haggard & Tsakiris, 2009). In educational contexts, sense of agency has also a role in equipping students with the ethical reasoning skills necessary to question complex societal challenges, promoting accountability and moral engagement with life (Bandura et al., 1999).

As educators (and society more broadly) continue to engage with systems dominated by techniques that prioritise rationality and efficiency (Ellul, 1954), which are deeply embedded in the secular liturgies of school practices, their stance is particularly necessary for reviving the spirit that animates schools and preserving their organic capacity to respond to their own experiences. Practically, this means that teachers are expected to play a meaningful role in shaping the policies they are tasked with enacting, alongside families and community members, who are critical partners in educational development (European Union, 2018). This need for agency is especially relevant in the context of technological advances and narratives surrounding AI's autonomy. Pasquinelli (2020) argues that labour is at the root of the historical development of AI, which is shaped by a composite of dynamics reflecting social

relations and power structures, where the current control is exerted by high-tech companies. This understanding highlights why agency cannot be separated from collective dynamics and historical contexts. Teacher agency, therefore, is also the result of interactions between personal capabilities and the environmental factors influencing their actions (Priestley et al., 2015). In an era of high automation, educators have a role in reclaiming the significance of agency and challenging the contested narratives surrounding automation and work. As AI evolves into a regime that replaces an episteme of causation with one of automated correlations (Pasquinelli, 2020), educators are called to cultivate spaces where collective knowledge and agency can flourish. This perspective has direct implications for the way teacher professional development is conceived and designed, as well as for a comprehensive understanding of human and societal development processes, positioning collective agency as one of the ultimate goals of education.

So, in the context of education, three levels of concern shape our understanding of how AI systems could impact agency: the subjective, intersubjective, and collective levels. On the subjective level, agency is explored through dimensions that examine how AI may affect human experiences in areas like decision-making, moral judgment, deliberate action, achievement, individuation, and project ownership. Intersubjectively, agency is negotiated among peers, between teachers and students, and among other stakeholders, including families. Here, power dynamics, teachers' status, sense of belonging, and scaffolding experiences come into play. Collectively, agency is examined through situational freedoms within institutional policies, collective deliberation and achievement, political development, and democratic practices.

2 Method

2.1 Research Question

In this context, this study aims to understand how teachers conceptualise the transformation of human agency in relation to AI in education, underpinned by the theoretical and methodological framework of dialogic ethics. The research is based on a qualitative analysis of data collected from focus groups consisting of teachers, professors, and teacher trainers, as part of a broader Educational Design Research project aimed at designing a training programme for K-12 educators that addresses the integration of AI in education, along with ethics and agency. While the initial purpose of the study was to explore collective perspectives on what a training programme on such themes should include, a significant portion of the generated text – though not directly focused on that theme – was closely related to sociotechnical AI systems and education. This prompted a new approach to data analysis, leading to a revised goal deemed highly significant for understanding the role of agency at the intersection of these technologies and pedagogical practices, as perceived by one of the key educational actors: the teachers. Therefore, the research question for this phase was defined as follows: How do teachers conceptualise the transformation of human agency in relation to AI in education on subjective, intersubjective, and collective levels?

2.2 Participants, Implementation, and Data Analysis

Participants were selected based on their experience in K-12 teacher education and their proficiency in Spanish, the primary language of the research centre. Initially, convenience sampling was used, followed by snowball sampling to ensure cultural diversity. Out of 40 invitations, 19 educators from five countries (Colombia, El Salvador, Portugal, Mexico, Spain) participated, all of whom held roles in pre-service or in-service teacher education. The study involved a balanced representation of males ($N=10$) and females ($N=9$), with a mean age of approximately 45 years. Among them were professionals in various roles, including a Director of Digital Transformation in Higher Education, several associate professors, and former K-12 teachers who held significant positions in the Ministry of Education's Master Plans. The collective expertise of these participants spanned various aspects of education, including research methods, teacher training, and the integration of information technology in educational settings, reflecting a strong emphasis on the intersection of education and technology.

In 2022, they engaged in discussions of hypothetical scenarios – dystopian perspectives on AI use in education derived from a Delphi method expert group – based on earlier research (Mouta et al., 2023) across four sessions to ensure data saturation. Conducted virtually via Zoom due to pandemic constraints, these sessions gathered a range of perspectives that contributed to the programme's development. Sessions were recorded, transcribed using AI software, and verified by two researchers. This phase of the analysis combined deductive and inductive methods, identifying themes based on existing literature, earlier project phases, and insights from raw data. As the purpose for this phase of the study is to understand teachers' perspectives on how the sense of agency can be influenced or transformed by the use of AI systems in education, the entirety of the data was once again explored through different search strings, namely: “agency”, “autonomy”, “author”, “authorship”, “capacity for action”, “control”, “decision-making”, “independence”, “initiative”, “personal sovereignty”, “self-determination”, and “self-governance”. This approach emerged from the observation that participants often discussed the topic of human agency using various related terms, which, although not explicitly labelled as “agency”, resonated within the same conceptual universe. This practice of broadening the scope to include associated concepts ensured a comprehensive understanding of how teachers perceive agency. The final selection of categories was determined through two reliability tests conducted with a two-month interval between coding stages, as well as an inter-rater analysis performed by two researchers who compared the data directly (in the second stage) and reached a consensus.

2.3 Results Communication

This section outlines the main findings from the focus group discussions, addressing the research question that guided this phase of the project: How do teachers conceptualise the transformation of the sense of agency in relation to AI in education on subjective, intersubjective, and collective levels?

The results are presented in a structured table format that delineates the findings across different levels of agency (cf. Table 1). Each row of the table is categorised by (1) “Level” (subjective, intersubjective, collective), highlighting the context in which teachers discussed agency. Within each level, specific (2) “Dimensions” are outlined, reflecting the key aspects of agency that emerged during discussions. Accompanying each dimension is a (3) “Description”, which contextualises the findings and provides a concise overview of the issues identified. Furthermore, (4) “Examples” from the focus group discussions are included to illustrate the points made within the descriptions.

3 Discussion

This discussion section is organised into two main parts, drawing on the qualitative data presented in Table 1. The first explores how agency is perceived by teachers across individual, relational, and systemic dimensions of school life, and how AI sociotechnical systems may interfere at each of these levels. Particular attention is given to how teachers interpret the evolving role of AI in relation to students’ developmental processes, the professional identity and responsibilities of educators, and the democratic ethos of educational institutions. The second part presents an integrative analysis that identifies key trends across the three levels, highlighting patterns and interdependencies between forms of agency. These include tensions between automation and authorship, the evolving nature of proxy agency, and the need for moral scaffolding and shared responsibility.

3.1 Exploring Agency across Subjective, Intersubjective, and Collective Dimensions

3.1.1 Subjective Level

Teachers expressed concerns about fundamental developmental processes, such as individuation and moral development, acknowledging their role as committed keepers of the educational mission, which extends far beyond the mere goal of content learning (Biesta, 2020). They also highlighted specific meta-cognitive and affective processes, which serve as qualitative differentiators for holistic development – a responsibility shared by all educational actors and formal institutions. These processes include the architecture of decision-making, encompassing critical reasoning, recognition of power structures, and the consideration of diverse options.

In terms of critical reasoning, teachers recognise that schools have a responsibility to create conditions that help students understand the personal variables influencing their choices throughout life; factors which, as many vocational psychologists have suggested (e.g., Guichard, 2009), directly impact vocational exploration, career self-determination, and work well-being. This task, which can be pedagogically designed, is significant for exploring personal critical development dimensions, such as interests, competencies, values, influences (both proximal and distal), and both positive and less positive personal bonds. This can be achieved through curriculum infusion

Table 1 Teachers' perspectives on AI and human agency: focus group findings

Level	Dimension	Description	Examples
SUBJECTIVE	DECISION-MAKING PROCESSES	<p>Critical Reasoning AI-driven decisions may undermine students' ability to reason about personally significant variables.</p> <p>Power of Structure Influence of social, cultural, and environmental factors that shape and may constrain individuals' choices based on available resources and visible opportunities, underscoring the need to critically assess whether technology and AI can expand or restrict these options.</p> <p>Reduction of Diversity and Circumscription of Choices AI systems that rely on pre-defined learning paths may restrict students' ability to explore a diverse range of possibilities.</p> <p>Differentiation and Authorship AI systems are expected to support the educational community exploring the diverse developmental trajectories and unique personal expressions of each student (e.g., learning pace, behavioural repertoire, specific manifestations) rather than enforcing standardised measures.</p>	<ul style="list-style-type: none"> • "When AI decides what is best for the student... we're taking away the student's ability to decide, what they want to learn, or how they want to approach a topic." • "It makes the student unaware of why a certain personalisation process is beneficial. There is an entire external system that is managing their personalisation, which means that instead of fostering autonomy, what is actually being encouraged is heteronomy." <p>"the reality is that the decisions we make are constrained by our surroundings, by what we can see. We choose to study from the options available in our community. (...) we shouldn't assume from the outset that technology or AI will restrict that freedom. Instead, we need to assess whether we truly have those freedoms today."</p> <p>"I think there's a loss of autonomy because if the student is only able to choose from a list of options... they're just choosing what was already decided."</p> <ul style="list-style-type: none"> • "This individual vision of each student and their needs contrasts with the generic view that ultimately states we don't want everyone to be the same, we need to delve deeper into each person being distinct and unique. Therefore, this type of system that attempts to homogenize and make everyone equal ends up breaking down the personal characteristics of each individual." • "what we do is help the student, year after year and day by day, develop their capacity to take ownership of the condition of being the author of their own life. They must create their life, but that is not easy. It is definitely not easy."
	INDIVIDUATION		

Table 1 (continued)

Level	Dimension	Description	Examples
	FORESIGHT, SELF-REGULATION AND SELF-REFLECTIVENESS	Agentic Properties Qualities that enable individuals to take ownership of their lives and decisions, enabling individuals to actively shape their experiences and achieve their goals.	“The ability to take ownership of one’s own life (...) the day my time runs out, my life runs out. Therefore, autonomy is about development; autonomy is about beginning to give the student time so that they can take ownership of their time, envision their future, plan ahead, and be capable of committing their ability to act over time, almost to the point of being able to commit their capacity for work.”
	MORAL DEVELOPMENT	Moral Disengagement Progressive detachment from responsibility and erosion of ethical standards that occurs when individuals justify or rationalise actions, often by attributing decision-making to external forces, such as technology or authority figures. This can diminish personal accountability and moral reasoning. Teachers’ roles Erosion of teachers’ roles and purpose, as AI begins handling core pedagogical tasks, potentially turning their role into one of mere oversight.	“efforts are being made to incorporate Artificial Intelligence into these autonomous cars (...) Artificial Intelligence can reduce our capacity for agency. For some people, this may be satisfying because making those decisions is certainly not easy, and there are individuals who may feel relieved not having to make those choices.”
	INTERSUBJECTIVE PURPOSE AND VALUES	Proxy Agency AI systems are viewed as supportive tools, with the expectation that teachers will maintain their role as proxy agents, preserving their professional autonomy and authority.	“If AI starts taking over the planning and grading, then where does that leave us as teachers? We don’t want to just be there to supervise what the machine is doing.”
	AUTHORITY AND PROFESSIONAL STATUS	Scaffolding Supportive process through which teachers leverage their own experiences and socio-emotional skills to guide students in developing critical competencies.	“AI can make suggestions, but it shouldn’t make decisions. It should be the teacher who decides... because we know the students better than any algorithm.”
	MODELLING AND CONFIDENCE		“But the experience and socio-emotional skills that they have developed, along with the failures they have encountered in the past, will allow the teacher to guide students in discerning what information is good, what information is not good, how to work in teams, how to communicate better, how to collaborate, and how to solve problems (...) their experience, can help their students develop these socio-emotional skills”

Table 1 (continued)

Level	Dimension	Description	Examples
Collective	POLITICAL SPACE	<p>Participation Importance of a participatory approach where all stakeholders (e.g., students, teachers, parents) are involved in shaping AI's role in education.</p> <p>Distributed Decision-making Call for collaborative decision-making in AI integration, with teachers advocating for a role in shaping its implementation, highlighting the need for democratic practices at an institutional level.</p> <p>Democracy and Rule of Law Importance of participation, agency, and responsibility, rooted in the needs of the community. It promotes a society guided by established institutional frameworks that protect individual freedoms and enhance social welfare.</p> <p>Sociotechnical Assemblages Dynamics and shared agency between human and technological actants create a unique context for learning and reinforce the idea that the educational experience is co-constructed, shaped by the interplay of diverse influences.</p>	<p>“There needs to be a way for students, teachers, and even parents to give input on how AI is used... so we should all have a voice.”</p> <p>“We need to have a say in how AI is being used in our schools. It shouldn't be a top-down decision...”</p> <p>“In my view, I believe that two beautiful aspirations in a person's life can be striving to live in a democratic system and being guided by the factors established within the framework of institutions. This means aspiring to these two ideal goals that exist in most countries.”</p> <p>“We could engage in dialogue with parents and teachers about viewing Artificial Intelligence and other aspects of information and communication technologies as added values to the learning processes and the construction of knowledge that have been developed. The very systems of life and the evolution of societies drive us to recognise that there must be different approaches to build learning.”</p>

activities, project-based learning, and role-taking experiences, where intentionality and systematicity are key. In this context, the role of AI systems is expected to be critically examined. AI systems that provide overly rapid feedback and excessive automation may progressively be interpreted as authoritative figures that “hold the truth”, negatively impacting students’ sense of control. As research suggests, this perceived loss of control can be explained by high levels of automation and hierarchical structures within collaborative tasks. According to Gozli’s (2019) hierarchical model of agency, such a loss is not merely cognitive but begins at the foundational level of sensorimotor fluency, the capacity to skilfully and confidently engage with one’s environment. When this fluency is disrupted, students may experience agency as distant or inaccessible, leading to increased passivity and conformity. Consequently, such systems can reinforce passivity and conformity while distancing students from engaging with diverse significant others, potentially limiting the freshness of perspectives and contributing to emotional blunting.

Teachers also demonstrated sensitivity in recognising that social, cultural, and environmental factors shape personal agency (Larreamey, 2011) at various instances and levels throughout different moments in life. This expanded awareness opens up the possibility of understanding agency as emerging from the interplay of heterogeneous elements, where objects and technologies do not merely assist human action but actively participate in shaping outcomes (Latour, 2005). In schools, this calls for a shift in how learning experiences are understood: rather than viewing agency as a purely human trait expressed through choice or autonomy, it becomes something co-constructed with the tools, spaces, and systems that surround learners. As the human-machine boundary becomes increasingly blurred (Haraway, 1985), and as material environments exert their own quiet pressures and affordances (Bennett, 2010), schools are called to move beyond simplistic binaries, such as active/passive or autonomous/dependent, and instead design for agency as something distributed, situated, and always in-the-making.

The power of structures, and their tendency to push the margins towards the centre, is reflected in education through the dominant discourse of efficacy that permeates the Big EdTech world and underpins AI applications in learning environments. This often places the burden of adoption and exploration of associated challenges on individuals, irrespective of the structural inequalities they may face, reinforcing a model shaped by minimal governance and free-market dynamics (Selwyn, 2016). Rather than fostering individual agency, such a distanced and individualised approach to what Zuboff (2019) describes as an unprecedented transformation in the exercise of power may lead to a diminished, if not detached, experience of agency, ultimately weakening the capacity to connect with others and build agency collectively. Furthermore, as Watters (2021), Williamson (2022), and Saltman (2020) highlight, educational platforms and AI applications mirror Big Tech strategies by embedding practices of categorisation, optimisation, and impact measurement, promoting models of learning that reinforce privatisation, outsourcing, and labour exploitation. Such logics, while marketed as efficient, risk narrowing the space for variation, error, and alternative pedagogical approaches. This structural emphasis on predictability also affects how uncertainty is valued in emotional and cognitive development. Yet uncertainty, can bolster creative and resilient approaches to dealing with circumstances and

chance, and it may even contribute to the concretisation of technological tools in the process of their appropriation.

The role of education in supporting the development of students' individuation through differentiation authorship was also acknowledged, recognising a process in which one's voice, the diversity of experiences and actors, and the ability to articulate one's experiences through timely discursive connectors set the stage. How can a speculative argument on personalisation account for processes of meaning-making and signification? Adaptive learning systems may weaken or disconnect knowledge and learning from the cultural backgrounds, unique identities, experiences, and subconscious processes that are foundational to subjectivities. While educational sciences and psychology have moved beyond strictly behavioural and cognitive models, the continued reliance of adaptive learning systems on these frameworks, as both inputs and outputs, raises questions about their adequacy in addressing the complexity of human learning.

The properties of agency, as described by Bandura (2001, 2006), were referred to. Once again, teachers demonstrated their awareness of the conditions that enable their students to become not only knowledgeable but also the architects of their own lives. They recognise that intentionality and planning are only possible when opportunities for self-regulation and self-reflectiveness are experienced. These factors are, in fact, critical ingredients for learning and meaningfully exploring the world – enabling individuals to act and integrate through processes of signification within meaningful relationships. As AI systems grow more powerful and sophisticated, their explainability becomes increasingly difficult. Moreover, a greater sense of agency arises from outcomes generated by humans rather than machines, with emotional feedback playing a crucial role in behavioural adjustment. This leads to speculation that automated feedback may be less relevant for fostering lasting behavioural changes. Moreover, AI evaluations used for grading, particularly when lacking critical discussions from peers and teachers, can negatively impact self-regulation and self-reflectiveness processes. This may result in students struggling to connect curricula with personal development, leading to disengagement from the learning process and transforming the experience of achievement into a dualistic and impoverished experience of success versus failure.

Finally, teachers recognise that values and norms are context-dependent and can vary within a classroom, from classroom to classroom, from school to school, and from region to region, as well as among different stakeholders. With the advent of Generative AI, it may be challenging to harmonise perspectives on what should be accepted from this tool. For instance, can a student's work that utilised Generative AI be considered original? If students and researchers are honest about their use of the tool, can their work genuinely be regarded as their own? Are honesty and transparency equivalent? Furthermore, can Generative AI be seen as a proxy agent? The answer to these questions, as well as the plurality of these questions and their capacity to provoke further inquiry will be critical for moral reasoning and development moving forward.

3.1.2 Intersubjective Level

In terms of the intersubjective dimension, teachers primarily emphasised their role, which can be explained by narratives that view AI as a potential replacer or, at the very least, as a tool for fulfilling several critical functions traditionally held by teachers, while stripping away their core purposes and values. This raises the question: when teachers are relieved of administrative tasks, what other responsibilities are they freed from, and where are they displaced to? Some insights into possible answers may help explain why teachers highlighted their importance as role models, as well as their unique capacities for scaffolding. They also stressed their significance as proxy agents, possessing an authority that must be preserved to maintain their professional status and consistency.

However, the teachers involved in this research were notably silent regarding transformations among peers and the relevance of a sense of belonging in the learning experience. They did not address parents' concerns, nor did they discuss how distributed agency – where parental engagement is key – could serve as a touchstone for a meaningful accomplishment of educational initiatives. Such an approach to parents could help ensure that learning is transferable and generalisable to other contexts while preserving both the conditions and content relevant to learning. Furthermore, such a commitment may support the rise of AI literacy levels within educational communities.

Considering how detailed and prolific teachers were regarding the potential effects of AI systems on the subjective dimension of agency, this focus on their roles on the intersubjective level may express not only how hegemonic narratives from BigTech can affect their professional autonomy and deliberation but also their capacity to critically examine the impacts on various relationships and actors within schools. This narrowing of thought can be a consequence of these procedures, as well as the processes through which these sociotechnical systems reinforce their own entrenched status and solidify their influence, pushing aside the relevance of, or obscuring, the various levels of relationships that education encompasses.

3.1.3 Collective Level

From a collective perspective, teachers continue to uphold the vision of schools as democratic spaces, echoing Dewey's principles. In today's society, where political spaces are increasingly under threat – dialogues being replaced by (short and efficient) talks, plurality giving way to homogeneity, and politics overshadowed by policies – this vision holds significant importance. Creative writing and expression (often diminished by Large Language Models), for instance, do not merely produce novel outputs; they provide intellectual nourishment, prompting meaningful, context-driven innovation. Furthermore, achieving shared flow and emotional synchrony can foster a strong sense of belonging, enhancing participation, accountability, and distributed agency with expanded roles and synergies.

Situational freedoms have long been threatened by psycho-pedagogical techniques that presented themselves as a pharmakon in education, obscuring the vocation of education as a collective foundation for experiencing societal challenges and explor-

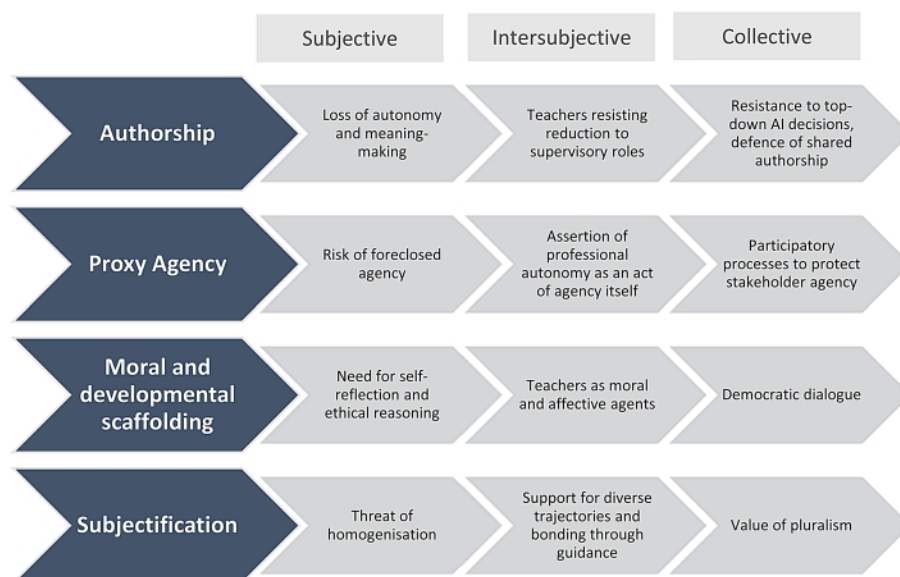


Fig. 1 Thematic flowchart of core categories across the subjective, intersubjective, and collective levels. Note. The diagram illustrates how the four key thematic categories are articulated across three analytical levels. Each level reflects a specific dimension of agency: subjective, intersubjective, and collective. Source: Authors' own elaboration based on qualitative analysis (2025)

ing personal genealogies of feelings and bonds. Technologies like automated essay scoring, predictive learning analytics, and AI-enhanced classroom management tools may offer immediate or short-term benefits but pose potential threats in the medium to long term. Critical pedagogy, which has traditionally integrated technology to support self-directed and heutagogic learning processes, is at odds with behaviour conditioning practices, such as gamified engagement. For instance, AI-enhanced peer tutoring systems can provide guidelines for constructive feedback, and gamified learning platforms powered by AI can facilitate role assignments tailored to students' unique characteristics. Teachers and students are less likely to view this technological support with suspicion if there are opportunities for organic interactions and spaces for personal and collective recognition of each individual's contributions. Such an environment can provoke affective-cognitive dissonance, which is crucial for embracing contradiction and divergence in social interactions and for building genuine common ground, far removed from illusions of full mastery.

3.2 Cross-Level Patterns and Interdependencies

A cross-dimensional analysis of the qualitative (cf. Fig. 1) data reveals thematic patterns that cut across the subjective, intersubjective, and collective levels. These themes illustrate the intricate nature of agency in educational contexts and demonstrate how AI sociotechnical systems may influence or reconfigure agency across these domains. Particularly, teachers' perspectives emphasise that agency is not a

fixed attribute but a dynamic, relational process, continually shaped by institutional structures, technological environments, and interpersonal interactions.

3.2.1 Tension Between Automation and Authorship

Across all three levels, teachers voiced concerns about how automation might displace human authorship and control. At the subjective level, this concern manifests in the perceived loss of students' autonomy, especially when AI-driven decisions bypass personal meaning-making. At the intersubjective level, the fear of being reduced to mere supervisors threatens teachers' roles as intentional and relational figures; roles that are critical to education's dual desideratum of subjectification (which is always made possible through the alterity of significant others) and socialisation (Biesta, 2020). Collectively, the emphasis on distributed decision-making reflects a broader anxiety over top-down AI implementations that may further erode stakeholder agency.

This analysis of the recurrence of the themes of authorship, meaning-making, and control across the three levels offers a comprehensive vision of agency as a composite phenomenon that intrinsically nurtures itself: the more it grows subjectively, the more it can be gained collectively; and the more it grows collectively, the more it can be regained subjectively.

3.2.2 Foreclosed Agency as a Substitute for Proxy Agency

Another thematic recurrence concerns the complex balance between proxy agency (teachers and technologies extending the agency of students) and a kind of foreclosed agency, in which both significant others and tools acquire the power to *resonate* or *act* on behalf of students. While teachers accept a mediating role, they caution against the complete outsourcing of educational decisions to AI. At the intersubjective level, teachers assert that professional autonomy must be preserved; this very assertion constituting a foundational act of agency. At the collective level, they advocate for participatory processes in AI implementation, ensuring that decisions are not dictated solely by external agents or commercial interests.

Once again, this pattern illustrates that agency is deeply relational. The development of autonomy in students is conditioned by the autonomy and reflexivity of educators, which, in turn, depends on democratic and transparent policies and institutional frameworks.

3.2.3 Moral and Developmental Scaffolding

Teachers repeatedly highlight their role in guiding moral development, fostering foresight, and supporting emotional and meta-cognitive growth, especially at the subjective and intersubjective levels. These processes, including the development of critical reasoning and self-reflection, are seen as critical for meaningful and context-sensitive learning. At the collective level, teachers view democratic participation and dialogue as key not only for institutional functioning but for cultivating students' sense of social responsibility.

Teachers conceive themselves not merely implementers of curriculum or AI facilitators working solely toward qualification (Biesta, 2020). They not only recognise, but seem to embody, a professional and expressive dimension: to be moral and affective agents in the lives of their students.

3.2.4 Subjectification and Diversity

The theme of diversity, both in learning experiences and in the exploration of possible developmental paths, is central. At the subjective level, AI's potential to homogenise education poses a threat to processes of subjectification (Biesta, 2020), where learners construct their identities through engagement and choice. Within the intersubjective space, teachers draw on their socio-emotional skills to support diverse student trajectories and to cultivate bonds and context-sensitive socio-learning experiences. At the collective level, an emphasis on inclusive, participatory decision-making reflects the value of diverse perspectives, both in fostering political and democratic development and in reinforcing a sense of belonging, shared flow and emotional synchrony, which significantly contributes to a sense of collective achievement (Zumeta et al., 2016). Therefore, as individuals become more attuned to their environment and to one another, they are also better positioned to exercise agency in transformative ways. This aligns with Archer's (2015) concept of double morphogenesis, which describes how structural changes create new conditions that elicit responses from social actors, who in turn reshape those very structures. In schools, evolving educational policies, social expectations, and technological shifts do not simply impose change from above; rather, they invite reflective engagement, collaborative tinkering, and new practices within communities. Through this recursive process, teachers and students move from merely responding to projects and initiatives to actively shaping them, making collective agency both a consequence and a catalyst of meaning-making and transformation.

These perspectives point to a vision of agency that is developmental, dialogical, and contextually situated (Archer, 2003; Biesta & Tedder, 2006; Emirbayer & Mische, 1998). Moreover, the analysis of these aspects is valuable not only as a safeguard for the mechanisms that sustain human authorship at each specific level, but also as a recognition of agency as a compositional phenomenon, where the diminishing of agency at one level appears to reinforce its erosion at others, ultimately cutting off the roots of schools as spaces for the joint elaboration of possible subjective and collective futures. And this goes beyond the idea of merely creating futures from which the present condition deprives us. It raises a more fundamental question: what kinds of futures can we imagine without a present that nourishes desire? In digital capitalism, the collapse of desire manifests as a frantic urgency to appear alive, set against an ever-expanding void of meaning. A truly collective future is to be conceived as a project both dictated by and dedicated to attentional forms of care, at once psychic and social (Stiegler, 2013). This is a project that understands desire as the intensification of individuation, rooted in the dynamics of attention, memory, and becoming. In this sense, agency can be understood as *protention*, in Stiegler's (2015) terms: the capacity to anticipate and project oneself into the future through the practice of imagination and desire. It is a temporal orientation that links subjectivity with the

not-yet, with what can still be created, decided, or hoped for. Within this framework, technologies, particularly AI, could occupy a meaningful role as forms of *tertiary memory*, that is, as externalised supports of human thought and collective experience. Rather than replacing memory, judgment, or intentionality, they could reconfigure traces of past knowledge and imaginative projections, helping individuals and communities to sustain or question their orientation toward meaningful futures. However, for this potential to be realised, we must break the cycle of closure, standardisation, and overdetermination that shapes these technologies; and this means the obvious: they must be reinvented.

4 Conclusion

This study highlights how human agency may be transformed through the use of sociotechnical AI systems in education. It explores the potential impact of these systems at subjective, intersubjective, and collective levels, drawing on the perspectives of teachers and professors responsible for teacher training. Concrete AI applications are presented in relation to their potential dimensions of impact, particularly concerning metacognitive and affective processes that contribute to long-lasting developmental processes, such as vocational, moral, and political growth.

The insights gathered reveal that teachers demonstrate high levels of differentiation when considering their students' subjective agentic dimensions. However, their focus appears more constrained when exploring the impact of AI systems on the intersubjective substratum. This narrowing may indicate an epistemic encroachment into narratives of dystopian futures devoid of teachers, paving the way for their own alienation from meaningful references within the framework of school bonds and strategic connections necessary for bringing projects to life and achieving impact.

This study is not without limitations. While the qualitative approach provides rich insights, it may not comprehensively capture the breadth of experiences and perspectives across various educational contexts and regions. Additionally, the focus on a specific demographic of educators may not represent the wider educational landscape. Future studies could include a broader range of voices, particularly from students, parents, educational administrators, and even vendors, to gain a more comprehensive understanding of the potential impact of AI on agency and institutional powers. Moreover, although the focus groups included educators from diverse backgrounds across K-12 and higher education, the findings were not differentiated systematically according to educational sector, subject area, or level of teaching experience. Future research could meaningfully explore these differences. For instance, teachers working in vocational education, operating closer to industry demands, might conceptualise agency differently from those in more academically oriented settings. Similarly, younger or less experienced teachers, potentially more familiar with digital tools and AI applications, may engage with these technologies in ways that differ from more experienced educators, who may be more sensitive to relational disruptions and the complexities of situated educational contexts.

Furthermore, while this study discussed AI in broad terms, it did not distinguish between the implications of different types of AI systems. Future work could usefully

examine whether co-creative AI tools, which invite students to generate, reflect, and revise, better foster agency than prescriptive systems that may automate decisions and narrow learner autonomy.

Regulatory frameworks to support agency-centred integration of AI into education were also beyond the primary focus of this study. Nonetheless, the findings highlight the absence of such frameworks as a critical gap, which future research should address, particularly in designing participatory and ethical approaches to AI governance in education.

Finally, although this study draws on a triangulated Educational Design Research methodology involving a systematic literature review, a Delphi study, and focus groups, future stages, including the piloting and wider implementation of a professional development course, will offer further opportunities for empirical triangulation, refinement, and extension of the frameworks developed here. The findings are intended to inform the design of a framework on human agency and AI systems, serving as a foundation for developing sections of a professional development course for teachers on these topics. Furthermore, there is a pressing need for intervention studies focused on developing pedagogical sequences that integrate AI while maintaining the core values of education across all three levels of agency. Future research could also explore longitudinal studies examining the long-term effects of AI on agency dynamics across diverse educational settings, building on the recommendations outlined in this paper, as well as drawing from prior neuropsychological research on sense of control, intentional binding, and agency, alongside educational studies focused on agency development.

As we continue to explore the integration of evolving AI sociotechnical systems in education, we are called to strengthen distributed forms of agency, where potential and challenge coexist in a balanced manner, ensuring that collectives preserve their roles as critical actors while broadening their aspirations and renewing their capacity to desire and to hope as they move into the future.

Author Contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ana Mouta. The first draft of the manuscript was written by Ana Mouta and all authors commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open access funding provided by FEDER European Funds and the Junta de Castilla y León under the Research and Innovation Strategy for Smart Specialization (RIS3) of Castilla y León 2021–2027. Open access funding provided by FEDER European Funds and the Junta de Castilla y León under the Research and Innovation Strategy for Smart Specialization (RIS3) of Castilla y León 2021–2027. The research leading to these results received funding from Fundação para a Ciência e Tecnologia under the grant DOI: <https://doi.org/10.54499/2020.06258.BD>.

Data Availability The datasets generated during and/or analysed during the current study are not publicly available to protect study participant privacy but are available from the corresponding author on reasonable request.

Declarations

Competing Interests All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Archer, M. S. (2003). *Structure, agency and the internal conversation*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139087315>
- Archer, M. S. (2015). How agency is transformed in the course of social transformation: Don't forget the double morphogenesis. In M. S. Archer (Ed.), *Generative mechanisms transforming late modernity* (pp. 135–157). Springer. https://doi.org/10.1007/978-3-319-10788-6_6
- Bandura, A. (1999). Moral disengagement in the perpetration of inhumanities. *Personality and Social Psychology Review*, 3(3), 193–209. https://doi.org/10.1207/s15327957pspr0303_3
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1, 164–180. <https://doi.org/10.1111/j.1745-6916.2006.00011.x>
- Bennett, J. (2010). *Vibrant matter: A political ecology of things*. Duke University Press.
- Berberian, B., Sarrazin, J. C., Le Blaye, P., & Haggard, P. (2012). Automation technology and sense of control: A window on human agency. *PLoS One*, 7, e34075. <https://doi.org/10.1371/journal.pone.0034075>
- Biesta, G. (2020). Risking ourselves in education: Qualification, socialization, and subjectification revisited. *Educational Theory*, 70(1), 89–104. <https://doi.org/10.1111/edth.12411>
- Biesta, G., & Tedder, M. (2006). *How is agency possible? Towards an ecological understanding of agency-as-achievement* (Working Paper Five). Learning Lives: Learning, Identity, and Agency in the Life Course. Exeter: Teaching and Learning Research Programme.
- Bourdieu, P. (1990). *The logic of practice*. Stanford University Press.
- Davis-Keen, P. E. (2005). The influence of parent education and family income on child achievement: The indirect role of parental expectations and the home environment. *Journal of Family Psychology*, 19, 294–304. <https://doi.org/10.1037/0893-3200.19.2.294>
- Deleuze, G., & Guattari, F. (1980). *A thousand plateaus: Capitalism and schizophrenia*. University of Minnesota Press.
- Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. Macmillan.
- Ellul, J. (1954). *The Technological Society*. Vintage Books.
- Emirbayer, M., & Mische, A. (1998). What is agency? *American Journal of Sociology*, 103(4), 962–1023. <https://doi.org/10.1086/231294>
- European Union. (2018). *Supporting school innovation across Europe*. Publications Office of the European Union. <https://doi.org/10.2766/466312>
- Foucault, M. (1975). *Discipline and punish: The birth of the prison*. Vintage Books.
- Giddens, A. (1991). *Modernity and Self-identity: Self and society in the late modern age*. Stanford University Press.

- Gozli, D. (2019). Sense of Agency. In D. Gozli (Ed.), *Experimental Psychology and Human Agency* (pp.137–164). Springer, Cham. https://doi.org/10.1007/978-3-030-20422-8_7
- Guichard, J. (2009). Self-constructing. *Journal of Vocational Behavior*, 75(3), 251–258. <https://doi.org/10.1016/j.jvb.2009.03.004>
- Haggard, P., & Tsakiris, M. (2009). Peripersonal space: A psychological perspective on the multisensory aspects of the body. *Psychological Bulletin*, 135(3), 440–464. <https://doi.org/10.1007/s00221-022-06324-8>
- Haraway, D. (1985). A cyborg manifesto: Science, technology, and Socialist-Feminism in the late twentieth century. *Simians, cyborgs, and women: The reinvention of nature* (pp. 149–181). Routledge.
- Haraway, D. J. (2016). *Staying with the trouble: Making kin in the Chthulucene*. Duke University Press.
- Jones, L. (2007). *The Student-Centered classroom*. Cambridge University Press.
- Kant, I. (2005). *A metafísica dos costumes*. Fundação Calouste Gulbenkian.
- Khosravi, H., Buckingham Shum, S., Chen, G., Conati, C., Tsai, Y. S., Kay, J., Knight, S., Martinez-Maldonado, R., Sadiq, S., & Gašević, D. (2022). Explainable Artificial Intelligence in education. *Computers and Education: Artificial Intelligence*, 3, 100074. <https://doi.org/10.1016/j.caeai.2022.100074>
- Larreameydy, J. (2011). Aprendizaje Como reconfiguración de Agencia. *Revista De Estudios Sociales*, 40, 33–43. <https://doi.org/10.7440/res40.2011.04>
- Latour, B. (2005). *Reassembling the social: An introduction to Actor-Network-Theory*. Oxford University Press.
- Latour, B. (2014). Agency at the time of the anthropocene. *New Literary History*, 45, 1–18. <https://doi.org/10.1353/nlh.2014.0003>
- Latour, B. (2018). *Down to Earth: Politics in the new climatic regime* (C. Porter, Trans.). Polity Press.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.
- Marx, K., & Engels, F. (1976). In C. J. Arthur (Ed.), *The German ideology*. International.
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education— Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. <https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- Moore, J. W. (2016). What is the sense of agency and why does it matter? *Frontiers in Psychology*, 7, 1272. <https://doi.org/10.3389/fpsyg.2016.01272>
- Mousavinasab, E., Zarifasanaiey, N., Kalhori, N., Rakhshan, R. S., Keikha, M., L., & Ghazi Saeedi, M. (2018). Intelligent tutoring systems: A systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, 29(1), 142–163. <https://doi.org/10.1080/10494820.2018.1558257>
- Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2023). Design of a future scenarios toolkit for an ethical implementation of Artificial Intelligence in education. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12229-y>
- Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2024). Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12946-y>
- Obhi, S., S., & Hall, P. (2011). Sense of agency in joint action: Influence of human and computer co-actors. *Experimental Brain Research*, 211(3–4), 663–670. <https://doi.org/10.1007/s00221-011-2662-7>
- OECD (2012). *PISA 2012 Database*. <http://www.oecd.org/pisa/pisaproducts/pisa2012database-downloadabledataba.htm>
- Pasquinelli, M. (2020). *The eye of the master: A social history of Artificial Intelligence*. Verso.
- Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury.
- Rogers, C. R. (1983). *Freedom to Learn for the 80's*. Charles E. Merrill Publishing Company.
- Saltman, K. J. (2020). Artificial intelligence and the technological turn of public education privatization: In defense of Democratic education. *London Review of Education*, 18(2), 196–208. <https://doi.org/10.14324/LRE.18.2.04>
- Sartre, J. P. (1943). *Being and Nothingness: A Phenomenological Essay on Ontology* (H. E. Barnes, Trans.). Washington Square Press.
- Selwyn, N. (2016). *Education and technology: Key issues and debates* (2nd ed.). Bloomsbury Academic.
- Stiegler, B. (2013). *Pharmacologie du Front National: Suivi du vocabulaire d'Ars Industrialis* (V. Petit, Ed.). Flammarion.
- Stiegler, B. (2015). *States of shock: Stupidity and knowledge in the 21st century*. Polity. D. Ross.

- U.S. Department of Education, Office of Educational Technology. (2023). *Artificial intelligence and the future of teaching and learning: Insights and recommendations*. U.S. Department of Education. <https://www.ed.gov/sites/default/files/documents/ai-report/ai-report.pdf>
- UNESCO (2021). *Reimagining our futures together: A new social contract for education. students*. Retrieved February 2, 2023, from <https://creativecommons.org/licenses/by-sa/3.0/igo/>
- Vera, M., Le Blanc, P., Taris, T., & Salanova, M. (2014). Patterns of engagement: The relationship between efficacy beliefs and task engagement at The individual versus collective level. *Journal of Applied Social Psychology, 44*, 133–144. <https://doi.org/10.1111/jasp.12219>
- Wang, H., Tlili, A., Huang, R., Cai, Z., Li, M., Cheng, Z., Yang, D., Li, M., Zhu, X., & Fei, C. (2023). Examining the applications of intelligent tutoring systems in real educational contexts: A systematic literature review from the social experiment perspective. *Education and information technologies*, 1–36. Advance online publication. <https://doi.org/10.1007/s10639-022-11555-x>
- Watters, A. (2021). *Teaching machines: The history of personalized learning*. The MIT Press.
- Williamson, B. (2022). Big EdTech. *Learning Media and Technology, 47*(2), 157–162. <https://doi.org/10.1080/17439884.2022.2063888>
- Zuboff, S. (2019). *The Age of Surveillance Capitalism: The Fight for a Human Future at the New Frontier of Power*. Public Affairs.
- Zumeta, L., Oriol, X., Telletxea, S., Amutio, A., & Basabe, N. (2016). Collective efficacy in sports and physical activities: Perceived emotional synchrony and shared flow. *Frontiers in Psychology, 6*(1960), 1–12. <https://doi.org/10.3389/fpsyg.2015.01960>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Ana Mouta¹  · Ana María Pinto-Llorente²  · Eva María Torrecilla-Sánchez² 

✉ Ana Mouta
mouta.ana@usal.es

¹ Programme in Education in the Knowledge Society, University of Salamanca, P.º de Canalejas, 169, Salamanca, España 37008, Spain

² Instituto Universitario de Ciencias de La Educación (IUCE), University of Salamanca, P.º de Canalejas, 169, Salamanca, España 37008, Spain

APPENDIX D

PUBLISHED MANUSCRIPT IN EDUCATION AND INFORMATION TECHNOLOGIES

Mouta, A., Torrecilla-Sánchez, E.M. & Pinto-Llorente, A.M. (2025). Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-024-12946-y>

Journal Quality Indicators:

JCR – Q1. Journal Impact Factor (2024): 5.4. JIF Rank: 22/756. Percentile: 97.2%

SJR – Q1. SJR (2024): 1.654. H-Index: 97



Comprehensive professional learning for teacher agency in addressing ethical challenges of AIED: Insights from educational design research

Ana Mouta¹ · Eva María Torrecilla-Sánchez² · Ana María Pinto-Llorente²

Received: 9 April 2024 / Accepted: 23 July 2024
© The Author(s) 2024

Abstract

Continuing professional development plays a pivotal role in creating opportunities for teachers to explore the evolving educational landscape. With the integration of Artificial Intelligence into education, these opportunities involve grasping teachers' attitudes, expectations, and pedagogical approaches, with a focus on ethical considerations. Nevertheless, existing research and professional learning opportunities often overlook the perspectives of educators on these themes. Aiming to bridge this gap, this Educational Design Research approach began with a systematic literature review, followed by a Delphi study to gather educational stakeholders' insights on the ethical concerns of using AI in education. The current study presents the research third phase. It explores findings from focus groups with educators responsible for K-12 teacher education, informing the design of a training programme that addresses ethical concerns and agency. Four groups were conducted using a semi-structured script, centred on pre-provided scenarios. Data analysis involved thematic coding using both deductive and inductive approaches, revealing key themes related to: employing AI applications in course delivery for a mindful, hands-on exploration; discussing ethical and policy frameworks with a focus on contextual needs and *technogeographies*; addressing uncertainty, resistance, and transition; and fostering individual and collective agency regarding ethical issues through informal learning channels to build a nuanced narrative that challenges the corporate one. The study concludes by highlighting the importance of greater investment in professional development to enable educators to critically assess and reshape the values associated with education in the context of Artificial Intelligence, thereby contributing to the movement of aligning AI with our humanity.

Keywords Ethics · Agency · Continuing professional development · Training · Artificial Intelligence in Education · Teachers

Extended author information available on the last page of the article

Published online: 15 August 2024

Springer

1 Introduction

1.1 AI technologies for education, challenges, and ethical considerations

In recent years, Artificial Intelligence in Education (AIED) has made significant strides, with proponents asserting the potential for personalised learning experiences for students, increased efficiency for teachers, and automation of administrative tasks in schools (Wang et al., 2023). Adaptive learning systems and Intelligent tutoring systems (ITS) are often touted for providing tailored support and interactive learning experiences, ultimately enhancing overall learning outcomes (Mousavinasab et al., 2018). Platforms can integrate teachers' assessment standards, while AI agents assist in pre-service training with subject-specific scenarios and pedagogical tips tailored to future educators' proficiency levels (Liu, 2023). AI assistants can also facilitate group projects, assigning roles, guiding discussions, tracking progress, and promoting equal participation. Virtual facilitators based on Natural Language Processing (NLP) can be used to develop professional learning opportunities for teachers (Copur-Gencturk et al., 2024). Additionally, AI-powered chatbots, like Andy English Chatbot, are proving to be valuable tools in language learning, providing learners with a low-pressure environment to practice speaking and receive feedback without fear of judgment, promoting fluency and self-confidence (Fathi et al., 2024). AI-driven language learning platforms like Duolingo have demonstrated their significance in lifelong learning of a second language, supported by evidence from both written and oral proficiency assessments characterised by high validity and practicality measures (Sudina & Plonsky, 2024).

Albeit the potential benefits these technologies offer, concerns have surfaced regarding their potential drawbacks. For instance, Learning Analytics platforms can provide insights into student engagement, performance trends, and learning trajectories, with descriptive and predictive value. Nevertheless, the Council of Europe (2022) raises ethical concerns regarding the potential impact of these technologies on student and teacher agency. In fact, research suggests that interactions with automated artificial systems can diminish human operators' sense of control (Berberian et al., 2012; Obhi and Hall, 2011). Additionally, ethical considerations surround the collection of private and sensitive data such as mood analysis and activity logs concerning political views, ethnic identity, and sexual orientation (Tundrea, 2020).

NLP tools, while providing feedback on grammar, syntax, and coherence, may prioritise quantity over content quality and exhibit unreliability in error detection (Liu & Kunnan, 2016; Miao et al., 2021). Other AI tools for mathematics, like Photomath, offer opportunities for students to engage with mathematical concepts. However, they can also pose challenges in interpreting the information provided. This difficulty can stem from users' varying levels of prior mathematical knowledge (Gaona et al., 2022), potentially impacting their self-concept in relation to mathematics. Ensuring transparency and explainability in AI systems is key to uphold ethical principles in education. If educators and students cannot understand how these systems reach decisions, it can detach them from the learning experience.

Moreover, ITS have faced scrutiny concerning learners with diverse backgrounds or special needs (Alrakhawi et al., 2023). Indeed, there are apprehensions that these technologies may exacerbate existing educational disparities and expose students to bias and cultural barriers (Miao et al., 2021), impacting their learning processes and assessment. Pedagogical agents, for example, while aiming to enhance learning experiences by simulating human-like interactions and positive emotions, raise concerns about affective privacy and virtual relationships (Hudlicka, 2016). Ultimately, ethical challenges may encompass competing interests and extend beyond individual users to encompass their families, and discussions on these matters often reflect the perspectives of stakeholders, which can vary depending on their affiliations with private organisations, government agencies, research centres, universities, or schools (Popenici & Kerr, 2017).

Drawing on research, a comprehensive examination of AIED studies from 1970 to 2020 highlights the insufficient focus on ethics in the implementation of these technologies and the overlooking of viewpoints from teachers, parents, and students, who could provide valuable insights for ethical considerations (Bozkurt et al., 2021). This is consistent with the findings of a UNESCO report, which revealed that although 11 Member States have endorsed AI curricula, some either do not engage or engage only to a limited extent with AI ethics and its social implications (UNESCO, 2022).

Utilising Biesta's (2020) principles of subjectification, socialisation, and qualification, which are key to delineating educational endeavours, it becomes crucial to understand the educational objectives advocated by AIED and how they balance these principles. At a pedagogical level, certain AI applications may not align with socio-constructivist learning frameworks, resulting in a disjointed integration of technology into formal education settings. While some AIED resources are marketed as personalised learning tools, they often provide individualised pathways that ultimately converge toward a common goal, not considering the variance of the impact they may have on singularities and their ecosystems. Both scenarios can be argued to be a deliberate preference for mass education influenced by political and economic considerations, undermining the processes of subjectification and qualification. Are these AI-driven resources designed and used in a way that upholds the underlying foundation of qualifications, which includes self-efficacy? How can they be used to support the core properties of human agency: forethought, self-reactiveness and self-reflection (Bandura, 2023)? Regarding socialisation, a study by Hrastinski et al. (2019) highlights a crucial ethical concern surrounding AIED – its potential impact on socialisation within the classroom. The authors discussed concerns among teachers, researchers, and pedagogical developers related to shifts in the teacher's role and teacher-student relationships due to AI, largely emphasising the importance of teacher professional development for a relevant implementation of AIED in K-12 classrooms.

1.2 Ethical and meaningful pedagogical standards through continuing professional development

The significance of Continuing Professional Development (CPD) for teachers in enriching pedagogical practices is well-established (Abakah et al., 2022;

Darling-Hammond et al., 2017; UNESCO, 2015). CPD tends to foster heightened expectations that indirectly influence student achievement (Rubie-Davies, 2015) and holistic development by enhancing teachers' capacity for utilising evidence-based pedagogical methods, designing intentional and meaningful learning experiences, and implementing relevant assessment techniques. It can lead to experiences of rupture with methodological normativity and reaffirm the inevitability of pedagogical mediation (Trindade, 2012). Teachers involved in CPD can also serve as role models for their students and peers, promoting the value of learning, knowledge sharing, and critical thinking within learning communities. It may also enhance teachers' influence in decision-making processes, improving their decisional capital (Hargreaves & Fullan, 2012). Additionally, CPD may equip teachers with the ability to adapt to changing contexts, address the evolving needs of their students more confidently, meet societal expectations, and be accountable to the entire school community (Learning Forward, 2022). This seems particularly relevant for teachers considering technological advancements and their influence on students' lives.

Following an Ellulian perspective, CPD can serve as a catalyst for questioning the underlying mindset that governs our society and for the critical examination of the purposes and dynamics of technique, assisting educators in comprehending and navigating the intricate interplay between technology, pedagogy, and the broader societal context. Only an approach responsive to the challenges posed not only by AI technologies, but by a world driven by technique, can reveal how its dominance has permeated all layers of society and institutions, extending even into education. Such an approach can redirect discussions on educational ethics far beyond superficial debates centred solely on AI ethics.

Kowalczyk-Waledziak et al., (2019) also assert that, given the intersection of ethical considerations with pedagogical approaches, there is a need to reconsider teacher education. This could be done through a participatory approach involving educators, to make meaningful progress in pedagogical terms (Chichekian & Benteux, 2022). While UNESCO (2023a) recommends governmental agencies to build capacity for the proper use of Generative AI (GenAI) in education and research, a systematic literature review exploring the ethical dimensions of AIED technology implementation in education from 2011 to 2022 revealed a noticeable absence of emphasis on ethical considerations, despite some training projects being introduced in schools (Mouta et al., 2023a). However, more recently, some initiatives are currently underway to address these concerns and requirements. Set to launch in 2024, UNESCO's AI competency framework for educators takes a step further by outlining a three-stage progression that covers areas such as fostering a human-centric approach, addressing AI ethics, understanding AI fundamentals and applications, incorporating AI into pedagogy, and utilising AI for professional growth (UNESCO, 2023b). A training practical example comes from AI4T (2024), an Erasmus+K3 initiative designed by France, Slovenia, Italy, Ireland, and Luxembourg, which emerged to facilitate AI education specifically for teachers and school leaders. Understanding the ethical implications of employing AI in educational environments was included as a component of the shared learning objectives across all countries' professional development pathways.

1.3 Teacher agency

The dominance of a world of technique (the pre-eminence of rationality and effectiveness), the increasing diversity of students, the necessity to prepare learners for environmental transitions and to critically examine what development might mean, and the attempt to enhance learning and assessment through digital means greatly influence the direction of education. Thus, it becomes central to incorporate the needs of students and society into an andragogic narrative that specifically caters to teachers, recognising them as active contributors in co-authoring the educational plot. Intelligent professionalism suggests educators to be consistently involved in the processes that educational systems claim are enhancing education (Thompson, 2021). This is expected to come alongside the commitment to change within schools and among teachers (Kukulska-Hulme et al., 2020) and the recognition of families and community members as vital partners (European Union 2018). Being a teacher involves taking a public stance on important educational issues and public policy development, elevating the professional status across all systems (Thompson, 2021; UNESCO, 2021). Such an approach to professional development can strengthen teachers' shared epistemic agency (Damşa, 2014) and reinforce ethical practices related to AIED knowledge. Moreover, it can promote teachers' relational and collective agency, reassuring the structural context supports and facilitates the interdependence, while understanding how agency unfolds over time (Damşa, 2014). This perspective aligns more closely with an ecological approach advocated by Priestley et al. (2015), which suggests that teacher agency isn't an inherent attribute of individuals but rather a result of the interaction between personal capabilities and the environmental factors that influence their actions.

Ultimately, teacher agency is recognised as pivotal, as educators possess the ability to actively approve, adapt, or resist policies and programmes (Severance et al., 2016). While it is true that numerous decisions regarding educational technologies are often made from the top-down, involving government bodies, big tech companies, national research, education networks, and appointed staff from learning institutions, the ultimate realisation of these technological initiatives heavily relies on teachers themselves. Their awareness, knowledge, and, most importantly, their attitudes and level of engagement play a key role. To inform these aspects, training may be significant. Once educators are active agents in designing, implementing, evaluating, advocating for, and experiencing quality professional learning and the systems that support it (Cochran-Smith et al., 2022; Learning Forward, 2022), training initiatives built upon their knowledge and deliberation may yield meaningful investments for schools. Teachers are also expected to undergo training centred on the same principles of agency they are to instil in their students, enabling learners to exercise it when interacting with autonomous AI agents.

In this context, this paper presents a specific segment of a participatory research aimed at strengthening teachers' agency in training related to ethics and AIED. It is rooted in the theoretical underpinning of dialogic ethics, which acknowledges the multifaceted nature of ethical concerns and the need for a thorough examination involving various stakeholders (Santos, 2012). While existing research has explored the theoretical aspects of AIED ethics, it often overlooks the perspectives

of educators themselves. This study distinguishes itself by focusing on educators' voices through the lens of Continuing Professional Development. Specifically, the gap in participatory studies in this field is addressed by utilising focus groups with those responsible for shaping teacher education.

2 Method

This component is part of a comprehensive Educational Design Research, which began with a systematic literature review (SLR) to pinpoint major gaps in AIED and ethics (Mouta et al., 2023a). The findings from the SLR guided subsequent research steps, including the recognition of the need for the involvement of stakeholders from diverse educational backgrounds in research and for capacity building in AIED and education ethics, especially among teachers. As a result, the Delphi method was employed to investigate educational experts' ethical concerns regarding AI in education (Mouta et al., 2023b). With the goal of encouraging comprehensive debates on the development of professional learning in AIED, education ethics, and agency, these concerns informed the design of futures scenarios to facilitate in-depth focus groups discussions. Focus groups are commonly employed in educational research because of their usefulness for programme development (Nagle & Williams, 2013) and their ability to yield valuable data to shape policies and practices. This section delves into the initial inputs arising from those focus group discussions among professors, teachers, and trainers regarding the attributes of such a training. While the development of the professional learning materials is planned for later stages, this phase provides insights and discussion around the course's objectives, implementation strategies, and design aspects. This study's third stage is the main focus of the upcoming sections and is structured according to the framework proposed by Silva et al. (2014), which includes planning, preparation, moderation (referred to here as "implementation"), data analysis, and results communication.

2.1 Planning

The planning phase of focus group research is key to ensuring that clear objectives guide the moderation process.

2.1.1 Research guiding questions

To inform the design of a meaningful training programme for K-12 educators on AIED, ethics, and agency, this stage of the study delves into the following key research questions (RQ):

- RQ 1: Which structuring dimensions do educators consider important for a training programme on AIED, ethics, and agency targeting K-12 pre-service and in-service teachers?
- RQ 2: What objectives do educators prioritise for such a training programme?

- RQ3: What specific aspects should be taken into account when developing a training programme in this field?
- RQ 4: What strategies can be used to update teachers' knowledge and skills on AIED while also fostering agency and ethical thinking about its potential impact on education?

2.1.2 The moderation scripts

To facilitate participants' reflections and deepen their responses, it is recommended to use a moderation script with a "funnel" strategy, where questions become increasingly specific throughout the session (Borges & Santos, 2005). While the main objective of these focus groups was to identify the goals, topics, and delivery methods for a training programme on ethical AIED use for teachers, some initial questions on AIED ethics were included to familiarise participants with the topic under discussion. In advance of the session, each participant was provided with two hypothetical scenarios that had been generated during the Delphi phase of this research (Mouta et al., 2023b). The aim was to motivate the participants for the discussions by providing these narratives. It is worth noting that these scenarios resulted from a three-stage iterative process, leading to their creation in 2021 for discussions held in 2022. AI applications to the field of education have progressed significantly since then, as emphasised in the introduction. Despite this, the scenarios remain relevant for the purpose of this research. However, their dystopian aspects and the potential shift towards vignettes illustrating current experimental cases during the training will be investigated in the discussion section.

The scenarios presented to the participants were the same for each element within a group, but they differed between groups. The script utilised during the session featured open-ended questions on participant attitudes towards AIED, ethical concerns, student agency, and the design of an educator training programme within the context of AIED (see Appendix 1).

2.1.3 The participants recruitment

To prevent dispersion, a relatively uniform group was established, selecting participants according to criteria that corresponded to the research objectives (Silva et al., 2014). For this study, participants with professional certification for teaching and experience in the field of education at any level were eligible for selection. Additionally, they were required to hold responsibilities in K-12 teacher pre-service or in-service education. Participants also had to be either native Spanish speakers or highly proficient in the language. This requirement was based in the primary language of the research centre where the study was conducted, facilitating the initial recruitment of participants through convenience sampling due to time constraints at the start of the research phase. However, to ensure diversity in cultural perspectives, a subsequent snowball sampling approach was employed. This involved efforts to purposively recruit participants from various geographical regions, thereby improving the generalisability of the study's findings. This approach balanced the need for active communication in the research language with the goal of capturing a broad

range of cultural insights and experiences related to the study objectives. Invitations were sent via email. A larger group than required was invited to tackle the challenge posed by potential refusals (Ogbeifun et al., 2016). Among those who declined the invitation, the explanation provided was either difficulty in participating due to professional commitments or a lack of confidence in engaging in discussions on such a topic. So, 19 participants from a pool of 40 invitations agreed to participate and were subsequently divided into four groups. Participants provided their informed consent, and the moderator explained the research objectives and ensured participants' positive responses (Silva & Fortunato, 2021). The mean age of the participants is 44 years old (SD 7.46) and 47.4% identify as women and 52.6% as men. The group comprises individuals from five countries, all of whom hold responsibilities in preparing pre-service or in-service teachers in the K-12 education sector. They possess pedagogical experience in either K-12 or higher education. Participants' general sociodemographic and occupational profiles are presented in Table 1.

In each group, there was an in-depth discussion of two scenarios selected from the set of hypothetical scenarios generated in the previous Delphi phase, making a total of eight vignettes analysed across the four groups. Hennink et al. (2019) proposed that the appropriate sample size to reach saturation in focus group research depends on the study purpose, code characteristics, group composition, desired type/degree of saturation. These focus groups usually consist of three to six sessions, with a larger number of sessions when accommodating a diverse study population (Coenen et al., 2012; Guest et al., 2016). Therefore, in this research code and meaning saturation were achieved through four discussions, aligning with the desired saturation levels for focus groups not stratified by demographic characteristics.

2.2 Preparation

Participants were notified through emails, which contained details about the research objectives, session duration, group specifics, focus group dynamics, session recording, and informed consent. The two hypothetical scenarios were dispatched 48 h prior to the session. Since the invitees originated from diverse regions and the world still faced pandemic restrictions, it was deemed more appropriate to conduct virtual meetings, through the Zoom platform. Despite some limitations, research has shown that virtual focus groups can be just as productive as in-person focus groups in generating rich and meaningful data (Krueger & Casey, 2015).

2.3 Implementation

Four focus groups were implemented in the first semester of 2022, with varying numbers of participants (6, 4, 5, and 4) to accommodate their availability.

The participants were unfamiliar with each other, except for two individuals in the first and in the second groups. In both cases, their professional roles were independent from each other. Across all sessions, there was a notable absence of dominant speakers. All participants tried to ensure everyone had an opportunity to share their thoughts, mitigating power differentials and ensuring that diverse viewpoints

Table 1 Teachers' sociodemographic profile
Focus group participants' profiles

Group no. 1	Professional Profile	Age range			Gender		Working Country
		25–35	36–45	46–55	F	M	
		Director of Digital Transformation in Higher Education (Tecnológico de Monterrey)		X			
Associate Professor in Research Methods and Diagnosis in Education (University Salamanca)	X				X	Spain	
Former k-12 teacher. Director of the Ministry of Education's Master Plan			X		X	El Salvador	
Associate Professor of Research Methods in Education (University León)			X		X	Spain	
Former k-12 teacher. Manager of Academic Development for Higher Education at the Ministry of Education			X		X	El Salvador	
Former K-12 teacher. Researcher in Teaching of Mathematics in K-12 education		X				El Salvador	
GROUP no. 2							
Associate Professor (University Salamanca). Coordinator of Primary Education Degree		X			X	Spain	
k-12 teacher. Director of the Fontán School Network			X		X	Colombia	
k-12 teacher. Pedagogical Coordinator at Fontán School		X			X	Colombia	
Teacher Trainer. Associate Professor in Information Technology (University El Salvador)		X			X	El Salvador	
GROUP no. 3							
Full Professor in Computer Systems (University Salamanca). Academic Director of Quality Assessment				X	X	Spain	
Full Professor in Research and Diagnostic Methods in Education (University Salamanca)		X			X	Spain	
Associate Professor in Didactics of Experimental and Social Sciences (University Valencia)		X			X	Spain	
Former k-12 teacher. Assistant Professor in Primary Education (Fray Luis de León University School)		X			X	Spain	
Teacher trainer and project Manager (University Salamanca)				X	X	Spain	
GROUP no. 4							
Assistant Professor in Educational Theory (University Salamanca)				X	X	Spain	

Table 1 (continued)

Focus group participants' profiles

Group no. 1	Professional Profile	Age range			Gender		Working Country
		25–35	36–45	46–55	F	M	
	Associate Professor in Didactics and School Organisation (University Santiago Compostela)		X			X	Spain
	Full Professor in Research and Diagnostic Methods in Education (University Oviedo)			X		X	Spain
	K-12 teacher trainer		X			X	Portugal

were thoroughly considered. The duration of the focus group discussions varied: the first session lasted 80 min, the second and fourth sessions lasted 64 min each, and the third session lasted 62 min.

2.4 Data analysis

In accordance with Roberts et al.'s (2019) proposal, the deductive phase of this study included creating a preliminary codebook, aiming to ensure credibility through three key elements: content (describing the sampling frame and creating research instruments), criterion-relatedness (testing research tools for consistency), and construct validity (ensuring inferences match the research question) (Long & Johnson, 2000). The subsequent sections detail the methodology, which includes iterative coding, category refinement, and reliability and validity testing.

All sessions were recorded using the recording option in Zoom, after obtaining email and verbal consent from all participants. The recordings facilitated the verbatim transcription of the session's content using an AI-powered software (Trint). All transcriptions were then thoroughly reviewed and cross-checked against the video recordings by the principal researcher. The analysis was based on the research question being asked ("How should a training programme targeting ethics in AIED for K-12 teachers or aspiring teachers be structured and organised?"), the initial analysis of the literature, the Delphi method approach undertaken as part of the project, and a preliminary scan of the raw data. Additionally, an inductive approach was employed, which allowed for the identification of any unexpected themes that emerged during the coding process. So, the preliminary data analysis entailed the researcher's close examination of the raw data, selectively marking segments for coding. The entire dataset, transcribed from audio recordings, was analysed using NVivo software. Moderator contributions were transcribed but not included in the analysis. The research team agreed on unit separation criteria upon importing the data into the software. A thematic approach was selected, taking into account the specific content associated with each element of the text. The structural matrix, employed for the coding of the data, was derived from both the extant literature and the aforementioned initial reading and identification of preliminary themes from the raw data. In cases where a significant proportion of the codable units could not be accurately captured by the existing codes, further codes were added to the analytical framework. This process was repeated multiple times until no new codes were identified, thereby confirming the matrix as a valid representation of the data. The identified codes were added to NVivo as nodes, and the coded text was matched to the nodes in a systematic way. As detailed in Fig. 1, the combination of deductive and inductive approaches enabled the study to identify patterns both within and outside of the predetermined codes, resulting in a comprehensive and nuanced analysis of the research question.

The final selection of nodes resulted from two reliability tests: a test-retest reliability and an inter-rater/coder analysis (*cf.* https://drive.google.com/drive/folders/13fF4h6fGbrwf9woDKApJ7NulfoeNzLEr?usp=drive_link), conducted by two researchers. The Pearson correlation coefficient between the initial coding and the

coding performed three weeks later was 0.86, indicating high consistency in the coding process over time. The percentage agreement between the coders was 81%, demonstrating the extent to which they agreed on coding decisions. The Cohen's Kappa score was 0.83, signifying substantial agreement beyond chance across different coders. This analysis allowed for the examination of information using a framework matrix consisting of four main categories with its own set of subcategories: attitudes towards AIED, benefits and opportunities, ethical challenges, teacher initial and CPD; plus, one category labelled as "other" was added to account for text segments that were tangential and not directly related to the ongoing discussion (cf. Table 2). The description of each node and respective examples from categorisation in NVivo is also available (see Appendix 2).

Instances arose where information exhibited overlapping characteristics, being assigned to multiple categories simultaneously, or left uncoded if it did not align with a specific category. For this analysis, categories encompassing all the gathered information were established, although only specifically studying the data directly relevant to the main research question. In fact, the initial open-ended questions were designed to prepare participants for the discussion. Since the topic under consideration is relatively new or often overlooked in current educational discourse, these questions aimed to provide the necessary context for discussing training methods. The need to code such information stems from the fact that coding ensures that no data is overlooked and that all related aspects are taken into account during the analysis process. However, for the purpose of this analysis, a deliberate decision was made to analyse and present the data that directly addresses teacher training in ethics in AIED.

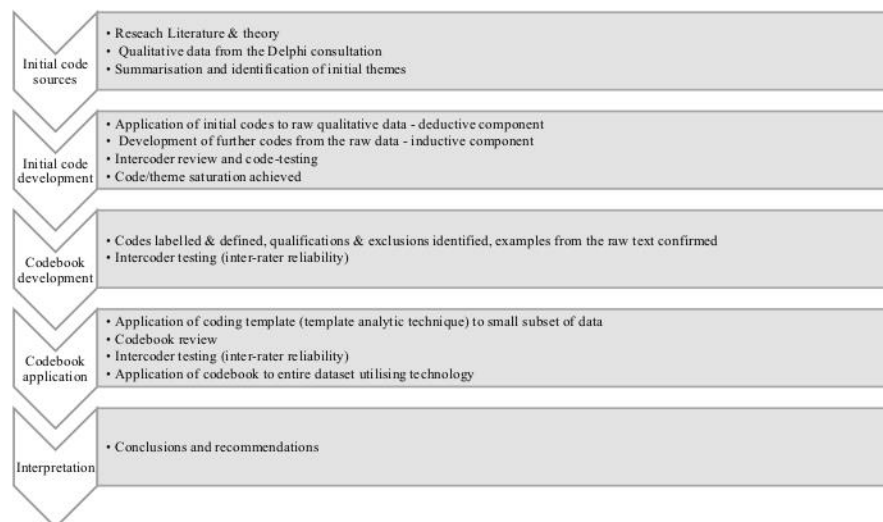


Fig. 1 Process of code creation and testing. Note. The process of code creation and testing was adapted from Roberts, K., Dowell, A., & Nie, J. B. (2019) Attempting rigour and replicability in the thematic analysis of qualitative research data; a case study of codebook development. *BMC Medical Research Methodology*, 19, 66. <https://doi.org/https://doi.org/10.1186/s12874-019-0707-y>

Education and Information Technologies

Table 2 NVivo framework matrix

Node name	
A. Attitudes towards AIED	C6. Education as a system
A1. Affective attitudes	C7. Examples
A1.1. Negative	C8. Families
A1.2. Positive	C9. Future of Society and Civilisation
A2. Cognitive Attitudes	C10. Investing in projects, research, and theories
A2.1. AI tech as neutral tools	C11. Lifelong Learning
A2.2. People's Agency	C12. Problems with Data sets
A2.3. Technological determinism	C13. School Administration
B. Benefits and opportunities	C14. Students
B1. Citizenship and Equity	C15. Teachers
B2. Education as a system	C16. Work and Employability
B3. Examples	D. Teacher initial and continuing professional development
B4. Families	D1. Broadened ethical thinking
B5. Future of Society and Civilisation	D2. Collaboration between teachers and with families
B6. Lifelong Learning	D3. Compliance with a policy framework
B7. People in general	D4. Dealing with resistance and transition
B8. School Administration	D5. Hands-on and skill-based training
B9. Students	D6. Informal self-directed CPD
B10. Teachers	D7. Innovative and purposeful pedagogical practices
B11. Work and Employability	D8. Learning by example
Ethical challenges	D9. New role requirements
A. C1. AIED tech design and developers	D10. Pointless without AI tech
C2. Citizenship and Equity	D11. Pre-Service Training
C3. Communities and cultural context	D12. Questioning Educational Paradigms
C4. Developing Frameworks and Policies	D13. Setting priorities
C5. Different interests between users and corporations	D14. Update of knowledge and skills
	E. Other

2.5 Results communication

This section presents the key findings from the focus group discussions in relation to the research questions guiding this stage of the project.

2.5.1 RQ1: Structuring dimensions for the training programme

Concerning the findings on teacher initial and Continuing Professional Development, the relationships between the coded text excerpts and their corresponding nodes are examined. This analysis uses the data shown in Fig. 2, which identifies the

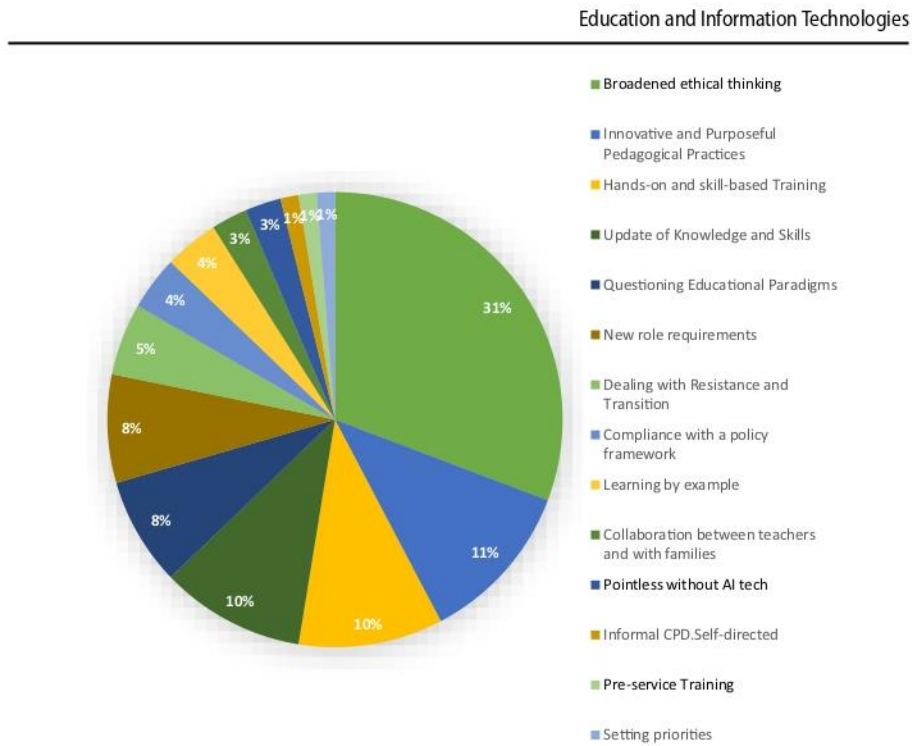


Fig. 2 Number of references per item

topics participants discussed the most. This figure provides an overview of the percentage frequency of each category mentioned across the entire dataset, highlighting its prominence in the overall discussions.

These discussions were further categorised based on different structuring dimensions of the training course. By analysing the arrangement of the nodes, it is possible to identify the following key structuring dimensions: training aims and objectives, design considerations, and implementation strategies. Table 3 provides the percentage of discussion dedicated to each category within these specific planning dimensions, offering insights into their relative prominence among them. For instance, while “broadened ethical thinking” constitutes 31% of the overall discussion dataset (cf. Figure 2), it comprises 51.06% of the references associated with the “aims/objectives” category in Table 3. This “aims/objectives” category was indeed the primary focus of the participants’ discussions, centring on the purpose of the training. With 60% of the discussions dedicated to appraising and explicitly articulating its objectives, it is clear that participants were committed to ensuring the course aligned with teachers’ needs.

To further analyse the relationships between these discussion topics, NVivo’s cluster similarity function was employed. This function measures similarity between nodes, considering the presence or absence of shared words between each pair of nodes. The results of this cluster analysis are displayed as a horizontal dendrogram (cf. Figure 3). The following paragraphs will discuss categories and subcategories, as well as the connections between them, highlighting those revealed by the analysis.

Table 3 Optimising teacher training: Organising items based on planning dimensions

Aims/Objectives Category	No	Design Category	No	Implementation Category	No
Broadened ethical thinking	51.06%	Dealing w/ Resistance and Transition	36.36%	Innovative and Purposeful Pedagogical Practices	45%
Update of Knowledge and Skills	17.02%	Collaboration between teachers and with families	18.18%	Hands-on and skill-based Training	40%
Questioning Educational Paradigms	12.77%	Pointless without AI tech	18.18%	Learning by example	15%
New role requirements	12.77%	Setting priorities	9.09%		
Compliance with a policy framework	6.38%	Pre-Service Training	9.09%		
		Informal CPD, Self-directed	9.09%		

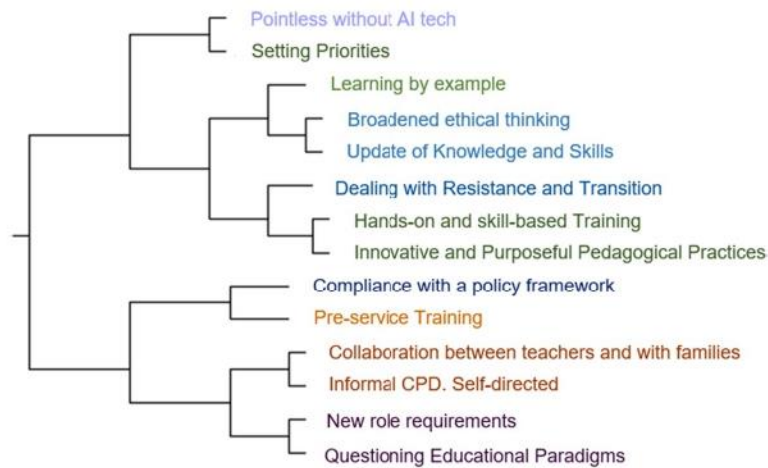


Fig. 3 Items clustered by coding similarity

2.5.2 RQ2: Aims and objectives prioritised by educators

Broadened ethical thinking The aims and objectives category accounted for a total of 47 references across five items. The emphasis on this category reveals a significant concern among the participants regarding the purpose and direction of an AIED ethics training course. Their main concern revolved around the necessity to foster broad ethical thinking; a topic extensively discussed in approximately 51% of the discourse segments. In this context, one of the participants expressed, "when I think about teacher training, I don't see it from a technician point of view (...) [it] is the space for reflection and construction, joint construction about what the challenges, the difficulties, the opportunities might be". This highlights the understanding that dealing with the ethical aspects of technological progress necessitates a holistic strategy that encompasses all disciplines. By cultivating a climate of ethical consciousness and accountability regarding technological advancements, individuals are given the opportunity to question the very ethical nature of the educational project.

Update of knowledge and skills Moreover, by recognising the significance of keeping knowledge and skills up to date, as evidenced by the subcategory that garnered 17% of the references, participants suggested that training is expected to question what is technically known, ethically debatable, and required to create a theoretical and conceptual rationale. The following excerpt illustrates this perspective: "we must perhaps have a significant ability to develop competencies for knowledge, to know the advances of science and their relationship with the benefits and problems that ethical aspects in sciences and technologies can generate for us". Participants also underscored the ever-changing nature of the field and the imperative for trainees to remain abreast of the latest advancements. In the context of CPD, offering training that aligns with teachers' current needs

while also anticipating and addressing upcoming changes and advancements may help them stay ahead of the curve. This dimension may also support educators to respond to any emotional discomfort that may arise from global shifts affecting their learning environments, whether due to emerging resources, technologies, or pedagogical trends. It may also encourage them to explore alternative approaches, experiment with emerging resources and help students engage meaningfully with the learning experience, keeping teachers motivated and satisfied in the long run.

The cluster similarity analysis in Fig. 3 highlights a connection between these two subcategories: broadened ethical thinking and knowledge/skills update. This association underscores that extensive ethical reasoning in AIED leads to the recognition of the importance of updating knowledge and skills in this field. Conversely, a strong knowledge base is crucial for informing ethical deliberations accurately.

Questioning educational paradigms As a result of this mindset and skills update, teachers will be instigated to question current educational paradigms, a subcategory that garnered 13% of the references. In this case, participants acknowledged the significance of motivating trainees to challenge the established norms and embrace a more comprehensive and ethical educational approach, facilitated by the dilemmas posed by this recent technological endeavour. As one of the participants stated, "It has to come accompanied (...) by a significant innovation (...) within the educational system. For example, it's about ending the concept of exams, the concept of a teacher, as we see it today".

New role requirements Once teachers begin to challenge current educational paradigms, their role in the classroom will also be challenged, as stated 13% of the times and as the cluster similarity analysis underscores (cf. Figure 3). This is exemplified by an excerpt from one participant: "becoming guides means positioning ourselves horizontally in the processes of development, in learning, in the construction process, in existence". So, training is expected to encourage teachers to re-evaluate the adequacy of their current roles and practices, critically question what it means to be a "guide" when integrating AI technologies, and understand data-driven insights in a holistic way.

Compliance with a policy framework Finally, participants recommended preparing teachers to comply with a policy framework, which accounted for approximately 6% of the recommendations, as expressed in this discourse: "teachers should indeed receive training (...) on the current legislation and how to apply it in the context of using applications that collect personal information, such as (...) artificial intelligence systems". This highlights their concern for achieving conformity once legal and regulatory guidelines are established. Furthermore, it underscores the significance of enhancing teachers' professional credibility and safeguarding students' rights by ensuring consistent practices across contexts.

2.5.3 RQ3: Design considerations for the training programme

The design category was the less mentioned, receiving a total of 11 mentions. However, the existence of six subcategories related to this topic suggests that the participants did differentiate their insights into the strategic and organisational aspects of the training programme.

Dealing with resistance and transition In 36% of the discourses, the subcategory dealing with resistance and transition was mentioned. It suggests a concern about how teachers might respond to the integration of these technologies in learning environments, due to their cognitive and affective attitudes or lack of familiarity with the theme of AIED. It encompasses resistance to change and managing transitions comprehensively. As one participant emphasised: “I see what I call the transition risk, meaning, in the initial stages, the first teachers may not adopt these new technologies. There will be a transitional phase”. Therefore, it is proposed that the training tackles concerns related to a smooth and contextually meaningful transition. The dendrogram’s analysis offers insights on this matter, indicating the design category’s proximity to two implementation categories (cf. Figure 3). It suggests that addressing concerns within the design category may necessitate practical interventions to gain teachers’ support. Hands-on strategies and innovative pedagogical methods aligned with a major purpose could prove instrumental in building teachers’ confidence.

Collaboration between teachers and with families The subcategory that pertains to the collaboration among teachers and between teachers and families was covered by 18% of the data, indicating a strong concern in engaging with key stakeholders who play a proximal role in shaping students’ ethical understanding and behaviour, as highlighted by this participant: “But it is here where we, especially teachers, need to take a stand, collaborating closely with families to enhance the quality of students’ learning”. This also recognises that these actors hold the power to either support or undermine an educational project.

Pointless without AI technologies The subcategory that questions the relevance of such a training without AI technologies available at a school level was covered by 18% of the references, shedding light on the belief that ethics training in AIED holds significance when it is integrated with and informed by AI technology. Participants acknowledged that educational endeavours should be purposeful and aligned with the actual needs and practices of stakeholders, grounding ethics training in the available current technological landscape, as demonstrated by this participant: “It wouldn’t make sense to do it too far ahead of the time when teachers will have the technologies to work with”. The cluster similarity analysis reveals an interconnection between this subcategory and the one that raises a thought-provoking question on setting priorities (cf. Figure 3).

Setting priorities In 9% of the discourses, the importance of adding more training specifically focused on AI was questioned. Other shortcomings in teacher education related to inclusion and gender perspective were mentioned as priorities. The participant expressing this concern highlighted the following: “Are we prepared to take on more training, specifically in the field of AI, when we still have so much pending training related to inclusion, gender perspective, and many other areas in which we are (...) failing in teacher education?”. Once again, the dendrogram provides insights on this matter, particularly highlighting the farthest neighbour clustering between the categories of training’s relevance without AI technology and the necessity for debates regarding new role requirements or educational paradigms (cf. Figure 3). According to this pragmatic approach, without setting priorities for training focuses and ensuring teachers have access to AIED, this discussion is spurious.

Pre-service training 9% of the references had also been made for equipping educators with the necessary knowledge and skills related to AIED ethics during their pre-service training, as demonstrated by this quote from one of the participants: “However, in many instances, the discussion of ethics needs to start from a more fundamental standpoint. This particularly applies to the generations now entering universities”. This integration has the potential to cultivate a culture of ethical consciousness and responsibility among future educators. Additionally, when teachers with different levels of experience and diverse knowledge exchange their perspectives in educational settings, positive outcomes can result.

Informal and self-directed continuing professional development 9% of the discourses have emphasised the importance of maintaining an informal and self-directed Continuing Professional Development (e.g., “And the positive awareness that comes with this type of tool, without the need for other types of courses.”), revealing the recognition that professional growth in ethical practices should extend beyond formal training programmes. This can be fostered through reflective practice, learning communities, diverse resources, and collaborative learning. Such subcategory and the one that pertains collaboration among educational actors appear connected in the cluster similarity analysis (cf. Figure 3). This suggests that teachers and families who actively engage in collaboration are more likely to be the ones that take ownership of their professional development and seek out opportunities for learning and growth independently. This proactive approach to professional learning and to forming alliances reflects a collective agency and commitment to improving educational outcomes.

Finally, the cluster similarity analysis reveals that the nodes “compliance with a policy framework” and “pre-service training” share some characteristics, despite being represented by different colours, indicating a less close association between them (cf. Figure 3). Nevertheless, this near neighbour clustering highlights the importance that some participants place on establishing a strong foundation for the integration of AIED, suggesting it be formalised through policy and incorporated into the curriculum for pre-service educators. This concern calls the attention to the

efforts that are expected to accompany teacher professional learning and that relate to policy development and curriculum design.

2.5.4 RQ4: Implementation strategies

The implementation category received 20 references, indicating that the participants placed importance on the delivery methodologies of the training programme. These references likely stem from the participants' experiences as trainers, suggesting the value of hands-on, skill-based training and innovative pedagogical practices in the context of AIED ethics.

Innovative and purposeful pedagogical practices The subcategory centred around innovative and purposeful pedagogical practices has received 45% of the references, indicating a strong desire for intentional and creative practices. In this regard, one of the participants stated:

"There, regarding the assessment aspect, there needs to be a modification. I still don't understand why teachers are reluctant to (...) all kinds of exploration, advancement, and the construction of students' knowledge with open-book assessments. For instance, presenting a case like the one you've just raised and, around it, initiating a discussion or debate. If knowledge and experiences are meant to be shared among everyone, to discuss common points, identify areas of differing thoughts, and establish consensus—which is the most challenging part. (...) I mean, the focus should be on the knowledge-building processes, in utilising what is being generated through artificial intelligence."

This passage illustrates the assumptions participants hold regarding the impact of traditional practices, particularly assessment, in education. To address this concern, they advocate for creativity coupled with purpose. Aligning intentionality with innovative practices reflects a deliberate desire to challenge existing paradigms of meaning-making in education and stimulate fresh approaches to thinking and action. From this viewpoint, participants propose that trainers not only consider how to assess students but also question whether the objectives and nature of what is being assessed remain static. These practices may initially disrupt conventional norms, but they hold the potential to spark meaningful shifts in the educational context, shifting from learning better to an overall experience of better learning.

Hands-on and skill-based training The subcategory concerning hands-on and skill-based training received 40% of the references, indicating a strong emphasis on experiential training and the desire to provide opportunities for applying knowledge in real-world scenarios. This perspective is exemplified by a participant who stated: "I believe that courses for professionals who are currently working should be very practical. They should not be awareness courses". As highlighted earlier, insights

gleaned from the cluster similarity analysis, illustrated in Fig. 3, suggest a connection between these two subcategories and indicate that both approaches can address potential resistance and ease transitions by instilling confidence in participants regarding their existing level of awareness in their field.

Learning by example The topic of learning through examples arose in approximately 15% of discussions, highlighting the significance of practical illustrations and real-life scenarios, through which participants can gain insights into ethical dilemmas and develop skills to navigate them proficiently, as demonstrated by this quote: “The only mechanisms that come to mind for training in that area would be to show examples of what not to do so that they can learn how to use them. (...) Teaching small cases tailored to the user and having them identify if they share that use and if they see potential risks in using it. And then, they are presented with half of the case and the consequences are observed”. This subcategory, as indicated in Fig. 3, is linked with the promotion of broadened ethical thinking and the update of knowledge and skills. It implies a recognition of the importance of examples, role-playing, and role-taking in stimulating perspective taking and the affective-cognitive dissonance that leads to the emergence of ethical dilemmas and to a deeper understanding of ethical decision-making processes. Furthermore, these examples serve as valuable learning tools by offering direction on which areas to prioritise for additional learning. This is especially important when considering the informed allocation of resources needed for continuing professional development.

3 Discussion

The findings of this research contribute to shaping the next steps in developing an AIED and education ethics training course for K-12 in-service and pre-service teacher programmes. This evolving framework aims to foster dialogic practices that allow teachers to reflect on the ethics of education in the presence of AI technologies. The discussion examines the findings concerning the research questions to guide the next steps in developing the training course. It integrates theoretical and conceptual elements from the current draft of the AI Competency Framework for Teachers’ Development (UNESCO, 2023c), emphasising how the findings align with ongoing work by experts in this field. Additionally, it interprets the results through the lens of the ik-model domains (Mouta et al., 2015). The ik-model builds upon the TPACK framework by introducing diachronic dimensions to enhance the understanding of technological integration in learning environments. These dimensions focus on relationships and processes, adding depth to the technological and content domains. They provide a perspective on how relationships between different educational stakeholders are shaped and change, and how pedagogical processes are initiated, evolve, and adapt contextually to foster meaningful learning experiences. Additionally, participants’ responses naturally encompassed the dimensions of the ik-model in their narratives, thereby demonstrating the model’s relevance and applicability for discussing the results. The forthcoming discussion on each research

question will utilise the findings to explore its practical implications for teacher training on these matters.

Regarding the first question, concerning the structuring dimensions of a training programme on AIED, ethics, and agency it is possible to acknowledge that participants expressed a nuanced perspective on what is expected to be implemented, far from a positivist approach to designing training experiences. In fact, teachers spontaneously covered all the ik-model's domains as relevant parts to be considered in such a training: (a) technological domain – AIED techniques and technologies; (b) content domain – ethical thinking, educational paradigms, teachers and students' roles, policies frameworks, best practices; (c) relational domain – roles and collaboration between key educational stakeholders; and (d) processes domain – dealing with resistance and transition, experiential practices, learning by example, self-directed learning. What is worth noting on establishing this parallel on the discussion, is that the relational and processes dimensions are often sidelined in training programmes on AIED for education, which predominantly focus on content knowledge and pedagogical content knowledge. The categories addressed within the processes domain also demonstrate participants' proficiency in andragogical principles, possibly stemming from their roles as teacher trainers, indicating a deep understanding of the dynamics involved in designing and facilitating adult learning.

On the second question, regarding objectives, participants covered the course technological and content domains. The first implication for preparing this course regarding its aims is the following:

- (1) Consider it an ethical decision in itself to discuss AIED frameworks as either supporting tools or constraints to meaningful practices and pertinent innovation.

This concern and recommendation regarding teacher competency align with the UNESCO draft, which proposes the inclusion of “ethics of AI” as an asset. This asset encompasses an understanding of fundamental ethical principles related to AI, as well as participation in communities for the iterative development of institutional and societal regulatory environments, bridging the gap from understanding to active engagement. In these focus groups, participants put emphasis on the importance of discussing AIED-oriented policy frameworks during training. The adoption of such frameworks is deemed suitable solely when enabling a comprehensive approach tailored to local interests and school microsystems. As these frameworks are typically developed through extensive research, consultations, and expert input, with the aim of reflecting societal values, legal requirements, and best practices in the field, they can provide reassurance to teachers during transition stages. For instance, the *Beijing Consensus on Artificial Intelligence and Education* (UNESCO, 2019), *The Ethical Framework for AI in Education* (The Institute for Ethical AI in Education, 2021), and the draft *AI Competency Frameworks for school students and teachers* (UNESCO, 2023c) can serve as valuable guidelines in this regard. These frameworks are scalable and allow for some contextualisation and progression, serving as an initial step in addressing the ethical challenges posed by AIED and ensuring its alignment with pedagogical principles.

On the one hand, the decision to follow these guidelines in the context of AIED can be seen as an ethical decision in itself. In fact, it can help reduce excessive reliance on singular experiences and personal biases. On the other hand, it may prompt considerations as to whether these frameworks might unintentionally impede meaningful innovation, creativity, or more dynamic and fully context-driven approaches. Nevertheless, exploring these frameworks may provide a basis for discussing a set of principles, considering the perspectives and interests of various stakeholders.

Regarding the third question about particular considerations for designing a training programme in this field, participants mainly reflected on the course processes domain. The second implication for preparing this training is as follows:

- (2) Provide opportunities for addressing teachers' concerns, needs, and uncertainty during transition processes throughout course's design and implementation phases.

The utopian and dystopian narratives surrounding AI, the current knowledge, and conditions for its integration, combined with past experiences of integrating digital technologies into learning environments, make it key to address the cognitive, affective, and behavioural aspects related to AI integration during training. In fact, participants underscored the necessity of assisting teachers in handling resistance and facilitating transitions to foster positive training effects. Optimising attention to change, uncertainty and transition processes can also be strategically used to measure the course's impact on educators' cognitive, affective, and behavioural dimensions of attitude.

Cukurova et al. (2023) devised a reliable instrument for assessing the comprehensive factors that impact teachers' adoption of adaptive learning platforms in educational settings. According to their study, it is imperative that these technologies do not impose an additional workload on teachers, requiring them to switch between different tools, or navigate various pedagogical practices during implementation, while ensuring the minimisation of ethical concerns. Certain frameworks concentrate on the emotional and cognitive phases that teachers undergo when faced with change, including endings, neutral zones, and new beginnings (Bridges, 1991; Hall & Hord, 1987). Frameworks like The Diffusion of Innovations Theory (Rogers, 2003) and Change Management Models highlight the importance of communication, social networks, stakeholder involvement, leadership support, and the perceived advantages of the innovation in influencing teachers' acceptance and positive implementation. The selection and combination of frameworks may vary depending on the context, nature of the change, and the specific requirements of teachers and educational institutions, recognising and exploring transition pathways throughout the implementation phase.

Nevertheless, in this regard, understanding uncertainty as the movement through which agency can be restored is crucial for enabling educational actors to lead implementation. The process of integrating new technologies is inherently non-linear and open-ended. Within this space, imagination, experimentation, and the recombination of old and new elements can thrive, fostering meaningful decision-making and

agency. Becoming is not only about moving towards new possibilities but also about letting go of what will no longer be or other alternative modes of being. Understanding and incorporating this dynamic movement into the strategic training axis is indeed particularly rich for engaging educational actors in practices of meaning and collective agency.

In relation to the fourth research question, focused on implementation strategies, participants covered the course technological, relational and processes domains. With respect to the technological domain, they proposed the following:

- (3) Use AI not only to support course design, delivery, and assessment but also to redefine and concretise AI tools within dialogical practices.

Participants highlighted the need for direct access to AI tools in order to comprehend their techniques and subsequently assess their potentials and pitfalls. They suggest complementing theoretical discussions of ethical concerns with practical training using various AI resources to support course delivery and assessment. Methods for implementation include offering guidance in course forums through AI teaching assistants, fostering collaboration with AI facilitators, and evaluating course activities using learning analytics. Additionally, GenAI can be trained to provide models for teacher assistance, incorporating school and national educational guidelines. This includes addressing several topics which can be chosen by the trainees, such as managing challenging classroom behaviours or offering pedagogical tips to improve student engagement with a given subject. The suggestion aligns with the competency outlined in the UNESCO draft, which advocates for teachers to progress from understanding fundamental AI concepts and functions to comprehensively mastering the adaptation of AI tools to develop solutions tailored to their own educational contexts.

This strategy should not be perceived solely through an instrumental lens. It goes beyond mere tool usage to understanding its functionality, enhancing learning experiences, and addressing the ethical challenges these resources may pose. A perspective aligned with Bardone et al.'s (2024) concept of tinkering would be more dynamic, exploratory, and diachronic for all educational actors. Through this approach, they may participate in the process of concretisation of these tools, as defined by Simondon (2008). A new iteration of the technological object would facilitate the concretisation of its overabundant functions, potentially leading to new technological possibilities within a technogeographical space situated between nature and technics. This process involves the integration of the technological object with its surroundings, demonstrating a level of internal compatibility that produces its external adaptability. Importantly, such an approach holds dual significance: it pertains to the concretisation of pedagogical practices themselves, reshaping evolving roles among all actors; and it also transforms power dynamics in shaping narratives surrounding AI for education.

The relational domain was also covered with respect to the fourth research question on implementation strategies, with participants advocating for:

- (4) Fostering individual and collective agency on AIED ethical issues through informal and self-directed learning channels.

In addition to engaging in critical reflection, prioritising training initiatives that enhance collective agency is crucial. In fact, teachers are expected to translate insights into meaningful actions aligned with the specific demands of their educational environments. So, another insight derived from this study is that any training fostering a reflective and collective agency journey in AIED and education ethics must persist through various channels, including professional learning communities, communities of practice, online professional learning networks, teacher associations and unions, faculty meetings and committees, school leadership teams, and action research projects. The UNESCO draft on AI for Professional Development also encompasses this perspective, viewing AI as a facilitator of learning for teachers. It emphasises the importance of teachers critically adjusting or altering AI tools to better meet the evolving needs of professional communities and contexts. This makes it clear that in the context of AI, professional learning is key, namely through informal communities and self-directed endeavours.

Collective agency can also be fostered in educational cultures that allow for contextualisation and autonomy, through curriculum flexibility, adapting to the diverse needs of student populations and broader community requirements in feasible timeframes (OECD, 2018). This work can support the building of alliances that advocate for the pertinence and sustainability of bottom-up decisions. According to Crary (2022), these bottom-up strategies are critical to collectively reshape the landscape of technology in education, contributing to a social movement that brings about collective benefits and shared risks. This assertion gains heightened significance in light of the prospect of AIED automating pedagogical practices that lack contextual meaningfulness, coupled with the risk of potentially disempowering teachers in the long run. In a more positive approach, by incorporating AI techniques thoughtfully and ethically, educators can amplify collective agency. This can be achieved by leveraging data in a contextualised and insightful manner to inform decision-making and collaborative planning processes. Moreover, collective agency can be fostered when educators are not only motivated to ethically integrate AI into educational practices (educating with AI) but also to develop students' understanding and critical thinking about AI itself (education for AI). Initiatives such as the European Commission's Code Week (European Schoolnet, 2021) and the MIT RAISE programme (RAISE, 2024) exemplify this approach by supporting or providing K-12 curricula to expose students to AI at an early age. As Gertz (2016) argues, technology not only mediates our nihilism but also mediates our responsibility, serving as a means to prompt reflection on the human values that shape and nurture AI today.

A fifth implication for the implementation dimension regarding the processes domain has to do with the need of:

- (5) Employing narratives to contextualise these AIED technologies, considering ecosystemic factors.

The subcategories that emerged from participants' discussions, focusing on dealing with resistance and facilitating change, are associated with the implementation of training through innovative and purposeful andragogic strategies. In line with this objective, the ethical scenarios derived from the initial stages of this educational

design research were considered relevant by participants for knowledge acquisition processes, aiming for consensus. Therefore, another recommendation is to employ narratives to contextualise these AIED technologies, considering ecosystemic factors, implementation strategies, and societal dynamics. Consequently, in the upcoming phase of this study, these dilemmas will undergo review to incorporate specific instances reflecting advancements in AIED technologies concerning the same ethical dilemmas. These scenarios, combining both dystopian and utopian elements, will function as heuristic tools for reimagining education from an ethical and agentic perspective. This can aid in developing the UNESCO-recommended competency focused on a human-centred mindset, ranging from recognising opportunities and risks to demonstrating a profound understanding of societal impact and the ability to engage in transformative actions to address related challenges.

Another implication in terms of implementation strategies related to the processes domain concerns the following aspect:

- (6) Incorporate meaningful experiential learning as a means to stimulate dialogic ethics in AIED-related discussions.

The broad acceptance of GenAI has sparked concerns among educators across secondary to tertiary education levels. Despite ChatGPT not being crafted for educational settings, it has raised worries, particularly in scenarios where it can complete homework for students. Nevertheless, these platforms offer an opportunity to reassess and redefine the learning processes and corresponding pedagogical methods, potentially leading to more meaningful outcomes. This can entail students interacting with these platforms and then assuming the role of teachers to correct the AI responses. They may also position themselves towards the information, taking on roles such as researcher or opponent. Additionally, involving other community actors in the learning process to diversify information sources and fostering group and project-based learning are viable approaches. Indeed, teachers are the ones who understand the microgenesis of learning, the progression that unfolds throughout a learning session. So, using large language models can, in fact, be an opportunity to challenge tasks of assimilation, accommodation, and evocation of knowledge, while creating agentic opportunities to deal with this technological novelty within school. This can be purposefully achieved by engaging students in other high-order learning processes, such as perspective-taking (from both humans and machines), introspection (focusing on the uniquely human process of self-reflectiveness), imagination (inspired by the vast possibilities enabled by AI), affective-cognitive dissonance (introducing dilemmas brought by the interaction with AI that stimulate ethical reasoning), and exploration of engagement with action through balanced power and agency (teachers share decisions with their students about how to learn and be assessed with GenAI).

So, another suggestion arising from this study is that training is expected to integrate meaningful experiential learning as a means to fully embrace dialogic ethics in learning settings. This can be maximised through activities such as role-playing and role-taking, enabling educators to assess and refine their strategies directly in their classrooms during the training period. Subsequently, they can return to the secure training space to evaluate their practices alongside peers and trainers. This also

enables creating a platform for debate on ethics among educators, between educators and trainers, among students, and between teachers and students, ensuring that inter-generational perspectives on AI challenges are not only debated but also infused into subjects, potentially leading to valuable outcomes. This strategy directly aligns with the UNESCO suggested competency of guiding teachers to progress from identifying the potential benefits of using AI systems to critically evaluating AI in pedagogical practice, ultimately moving towards transformative pedagogy.

4 Conclusion

Until recently, the ethical foundations of education in the context of AIED have not been thoroughly explored, and the corporate narrative continues to dominate public discourse on AI for education. Debates on this topic seem to be especially pertinent as we live in a world of technique, as Ellul (1964) defined it: a world embedded in the pursuit of methods derived through rational means and achieving optimal effectiveness across all areas of human endeavour, where the humanities have lost their pace. AI technologies have become more pervasive, and major multinational corporations, in cooperation with at least one local government, are attempting to influence educational practices. Non-governmental curricula are already incorporating AI learning outcomes at middle and high school levels (UNESCO, 2022). Furthermore, various research findings indicate that AI curricula only marginally address AI ethics, and there is a shortage of opportunities for teacher preparation to tackle contemporary challenges in this area. To bridge this gap and capture diverse stakeholders' perspectives, this study focused on insights on education ethics in the AI era stemming from teachers and teacher trainers focus group discussions. Envisioning the subsequent step as the development of a training course in this area, this research phase contributed significant insights on course objectives, implementation strategies, and design stages. By prioritising transversal assets beyond AI techniques and technologies, participants highlighted the significance of continuing professional development as a platform where teachers actively contribute to shaping their professional growth and standing, safeguarding their agency against potential encroachments by tools or policies. They were also recognised as key figures for challenging established knowledge in their field, while also questioning the very nature of epistemology and the conditions under which it is nurtured. Therefore, participants articulated their perspectives regarding the revitalisation of educational paradigms and teachers' roles, not only as gatekeepers of high-quality content, but as vital agents of thought and action in the face of the polysemic landscape of AIED technologies.

While education remains subject to state regulation, opportunities exist for elevating the status of teachers across diverse educational systems and impacting public policy development, in the interplay between technology, pedagogy, and society. Additionally, increased investment in professional learning and participatory policy reforms, leading to policies that are more comprehensive, has the potential to significantly enrich the outcomes of national investments in AI for the educational sector. In this context, the present research has sought to provide educators with an avenue to explore relevant educational possibilities and to assume proactive roles as agents of meaningful transitions in

their learning environments. Participants' insights shed light on the importance of a broad ethical exercise and strategic collaboration among educational actors. They also revealed a nuanced understanding of the cognitive and emotional challenges teachers face during the integration of AI technologies, where uncertainty plays a key role. This calls for a comprehensive consideration of meaning indeterminacy and underscores the importance of carefully considering signification processes as they unfold in pedagogical practices. While the recommendations arising from the study underline the critical role of teacher autonomy in professional development, especially in the era of AIED, realising this objective can be a complex undertaking due to constraints related to limited resources, hierarchical organisational structures, and cultural factors that may hinder teachers' capacity to exercise professional autonomy. This clearly exemplifies the complex interplay between structure and agency to which the current research and further training aim to offer actionable insights. The participatory methodology used in designing research and teacher training situates this study within the broader discourse of dialogic ethics. This approach is sought to be responsive to the challenges posed not only by AI technologies, but by a world driven by technique. It is only when uncovering the technique layer that is overarching education too, that a consistent dialogue on ethics for education can be established, going beyond mere debates on the ethics of AI.

Finally, while this research advances past studies that merely explore the ethical challenges of AIED by offering concrete recommendations for training programmes targeting teachers in this domain, it is necessary to acknowledge its challenges and limitations. One challenge arises from the disconnect between the primary examples of the scenarios toolkit and the rapid advancements in AI technologies. This underscores the imperative of integrating more relevant examples that align with the current state of the art, thereby enhancing the applicability of the toolkit for the upcoming training. The scenarios will be used with these updated examples, as they depict ethical dilemmas that remain relevant today. Other constraints in the study emerge from concerns about generalisability, given that the insights derive from a particular participant group, even though efforts were made to encompass participants from various geographical backgrounds. Moreover, the study primarily focused on providing recommendations for teacher training in AIED, agency, and dialogic education ethics, without exploring the actual implementation and effects of such a training programme. This research calls for the next step, which pertains an examination of its impact on agentic decision-making regarding the use of AIED in educational settings. Additionally, cultural and societal factors can significantly influence the level of autonomy granted to teachers, complicating efforts to introduce alternatives within existing structures, reform those structures, or challenge and resist them, as advocated by Wright (2010). In certain educational contexts, traditional norms and hierarchical power dynamics impose strict adherence to prescribed curricula and pedagogical methods, leaving limited room for individual autonomy or innovation. This lack of professional autonomy restricts teachers' ability to shape their own professional growth and may hamper their capacity to read their learning ecosystems in light of evolving educational demands, practices, and pressures. As a result, these contexts are less likely to play a critical role in challenging the predominant dominance of technique, reaffirming educational ethics and leveraging the potential of AIED technologies when integrated meaningfully. Consequently, they are more prone to encountering the ethical challenges associated with their use.

Appendix 1

Table 4 Focus groups script

FOCUS GROUP OPEN-ENDED QUESTIONS	SUPPORTING QUESTIONS
1. What caught your attention the most in this future scenario?	<ul style="list-style-type: none"> a. What has worried you the most? b. What advantages can the application of AI bring to Education, according to this scenario? c. What would you like to change from this perspective? d. How does this scenario present dimensions that compromise... <ul style="list-style-type: none"> d1...educational objectives? d2...assessment methods? d3...equity? d4...autonomy? d5...privacy? d6...transparency and accountability? d7...participation? d8...ethical design?
2. What ethical issues can arise in this scenario? For whom?	<ul style="list-style-type: none"> a. From the perspective of student development b. From the perspective of teachers c. From the perspective of the school as an institution d. From the perspective of families e. From the perspective of achieving the Sustainable Development Goals (ensuring quality education, reducing inequalities, ensuring peace and effectiveness of institutions)
3. A concern in the context of human-machine interaction has to do with our sense of agency when interacting with automated technologies, which assume much of the control that was previously in the hands of the user. How does the use of AI interfere with students' ability to think and act autonomously?	<ul style="list-style-type: none"> a. In what ways are teachers addressing these issues? b. What initiatives should be taken to ensure the best interests of students when using AIED?
4. Considering what we have discussed, what do you think is key to include in a training aimed at teachers and school coordinators on these topics?	<ul style="list-style-type: none"> a. Which topics should be focused? b. Which delivery methods should be privileged? c. How would you distribute the theoretical and practical components of this course? d. What topics should the theoretical part focus on and which dimensions could be worked on more practically? e. Regarding teaching resources, which options would you select and why? f. What could be the duration of such a training? g. Considering the availability of teachers, what would be the best time to have teachers involved in this course? h. What type of evaluation would be most appropriate for this training?

Appendix 2

Table 5 Table of node categorisation in NVivo

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
B. ATTITUDES TOWARDS AIED		
A1. Affective attitudes	Perceptions, beliefs, feelings, or behaviours individuals have towards AIED	<i>ND for the general category</i>
A1.1. Negative	Perceptions or feelings individuals have towards AIED that can range from the positive to the negative spectrum	"We have a segment of teachers who still fear the use of this type of resources and are becoming enemies of cell phones."
A1.2. Positive	Unfavourable or pessimistic perceptions or feelings towards AIED, such as fear, scepticism, or concern	"Artificial intelligence facilitates living conditions."
A2. Cognitive Attitudes	Favourable or optimistic perceptions or feelings towards AIED, such as enthusiasm, trust, and excitement	<i>ND for the general category</i>
A2.1. AI tech as neutral tools	Beliefs, opinions, and thoughts individuals hold towards AIED	"So, technology is not inherently bad, but it is true that its use can lead to unethical outcomes, not just unintended uses..."
A2.2. People's Agency	AIED tools are deemed to be free of any inherent bias, value or intent	"As a society, we have the ability to regulate the technology that is created and determine what should and should not be developed."
A2.3. Technological determinism	Capacity, control, influence, and ability to shape the use and development of AIED according to needs, preferences, and values	"It can be a case of falling into technological determinism, where what the robot or machine says is considered the ultimate truth."
C. BENEFITS AND OPPORTUNITIES		
	Belief that AIED will inevitably dictate or significantly influence educational practices, policies, and societal norms, often disregarding individual agency	<i>ND for the general category</i>
	Positive outcomes and possibilities that arise from the use of AIED	

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
B1. Citizenship and Equity	Principles of fairness, inclusivity, and equal access to educational resources and opportunities, emphasizing AIED potential to bridge educational gaps, provide tailored experiences for diverse learners, and promote equitable outcomes, irrespective of students' backgrounds or circumstances	"So, equity does not arise from equalisation. Equity arises from differentiation, from respecting differences. That's how we manage technology, in order to achieve that goal."
B2. Education as a system	Comprehensive infrastructure of educational institutions, policies, and practices positively affected by AIED, including curriculum development, pedagogical methods, assessment strategies, administrative tasks, and decision-making processes	"Leveraging the insights provided by artificial intelligence through research can support and strengthen educational processes within education systems."
B3. Examples	Instances or tangible demonstrations of how AIED has been successfully applied in educational settings, in areas such as adaptive learning platforms, intelligent tutoring systems, automated grading, or educational data analytics	"For example, there are artificial intelligence systems being developed to enhance the communication skills of individuals with autism, among other applications. These types of interactive tools, leveraging natural language processing, have shown significant success in the field of artificial intelligence."
B4. Families	AIED technologies provide parents and caregivers with tools and resources to support their children's learning, access educational materials, track progress, and collaborate with educators and schools to ensure the educational success and well-being of their children	"How can I improve living conditions within my family and for the community [with this AI tech]?"
B5. Future of Society and Civilisation	Long-term positive impact of AIED on the broader social, cultural, economic, and technological aspects of civilisation, contributing to the overall progress and development of future societies	"We should ethically utilize all the advancements in science and technology to grow as a society and as individuals."

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
B6. Lifelong Learning	AIED technologies support to individuals in acquiring new knowledge, developing skills, and accessing educational opportunities throughout their lives, enabling them to adapt to changing demands, pursue personal growth, and participate in a knowledge-driven society	"Each and every child, young person, and throughout their entire life, will have a personalised path that combines opportunities for work, business, or entrepreneurship with their personal interests, aptitudes, and learning style."
B7. People in general	Potential of AIED technologies to benefit various stakeholders by improving educational outcomes, enhancing teaching and learning experiences, and advancing educational practices	"For some individuals, this can be satisfying, as making those decisions can be challenging, and some may feel at ease not having to make those decisions."
B8. School Administration	AIED optimising administrative tasks, routine processes, data-driven decision-making, resource allocation, and school operational efficiency	ND
B9. Students	Primary beneficiaries of education provided with AIED tools, with benefits on their engagement, motivation, and achievement	"A better learning experience, increased productivity, and ultimately (...) greater happiness, a fuller sense of accomplishment, and access to better education and opportunities for growth."
B10. Teachers	Critical stakeholders in education can be provided with tools to improve their pedagogical practices, with benefits for themselves and their students	"(...) to contribute to developing creativity and innovation in the teacher."
B11. Work and Employability	Impact of AIED technologies on the workforce and the skills needed for future employment, enhancing job readiness, lifelong learning for career development, and responding to evolving demands of the labour market	"Utilising all this information to develop entrepreneurial skills and thereby changing the mindset of constantly thinking 'How am I going to find a job?' to 'How am I going to create a job?'"

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
D. ETHICAL CHALLENGES	Moral dilemmas and considerations that arise from the use of AIED, such as data privacy, algorithmic bias, transparency, accountability, fairness, and other implications of AI decision-making in educational settings	"I consider the topic (...) of paramount importance because we should indeed embrace information and communication technologies, including when artificial intelligence generates data based on a parameter, under the condition that individuals who manage these technologies do so appropriately."
C1. AIED tech design and developers	Criteria for developing AI algorithms, user interfaces, educational software, and systems that align with educational goals, ethical principles, and the needs of learners and educators	"Unfortunately, algorithms are created by human beings and are influenced by cultural patterns. So, what will happen? How will it be created?"
C2. Citizenship and Equity	Challenges in terms of fairness, inclusivity, and equitable access to educational resources/opportunities, emphasising digital divides	"They can replicate barriers, socio-economic barriers that currently exist."
C3. Communities and cultural context	Significance of local cultural, social, and community-specific factors when implementing AIED, respecting cultural practices and values, so that AIED aligns with the specific needs and contexts of different communities	"(...) certain technology in artificial intelligence is being developed in Europe, but how do I contextualise it? That is one of the main challenges. I believe that simply copying and trying to apply it is a mistake, and that's where we need to evaluate and understand our own characteristics, our population, what we have and what we don't have, and how we can apply these technologies to strengthen our situation."
C4. Developing Frameworks and Policies	Guidelines, regulations, and standards to govern the use of AIED in educational settings, ensuring that these technologies align with educational goals and support positive learning outcomes	"To foster development, it is crucial to play a significant role in terms of policies at the national level."

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
C5. Different interests between users and corporations	Potential misalignment of goals and motivations in the adoption and use of AIED, emphasising the varying priorities and perspectives of educational users (such as students, teachers, and families) and the commercial interests of corporations developing and providing these technologies	“When there is free software, which is what we try to use, ultimately that company has to make money somehow, and they do it through the sale of data. That’s why I believe the root of the problem lies there, and also because it’s a private company that manages this data.”
C6. Education as a system	Comprehensive infrastructure of educational institutions, policies, and practices challenged by AIED, including curriculum development, pedagogical methods, assessment strategies, administrative tasks, and decision-making processes	“(…) everything related to learning to interact with technology, with artificial intelligence, should be done through an innovative educational system and not from an educational system based on the principles of the industrial system. Because the two things are not compatible, they are inconsistent.”
C7. Examples	Instances or tangible demonstrations of how AIED has been ineffectively applied in educational settings, in areas such as adaptive learning platforms, intelligent tutoring systems, automated grading, or educational data analytics	“In Spain, for ex., based on the PISA tests, which is what I am most familiar with, regional governments are quite cautious about disclosing which schools have been evaluated to avoid creating rankings. In contrast, the United States is much more transparent about it.”
C8. Families	AIED technologies raise concerns for parents regarding their ability to meaningfully respond to the challenges posed when their children engage with these technologies, particularly concerning the collection, storage, and use of personal data, as well as biases in decision-making algorithms	“Educational institutions will come to life to the extent that educators and parents make an effort to keep up with the ways of thinking of our students. It is important to understand that they were born in an era where they have had access to information and communication in different modalities or aspects than those of us who are over 50 years old.”
C9. Future of Society and Civilisation	Long-term negative impact of AIED on the broader social, cultural, economic, and technological aspects of human civilisation, hindering overall progress and development of future societies	“(…) one situation is where we assume that the need to learn certain things by hand is already surpassed and that we can use technology for other tasks. I believe that this debate, for now, I’m not saying we’ve lost it, but certainly we’re progressing slowly, very slowly.”

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
C10. Investing in projects, research, and theories for integration	Allocation of resources and funding to initiatives that explore the meaningful integration of AIED into educational practices, supporting research endeavors, innovation, and the development of theoretical frameworks	"It is important not to lose sight of the fact that research is necessary, and of course, in research, not everything will conclude that what is known is true or positive. It is about exploring, investigating, and adapting those results to our diverse students and teachers."
C11. Lifelong Learning	AIED drawbacks for individuals in acquiring new knowledge, developing skills, and accessing educational opportunities throughout their lives, hindering their ability to adapt to changing demands, pursue personal growth, and actively participate in a knowledge-driven society	ND
C12. Problems with Data sets	Challenges related to the quality, reliability, representativeness, and bias present in the data used to train AI models, such as insufficient or inaccurate data, and lack of diversity	"We have serious doubts about whether the data source we are working with can provide us with the information to transfer what is happening in other fields to the educational field. (...) similar to what banks have to make their decisions. We are far from achieving that level of information access."
C13. School Administration	AIED hindering administrative tasks, routine processes, data-driven decision-making, resource allocation, and school operational efficiency	ND
C14. Students	Primary beneficiaries of education encountering issues using AIED tools and resources that challenge their autonomy, agency, engagement, motivation, and performance	"So, where does freedom fit in? Where does autonomy fit in? Where do those abilities that we assume our students have fit in?"

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
C15. Teachers	Critical stakeholders in education finding difficulties using AIED tools that bring challenges either for themselves or for their students	"It seems that if all these circumstances are present, this individual may require such adaptation or intervention. It is the professional who should consider all of these factors and determine whether it aligns with their own observations."
C16. Work and Employability	Impact of AIED technologies on the workforce and the skills needed for employment, supporting job readiness, skill development, career transitions, and preparing individuals for the evolving demands of the labour market	"It can hinder social mobility. I believe that one of the risks of the lack of social mobility is the absence of significant opportunities."
E. TEACHER INITIAL AND CONTINUING PROFESSIONAL DEVELOPMENT	Knowledge, skills, and competencies to deal with AIED challenges both in initial training provided to teachers entering the profession and during ongoing professional development throughout their professional journey	<i>ND for the general category</i>
D1 Broadened ethical thinking	Capacity to cultivate a comprehensive understanding of ethical principles, ensuring that AIED implications are thoroughly examined and addressed	"However, if we don't educate ourselves in a critical manner or analyse the reality more deeply, we may end up using it out of inertia."
D2. Collaboration between teachers and with families	Partnerships and communication between educators, parents, and caregivers, to support students' educational development	"But it is here where we, as educators, need to take a stand, collaborating closely with families to enhance the quality of students' learning."

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
D3. Compliance with a policy framework	Guidelines, regulations, and policies governing the use of AIED in education, ensuring that it is implemented in an appropriate and accountable manner	“(…) teachers should indeed receive training in this area, including current legislation and how to apply it within the framework of using applications that collect personal information, such as augmented reality systems or artificial intelligence. However, the challenge lies in the fact that these systems are sometimes opaque, making it difficult to understand how legislation is applied in practice.”
D4. Dealing with resistance and transition	Ways of addressing challenges, concerns, and resistance that may arise during the implementation and adoption of AIED	“We need to include in the chain to work with those teachers who do not really conceive these scenarios and want to distance themselves from them. I believe that this is another debate where the focus should be on how to attract those teachers.”
D5. Hands-on and skill-based training	Learning experiences that involve active participation, practical application, and the development of specific skills that enable learners to acquire and practice competencies directly related to their field of study or work	“Well, I agree with the rest of the people in the group that it should be a training, especially practical.”
D6. Informal self-directed CPD	Professional development activities that are not formally structured, implying educators taking ownership of their learning by pursuing self-directed initiatives, exploring resources, participating in online communities, attending conferences, or engaging in reflective practices	“And the positive awareness that comes with this type of tool, without the need for other types of courses.”

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
D7. Innovative and purposeful pedagogical practices	Creative and intentional didactic strategies to facilitate meaningful learning experiences, with the goal of enhancing student engagement, understanding, well-being, and achievement	<p>“We no longer have to provide content to the students; we have to guide them in learning how to analyse the vast amount of information available to them. They should be able to use that information to solve the problems they are facing within their communities and contribute to transformation. Of course, we must also consider environmental preservation. We need to strengthen initiatives that not only protect what little we have but also enhance it to improve our culture.”</p>
D8. Learning by example	knowledge, skills, or understanding acquired through the observation and emulation of successful models, examples, demonstrations, or case studies to guide and inspire learners in their own learning processes	<p>“Reflecting on that, I believe, is important. Practical examples, and if possible, having witnesses to what is being done in the United States or China, where I think there are already videos showing what happens in schools. Oh, I think it could be very interesting to consider from that perspective, understanding how things are done.”</p>
D9. New role requirements	Evolving expectations and responsibilities placed on educators, adapting to changes in pedagogical practices in the advent of AI/ED	<p>“I really liked it when Javier said that becoming guides means positioning ourselves horizontally in the processes of development, in learning, in the construction process, in existence.”</p>
D10. Pointless without AI tech	Educational practices or initiatives related to the use of AI/ED considered less relevant and impactful in the absence of these technologies	<p>“But then, there’s no discussion about the future, so it’s like I believe it’s a task we shouldn’t abandon because it’s important, but ultimately, it becomes very difficult to be monitored in real-time. I’m speaking in terms of infrastructure to technology in the classroom.”</p>

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
D11. Pre-Service Training	Training focusing on ethical considerations and requirements in the use of AIED during the preparation and development of future educators	<p>“However, in many instances, the discussion of ethics needs to start from a more fundamental standpoint. This particularly applies to the generations now entering universities. These young individuals have grown up with a mobile phone practically in their hands since a very young age and are accustomed to sharing absolutely everything, from their exam grades on Twitter to the view from their bedroom window, and so on. Therefore, I believe that this cyber ethics, which should also be integrated into education for future teachers, should first be instilled through a values-based approach before its implementation in the educational system.”</p>
D12. Questioning Educational Paradigms	Inquiries to conventional approaches to learning, entailing the re-evaluation of established educational practices and assumptions in the advent of AIED	<p>“Yes, and that also addresses the issue of training. I often wonder if we have adequate training for teachers or students in terms of digital competencies or skills. If we approach it from the traditional way of learning and educating, it creates a disconnect because the digital realm operates differently. In fact, it is a system that is forcing us to change and transform ourselves.”</p>

Table 5 (continued)

NODE NAME	DESCRIPTION	EXTRACTED EXAMPLES FROM CATEGORISATION IN NVIVO
D13. Setting priorities	Goals and needs for teacher training, involving educational objectives, resource allocation, stakeholder requests, and long-term sustainability	<p>"However, in any case, what concerns me the most about all this is that one of the pending subjects in our educational system is teacher training. So, the question I raise and that concludes this thematic section on teaching is the following: Are we prepared to take on more training, specifically in the field of artificial intelligence, when we still have so much pending training related to inclusion, gender perspective, and many other areas in which we are still, colloquially speaking, failing in teacher education?"</p>
D14. Update of knowledge and skills	Continuous process of acquiring new information, competencies, and proficiencies to stay current with advancements in the field of AIED	<p>"I mean, as an educational community, particularly in the teaching sector, we should strive to have a good level of proficiency and develop competencies to understand and navigate the advancements in science and their relationship with the ethical aspects and potential benefits or problems arising from science and technology."</p>
F. OTHER	Incorporation of supplementary remarks that were not critical to the AIED ethics discussion, potentially vague or tangential in nature	<p>"I would like to issue a cautionary note, urging us not to assume that artificial intelligence or technology will restrict our freedoms, because we are assuming that we have freedom today, right?"</p>

Author contributions All authors contributed to the study conception and design. Material preparation, data collection and analysis were performed by Ana Mouta. The first draft of the manuscript was written by Ana Mouta and the other two authors revised and commented on previous versions of the manuscript. All authors read and approved the final manuscript.

Funding Open Access funding provided thanks to the CRUE-CSIC agreement with Springer Nature. The research leading to these results received funding from Fundação para a Ciência e Tecnologia under Grant Agreement No 2020.06258.BD. <https://doi.org/https://doi.org/10.54499/2020.06258.BD>

Data availability The datasets generated and analysed during the current study are not publicly available to protect study participant privacy but are available from the corresponding author on reasonable request.

Declarations

Competing interests All authors certify that they have no affiliations with or involvement in any organisation or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Abakah, E., Widin, J., & Ameyaw, E. K. (2022). Continuing professional development (CPD) Practices among basic school teachers in the central region of Ghana. *SAGE Open*, 12(2), 21582440221094596. <https://doi.org/10.1177/21582440221094597>
- AI4T. (2024). *Artificial intelligence for and by teachers*. Retrieved March 1, 2024, from <https://www.ai4t.eu/>
- Alrakhawi, H. A., Jamiat, N., Umar, I. N., & Abu-Naser, S. S. (2023). Improvement of students achievement by using intelligent tutoring systems - A bibliometric analysis and reviews. *Journal of Theoretical and Applied Information Technology*, 101(10), 3793–3800. Retrieved November 29, 2023, from <https://www.jatit.org/volumes/Vol101No10/11Vol101No10.pdf>
- Bandura, A. (2023). *Social Cognitive Theory: An agentic perspective on human nature*. Wiley & Sons.
- Bardone, E., Möttus, P., & Eradze, M. (2024). Tinkering as a complement to design in the context of technology integration in teaching and learning. *Postdigital Science and Education*, 6(1), 114–134. <https://doi.org/10.1007/s42438-023-00416-6>
- Berberian, B., Sarrazin, J. C., Le Blaye, P., & Haggard, P. (2012). Automation technology and sense of control: A window on human agency. *PLoS ONE*, 7, e34075. <https://doi.org/10.1371/journal.pone.0034075>
- Biesta, G. (2020). Risking ourselves in education: Qualification, socialization, and subjectification revisited. *Educational Theory*, 70(1), 89–104. <https://doi.org/10.1111/edth.12411>
- Borges, C. D., & Santos, M. A. (2005). Aplicações da técnica do grupo focal: Fundamentos metodológicos, potencialidades e limites. *Revista Da SPAGESP*, 6(1), 74–80.
- Bozkurt, A., Karadeniz, A., Bañeres, D., Guerrero-Roldán, A., & Rodríguez, M. E. (2021). Artificial Intelligence and Reflections from Educational Landscape: A Review of AI Studies in Half a Century. *Sustainability*, 13, 800. <https://doi.org/10.3390/SU13020800>

- Bridges, W. (1991). *Managing Transitions: Making the Most of Change*. Da Capo Press.
- Chichekian, T., & Benteux, B. (2022). The potential of learning with (and not from) artificial intelligence in education. *Frontiers in Artificial Intelligence*, 5, 903051. <https://doi.org/10.3389/frai.2022.903051>
- Cochran-Smith, M., Craig, C. J., Orland-Barak, L., Cole, C., & Hill-Jackson, V. (2022). Agents, Agency, and Teacher Education. *Journal of Teacher Education*, 73(5), 445–448. <https://doi.org/10.1177/00224871221123724>
- Coenen, M., Coenen, T., Stamm, A., Stucki, G., & Cieza, A. (2012). Individual interviews and focus groups with patients with rheumatoid arthritis: A comparison of two qualitative methods. *Quality of Life Research*, 21, 359–370. <https://doi.org/10.1007/s11136-011-9943-2>
- Copur-Gencturk, Y., Li, J., Cohen, A. S., & Orrill, C. H. (2024). The impact of an interactive, personalized computer-based teacher professional development program on student performance: A randomized controlled trial. *Computers and Education*, 210, 104963. <https://doi.org/10.1016/j.compedu.2023.104963>
- Council of Europe. (2022). *Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law*. Retrieved October 21, 2023, from <https://rm.coe.int/prems-092922-gbr-2517-ai-and-education-txt-16x24-web/1680a956e3>
- Crary, J. (2022). *Scorched Earth: Beyond the Digital Age to a Post-Capitalist World*. Verso.
- Cukurova, M., Miao, X., & Brooker, R. (2023). Adoption of artificial intelligence in schools: Unveiling factors influencing teachers' engagement. In N. Wang, G. Rebollo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova (Eds.), *Proceedings of the international conference on artificial intelligence in education (AIED 2023), lecture notes in computer science* (pp. 151–163). Springer Nature Switzerland. <https://doi.org/10.48550/arXiv.2304.00903>
- Damşa, C. (2014). Shared Epistemic Agency and Agency of Individuals, Collaborative Groups, and Research Communities. In J. L. Polman, E. A. Kyza, D. K. O'Neill, I. Tabak, W. R. Penuel, A. S. Jurov, K. O'Connor, T. Lee, & L. D'Amico (Eds.), *Learning and Becoming in Practice: The International Conference of the Learning Sciences (ICLS) 2014* (pp. 440–447). International Society of the Learning Sciences.
- Darling-Hammond, L., Hyster, M. E., & Gardner, M. (2017). Effective teacher professional development. *Learning Policy Institute*, 10(54300/122), 311.
- Ellul, J. (1964). *The technological society*. Vintage Books.
- European Schoolnet. (2021). *Artificial intelligence role in K12 education: Agile collection of information*. Retrieved January 30, 2024, from <https://resetedu.eu/wp-content/uploads/2021/11/4.-LIDIJA-KRALJ-Group-Discussion.pdf>
- European Union. (2018). *Supporting school innovation across Europe*. Publications Office of the European Union. <https://doi.org/10.2766/466312>
- Fathi, J., Rahimi, M., & Derakhshan, A. (2024). Improving EFL learners' speaking skills and willingness to communicate via artificial intelligence-mediated interactions. *System*, 121, 103254. <https://doi.org/10.1016/j.system.2024.103254>
- Gaona, J., López, S. S., & Montoya-Delgadillo, E. (2022). Prospective mathematics teachers learning complex numbers using technology. *International Journal of Mathematical Education in Science and Technology*, 1–30. <https://doi.org/10.1080/0020739X.2022.2133021>
- Gertz, N. (2016). Designing responsibility. In M. van der Velden, M. Strano, H. Hrachvec, J. Abdelnour Nocera, & C. Ess (Eds.), *Culture, technology, communication: Common worlds, different futures? Proceedings of the tenth international conference on culture, technology, communication* (pp. 191–201). Retrieved May 31, 2024, from https://www.researchgate.net/publication/312591793_Designing_Responsibility
- Guest, G., Namey, E., & McKenna, K. (2016). How many focus groups are enough? Building an Evidence Base for Non-Probability sample sizes. *Field Methods*, 29(1), 3–22. <https://doi.org/10.1177/1525822X16639015>
- Hall, G. E., & Hord, S. M. (1987). *Change in schools: Facilitating the process*. State University of New York Press.
- Hargreaves, A., & Fullan, M. (2012). *Professional capital: Transforming teaching in every school*. Teachers College Press.
- Hennink, M. M., Kaiser, B. N., & Weber, M. B. (2019). What Influences Saturation? Estimating Sample Sizes in Focus Group Research. *Qualitative Health Research*, 29(10), 1483–1496. <https://doi.org/10.1177/1049732318821692>

- Hrastinski, S., Olofsson, A. D., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., Jaldemark, J., Ryberg, T., Öberg, L.-M., Fuentes, A., Gustafsson, U., Humble, N., Mozelius, P., Sundgren, M., & Utterberg, M. (2019). Critical Imaginaries and Reflections on Artificial Intelligence and Robots in Postdigital K-12 Education. *Postdigital Science and Education*, 1(2), 427–445. <https://doi.org/10.1007/s42438-019-00046-x>
- Hudlicka, E. (2016). Virtual affective agents and therapeutic games. In D. D. Luxton (Ed.), *Artificial Intelligence in Behavioral and Mental Health Care* (pp. 81–115). Elsevier.
- Kowalczyk-Waledziak, M., Korzeniecka-Bondar, A., Danilewicz, W., & Lauwers, G. (Eds) (2019). *Rethinking teacher education for the 21st century. Trends, challenges and new directions*. Verlag Barbara Budrich. <https://library.oapen.org/bitstream/handle/20.500.12657/23733/1006411.pdf?sequence=1&isAllowed=y>
- Krueger, R. A., & Casey, M. A. (2015). *Focus groups: A practical guide for applied research*. Sage publications.
- Kukulka-Hulme, A., Beirne, E., Conole, G., Costello, E., Coughlan, T., Ferguson, R., FitzGerald, E., Gaved, M., Herodotou, C., Holmes, W., Mac Lochlainn, C., Nic Giolla Mhichíl, M., Rienties, B., Sargent, J., Scanlon, E., Sharples, M., & Whitelock, D. (2020). *Innovating pedagogy 2020: Open University innovation report 8*. The Open University.
- Learning Forward. (2022). *Standards for professional learning*. Retrieved September 10, 2023, from <https://www.learningforward.org/standards-for-professional-learning>
- Liu, S., & Kunnan, A. (2016). Investigating the Application of Automated Writing Evaluation to Chinese Undergraduate English Majors: A Case Study of WriteToLearn. *CALICO*, 33(1), 71–91. <https://doi.org/10.1558/cj.v33i1.26380>
- Liu, D. (2023). *How Sydney educators are building 'AI doubles' of themselves to help their students*. Teaching@Sydney Homepage. Retrieved January 13, 2024, from <https://educational-innovation.sydney.edu.au/teaching@sydney/how-sydney-educators-are-building-ai-doubles-of-themselves-to-help-their-students/>
- Long, T., & Johnson, M. (2000). Rigour, reliability and validity in qualitative research. *Clinical Effectiveness in Nursing*, 4(1), 30–37. <https://doi.org/10.1054/cein.2000.0106>
- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education – Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization. Retrieved March 30, 2023, from <https://unesdoc.unesco.org/ark:/48223/pf00000376709>
- Mousavinasab, E., Zarifasanaiey, N., Niakan Kalhori, R. S., Rakhshan, M., Keikha, L., & Ghazi Saedi, M. (2018). Intelligent tutoring systems: a systematic review of characteristics, applications, and evaluation methods. *Interactive Learning Environments*, 29(1), 142–163. <https://doi.org/10.1080/10494820.2018.1558257>
- Mouta, A., Paulino, A., Ferreira, J., & Couto, F. (2015). Pedagogy as a Techné: Meaningful ICT integration in formal learning scenarios. In A. M. Teixeira, A. Szucs, & I. Mazar (Eds.), *Expanding learning scenarios: Opening out the educational landscape* (pp. 64–72). <https://doi.org/10.38069/edenc-onf-2015-ac-0009>
- Mouta, A., Pinto-Llorente, A. M., & Torrecilla-Sánchez, E. M. (2023a). Uncovering blind spots in education ethics: Insights from a systematic literature review on artificial intelligence in education. *International Journal of Artificial Intelligence in Education*. <https://doi.org/10.1007/s40593-023-00384-9>
- Mouta, A., Torrecilla-Sánchez, E. M., & Pinto-Llorente, A. M. (2023b). Design of a future scenarios toolkit for an ethical implementation of artificial intelligence in education. *Education and Information Technologies*. <https://doi.org/10.1007/s10639-023-12229-y>
- Nagle, B., & Williams, N. (2013). *Methodology brief: Introduction to the focus group*. Center for Assessment, Planning and Accountability. Retrieved December 2, 2023, from <http://www.mmconnect.com/projects/userfiles/file/focusgroupbrief.pdf>
- Obhi, S. S., & Hall, P. (2011). Sense of agency in joint action: influence of human and computer co-actors. *Experimental Brain Research*, 211(3–4), 663–670. <https://doi.org/10.1007/s00221-011-2662-7>
- OECD. (2018). *Curriculum flexibility and autonomy in Portugal - an OECD review*. OECD. Retrieved April 18, 2023, from [https://www.poch.portugal2020.pt/pt-pt/Candidaturas/Documents/Curriculum-Flexibility-and-Autonomy-in-Portugal-an-OECD-Review%20\(002\).pdf](https://www.poch.portugal2020.pt/pt-pt/Candidaturas/Documents/Curriculum-Flexibility-and-Autonomy-in-Portugal-an-OECD-Review%20(002).pdf)
- Ogbeifun, E., Agwa-Ejon, J., Mbohwa, C., & Pretorius, J. H. (2016, March). The Delphi technique: A credible research methodology [Paper presentation]. In *2016 international conference on industrial engineering and operations management*. Kuala Lumpur, Malaysia.

- Popenici, S. A., & Kerr, S. (2017). Exploring the impact of artificial intelligence on teaching and learning in higher education. *Research and Practice in Technology Enhanced Learning*, 12(1), 22. <https://doi.org/10.1186/s41039-017-0062-8>
- Priestley, M., Biesta, G., & Robinson, S. (2015). *Teacher agency: An ecological approach*. Bloomsbury.
- Raise. (2024). *Responsible AI for Social Empowerment and Education*. Retrieved March 1, 2024, from <https://raise.mit.edu/>
- Roberts, K., Dowell, A., & Nie, J. B. (2019). Attempting rigour and replicability in thematic analysis of qualitative research data; a case study of codebook development. *BMC Medical Research Methodology*, 19, 66. <https://doi.org/10.1186/s12874-019-0707-y>
- Rogers, E. M. (2003). *Diffusion of Innovations* (5th ed.). Free Press.
- Rubie-Davies, C. M. (2015). *Becoming a High Expectation Teacher: Raising the Bar*. Routledge.
- Santos, J. M. (2012). *Introdução à Ética*. Sistema Solar.
- Severance, S., Penuel, W., Sumner, T., & Leary, H. (2016). Organizing for Teacher Agency in Curricular Co-Design. *Journal of the Learning Sciences*, 25, 531–564. <https://doi.org/10.1080/10508406.2016.1207541>
- Silva, I. S., Veloso, A. L., & Keating, J. B. (2014). Focus group: Considerações teóricase metodológicas. *Revista Lusófona de Educação*, 26, 175–190.
- Silva, P. C., & Fortunato, M. (2021). Modera, observa, escuta e foca-te na conversa de grupo – uma reflexão crítica. In P. Sá, A. P. Costa, & A. Moreira (Eds.), *Reflexões em torno de metodologias de investigação: Recolha de dados* (vol. 2, pp. 9–10). UA Editora. https://ria.ua.pt/bitstream/10773/30772/3/Metodologias%20investigacao_Vol2_Digital.pdf
- Simondon, G. (2008). *Imagination et invention (1965–1966)*. La Transparence.
- Sudina, E., & Plonsky, L. (2024). The effects of frequency, duration, and intensity on L2 learning through Duolingo: A natural experiment. *Journal of Second Language Studies*, 7(1), 1–43. <https://doi.org/10.1075/jsls.00021.plo>
- The Institute for Ethical AI in Education (2021). *The Ethical Framework for AI in Education report*. Retrieved March 22, 2022, from <https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>
- Thompson, G. (2021). *The global report on the status of teachers 2021*. Education International. Retrieved June 18, 2023, from <https://www.ei-ie.org/file/230>
- Trindade, R. (2012). *O movimento da educação nova e a reinvenção da escola: Da afirmação de uma necessidade aos equívocos de um desejo*. U. Porto Editorial.
- Tundrea, E. (2020). Artificial Intelligence in Higher Education: challenges and opportunities. *INTED2020 Proceedings*, 2041–2049. <https://doi.org/10.21125/inted.2020.0644>
- UNESCO. (2015). *SDG 4 Indicators*. Retrieved February 2, 2023, from <https://www.education-progress.org/en/indicators>
- UNESCO. (2019). *Beijing Consensus on Artificial Intelligence and Education*. Retrieved February 2, 2023, from <https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- UNESCO. (2021). *Reimagining our futures together: A new social contract for education. students*. Retrieved February 2, 2023, from <https://creativecommons.org/licenses/by-sa/3.0/igo/>
- UNESCO. (2022). *K-12 AI curricula: A mapping of government-endorsed AI curricula*. Retrieved February 1, 2024, from <https://unesdoc.unesco.org/ark:/48223/pf0000380602>
- UNESCO. (2023a). *Guidance for Generative AI in Education and Research*. United Nations Educational, Scientific and Cultural Organization. Retrieved December 12, 2023, from <https://www.unesco.org/en/open-access/cc-sa>.
- UNESCO. (2023b). *Draft AI competency frameworks for teachers and for school students*. Retrieved November 24, 2023, from <https://www.unesco.org/en/digital-education/ai-future-learning/competency-frameworks>
- UNESCO. (2023c). *AI competency frameworks for school students and teachers*. Retrieved December 11, 2023, from <https://www.unesco.org/en/digital-education/ai-future-learning/competency-frameworks>
- Wang, T., Lund, B. D., Marengo, A., Pagano, A., Mannuru, N. R., Teel, Z. A., & Pange, J. (2023). Exploring the potential impact of artificial intelligence (AI) on international students in higher education: Generative AI, chatbots, analytics, and international student success. *Applied Sciences*, 13(11), 6716. <https://doi.org/10.3390/app13116716>
- Wright, E. (2010). *Envisioning real utopias*. Verso.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Education and Information Technologies

Authors and Affiliations

Ana Mouta¹  · Eva María Torrecilla-Sánchez²  · Ana María Pinto-Llorente² 

✉ Ana Mouta
mouta.ana@usal.es

Eva María Torrecilla-Sánchez
emt@usal.es

Ana María Pinto-Llorente
ampintoll@usal.es

- ¹ Grupo de Investigación en Interacción y Elearning (GRIAL), Programme in Education in the Knowledge Society, University of Salamanca, Salamanca, Spain
- ² Instituto Universitario de Ciencias de La Educación (IUCE), Grupo de Investigación en Interacción y Elearning (GRIAL), University of Salamanca, Salamanca, Spain

APPENDIX E

COURSE GUIDEBOOK AND SESSION PLANS (SESSIONS 1–8)

This appendix presents the full contents of the “From Automation to Autonomy: Educational Agency with AI” professional development course. It includes an overview of the course rationale, objectives, target audience, delivery modalities, session structure, pedagogical approaches, and digital tools used. In addition, it provides detailed descriptions of all four modules and their corresponding eight sessions, along with guidance on facilitation, assessment, and community-building practices.

**“FROM AUTOMATION TO
AUTONOMY: EDUCATIONAL
AGENCY WITH AI”**

**TEACHER PROFESSIONAL DEVELOPMENT
Course Guidebook and Session Plans**

COURSE OBJECTIVES

This course supports educators in developing critically informed and purposeful pedagogical uses of Artificial Intelligence in education, with a focus on strengthening educational agency in its subjective, intersubjective, and collective dimensions.

Participants are invited to:

- Critically reflect on socio-technical imaginaries and their influence on schools and educational futures.
- Develop meaningful criteria to support intentional and agency-oriented AI integration in schools.
- Integrate socio-technical perspectives into classroom practice to foster authorship, ethical awareness, and educational agency.
- Co-construct a sustainable Community of Practice grounded in collective agency, ethical responsibility, and shared pedagogical commitment.

TARGET AUDIENCE

K–12 teachers, teacher educators, school leaders, coordinators, administrators, and other education professionals interested in agentic pedagogical practices, digital ethics, and the meaningful integration of AI in education.

TOTAL DURATION

12 hours (across 8 core sessions of approximately 90 minutes each, combining in-person or online synchronous delivery with exploratory and reflective asynchronous tasks).

DELIVERY METHODS

MODALITY	DESCRIPTION
In-person	Sessions are delivered face-to-face in a shared physical space, allowing for real-time interaction, collaborative activities, and immediate feedback. Particularly relevant for fostering collective agency and promoting practice-based learning.
Blended (B-learning)	Combines face-to-face and online activities, offering flexibility while preserving opportunities for in-person collaboration and critical reflection.
Online	Fully remote format. Suitable when supported by well-structured platforms, clear communication and collaboration channels, and active facilitation to ensure participant engagement and continuity.

Recommended: In-person or Blended delivery, with Canva as the main support platform.

COURSE DYNAMICS

Sessions include:

1. Review of previous sessions
2. Hands-on and Minds-on activities
3. Tinkering with AI-integrated applications
4. Individual and collective reflection
5. Asynchronous task (task diary)
6. Self-assessment and peer feedback

Approaches include problem-based learning, ethical role-play, collaborative activity design, and classroom-based role-taking.

PLATFORMS & TOOLS

- Canva (course site): Visual support material and integrated activities
- Character.AI: Dialogic ethical simulations
- ChatGPT & Gemini: Prompt design and authorship
- Elicit: AI-supported inquiry planning and authorship through curriculum-infused research tasks
- Mentimeter: Interactive polls and reflections
- Miro: Concept mapping and collaborative timelines
- Wayground: AI-powered assessments

Continuity & Community: Session 8 focuses on final integration and the creation of a Community of Practice. A shared drive or digital platform is suggested for resource sharing and ongoing collaboration.

SUPPORT MATERIALS

- Session Plans (.pdf)
- Task Diaries (.pdf)
- Readings and video resources

COURSE STRUCTURE

MODULE 1: COLLECTIVE AGENCY. WHERE DO WE STAND IN THE EDUCATION AND AIED LANDSCAPE?

This module introduces the course by grounding the abstract discourse on AI in collective narratives and shared educational experiences. It affirms teachers' roles as epistemic agents in shaping socio-technical developments.

Sessions in this module:

- Session 1. EdTech Imaginaries and Socio-technical Assemblages
- Session 2. From Users to Shapers: Agency, AI, and Shared Responsibility
- Session 3. Beyond the Algorithm: Human Stories in AI Education
- Session 4. Collective Agency in Action: AI and Participatory Practices in Schools

MODULE 2: SUBJECTIVE AGENCY. RELATIONAL PEDAGOGY IN AN AI AGE

This module explores subjective agency, focusing on students' experiences of intention, self-regulation, and self-reflectiveness. Ethical dilemmas are used as pedagogical thresholds that provoke exploration and reflection, not only about students' agency, but also about the role of teachers as co-auditors of AI use in classrooms.

Sessions in this module:

- Session 5. Invisible Scripts: AI, Student Selves, and Agency
- Session 6. Reclaiming Agency and Meaning in AI-Classroom Interactions

MODULE 3: INTERSUBJECTIVE AGENCY. SOCIALISATION AND TOGETHERNESS

This module focuses on intersubjective agency, challenging narratives of automation and efficiency by prioritising relational pedagogies grounded in participation, care, dialogue, and human connection.

Session in this module:

- Session 7. Intersubjective Agency: Socialisation and Togetherness

MODULE 4: DISTRIBUTED AGENCY. SUSTAINING COLLECTIVE PRAXIS

This final module focuses on distributed agency, promoting co-responsibility and interdependence beyond the course itself. Participants are invited to reflect on their practice and co-create a sustained professional learning community.

Session in this module:

- Session 8. Integration and Future Praxis

ACTIVITY SHEETS

(SESSIONS 1-8)

MODULE 1. COLLECTIVE AGENCY. SESSION 1. EDTECH IMAGINARIES AND SOCIOTECHNICAL ASSEMBLAGES

MAIN GOALS

- To promote critical awareness of how educational technologies are historically imagined, socially constructed, and ideologically shaped through dominant narratives and pre-existing sociotechnical systems.
- To foster participants' sense of collective agency and ethical responsibility in shaping educational futures, by interrogating past imaginaries and exploring alternative, more comprehensive visions of AI in education.
- To enhance educators' ability to reflect on the conceptual foundations of intelligence, agency, and innovation, and to connect these reflections to their own pedagogical values, professional contexts, and technological choices.

* Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.

SUGGESTED RESOURCES FOR FURTHER EXPLORATION: <https://aeon.co/essays/on-the-dark-history-of-intelligence-as-domination> (**Online article**).

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: what do we expect from this journey?	1. PRESENTATION OF TRAINER AND PARTICIPANTS	To promote engagement by creating a collaborative learning environment, mapping participants' initial expectations, and setting a tone of collaborative meaning-making.	Open discussion	Exploration	NA	NA	10'
	2. DISCUSSION OF COURSE TERMS ("AGENCY", "AI", OR "EDUCATION"): VISUAL INTERPRETATION ACTIVITY	To stimulate a critical exploration of the concepts of "agency", "AI", and "education" to develop nuanced and comprehensive understandings.	Collaborative analysis, voting, discussion	Recall, exploration, perspective-taking, introspection, insight, and decision-making, synthesis, meta-reflection	Course's Canva platform; Projected Images, Mentimeter (AI tool) - https://tinyurl.com/yc57d5a4	Group discussion	20'
Conceptual Foundations: what do we mean when we say "intelligence"?	3. "INTELLIGENCE" AS A CONTESTED AND CONTROVERSIAL CONCEPT	To critically reflect on the historical and sociocultural construction of "intelligence", its appropriation within dominant technological paradigms, and the ways in which definitions influence educational governance and practice.	Expository and conceptual explanation, visual support, guided reflection	Abstraction, exploration, synthesis	Course's Canva platform; link to Synthesia video clip (AI tool) - https://shorturl.at/WFmj5	NA	15'

	<p>Description: Participants will explore a visual timeline presented through a Synthesia AI video, tracing the evolution of the concept of “intelligence” from 19th-century philosophy through developments in neuroscience and developmental psychology, up to contemporary debates shaped by AI. The timeline includes key thinkers, shifting definitions, and disciplinary perspectives. Participants will have the opportunity to critically reflect on how the meaning of “intelligence” has changed over time, how it has been shaped by sociocultural and scientific contexts, and why it remains a contested and controversial concept.</p>						
EdTech Imaginaries: who gets to imagine the future?	<p>4. VIDEO REMIX (1965 LEARNING MACHINE VS. GPT-4O DEMO); GROUP DISCUSSION ON EDUCATIONAL FUTURES</p>	<p>To explore how past and current visions of AI in education are shaped by</p>	<p>Expository presentation with visual support</p>	<p>Exploration, cognitive dissonance, insight</p>	<p>Course’s Canva platform; link to Video on YouTube: https://www.youtube.com/watch?v=u6YQfu0uTw</p>	<p>NA</p>	<p>5’</p>
	<p>Description: Participants will watch a remixed video sequence combining two representations of educational technology, a 1965 film on learning machines and a 2024 demonstration of GPT-4o in an educational context.</p>						
Closing: wrap-up	<p>5. TIMELINE OF EDTECH IMAGINARIES – INTERACTIVE MIRO BOARD TASK</p>	<p>To foster critical and forward-looking engagement with educational technology imaginaries, encouraging participants to recognise their own agency, values, and visions in shaping the role of AI in education.</p>	<p>Group task, conceptual mapping</p>	<p>Exploration, abstraction, meta-reflection, cognitive-emotional dissonance, perspective-taking, insight, integration</p>	<p>Miro board (AI tool) - https://shorturl.at/hZvKS; slides</p>	<p>Group discussion</p>	<p>30’</p>
	<p>Description: Working in small groups, participants will use an interactive Miro board to construct a visual timeline of educational technology imaginaries. They will begin by sequencing a set of grey boxes, each representing a reference point, in chronological order. Next, they will match a series of yellow quote clouds to the corresponding sources. Following the group activity, a series of slides will be presented showing the correct associations between each quote and its respective author or source. Participants will be invited to identify recurring themes, such as personalisation, automation, efficiency, and innovation, and to consider the role of educational stakeholders in sustaining or reshaping these imaginaries.</p>						
	<p>6. TASK DIARY: SCHOOL INNOVATION AND IMAGINED FUTURES</p>	<p>To prompt critical reflection on school innovation as a sociotechnical construct and to foster awareness of educators’ collective agency and responsibility in imagining, shaping, and narrating the role of technology in education.</p>	<p>Individual writing</p>	<p>Exploration, introspection, integration</p>	<p>Digital Task diary</p>	<p>Self-assessment</p>	<p>10’</p>
<p>Description: The activity concludes with a reflection on the question: “<i>What does school innovation associated with technology really refer to? Aren’t we simply re-enacting what we’ve already imagined?</i>”. After this group reflection, participants will individually respond to four prompts in their Task Diary, encouraging them to link the exercise to broader educational debates. Key questions include: “<i>Which imaginaries on the timeline feel still alive today?</i>”; “<i>What images of the future are circulating in schools today?</i>”; “<i>Do they still resonate with students?</i>”; “<i>How can we, as educators, create openings to imagine different futures: less predictable, and more our own?</i>”; “<i>Looking back on your experience using Miro, what potential did you see in it for supporting your pedagogical practice? What barriers did you encounter and do you think it’s worth overcoming them for the educational value it could offer?</i>”. Participants will be asked to consider how different images of the future shape our understanding of the present, and what role educators might play in creating new collective narratives.</p>							

MODULE 1. COLLECTIVE AGENCY. SESSION 2. FROM USERS TO SHAPERS: AGENCY, AI, AND SHARED RESPONSIBILITY

MAIN GOALS

- To support teachers in recognising how different levels of agency (subjective, intersubjective, collective, and distributed) operate in educational contexts, particularly in AI-mediated learning environments.
 - To deepen teachers' understanding of the distinct educational purposes of qualification, socialisation, and subjectification, and how these purposes interact with different forms of student, teacher, and institutional agency.
- * Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.*

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from Session 1	1. SESSION BRIDGE: FROM IMAGINARIES TO AGENCY	To help participants synthesise key insights from Session 1 by briefly revisiting its main themes, thereby connecting the conceptual groundwork to the current session's focus on agency.	Expository	Recall	NA	NA	5'
	Description: At the start of the session, the trainer will briefly revisit the main conclusions from Session 1, highlighting how educational technologies are shaped by imagined futures, how AI and EdTech influence assumptions about learning, intelligence, and innovation, and how these tools can either amplify or constrain the roles of teachers and learners.						
Conceptual Foundations: what do we mean when we say "agency"?	2. PRESENTATION OF DIFFERENT TYPES OF AGENCY (SUBJECTIVE, INTERSUBJECTIVE, COLLECTIVE)	To foster a comprehensive understanding of human agency in education.	Expository	Attention, exploration, abstraction	Course's Canva platform	Informal questioning	15'
	Description: The trainer will use the course's Canva site to introduce participants to the four key dimensions of human agency: subjective, intersubjective, collective, and distributed. Each type will be briefly defined and illustrated with quotes from teachers, drawn from an Educational Design Research project on AI, agency, and education.						
	3. MATCHING AGENCY LAYERS VIA MENTIMETER	To deepen understanding of the four dimensions of agency by enabling teachers to critically explore how educational resources and practices can either foster or constrain it, as reflected in the perspectives of other educators.	Interactive reflection, dialogic scaffolding	Exploration, perspective-taking, cognitive-emotional dissonance, differentiation, insight, collaborative reasoning, integration	Mentimeter	Peer-discussion-based reflection	25'
Description: Participants will work in pairs to explore a selection of teacher quotes, collected through focus group discussions on AI in education. Each quote reflects a specific dimension of agency. Participants will enter Mentimeter (https://tinyurl.com/ym9ezsn9), where the quotes are displayed. For each one, they will: (1) read and reflect on the quote together, discussing which layer of agency it best represents (subjective, intersubjective, collective, distributed) – 2 minutes per quote; (2) submit their response via Mentimeter. As results appear in real time, participants will reflect on similarities or differences between their interpretations and those of the wider group, with the trainer's support.							

	4. TASK DIARY: AGENCY IN OUR OWN PRACTICES	To integrate learning about agency with teachers' own experiences, enabling connection, generalisation, and learning transfer.	Reflective writing, autobiographical mapping and connection	Exploration, differentiation, integration	Task Diary. M1.S2.TD. Agency: https://tinyurl.com/3swceszn	Self-assessment through written reflection	10'
Description: Participants map a moment from their own teaching experience where their agency was strongly present, limited, or contested, whether by a system, another stakeholder, a top-down decision, or a technological support or constraint, using the Task Diary sheet "M1.S2.TD.Agency".							
	5. PROMPTING EDUCATIONAL PURPOSE	To foster teachers' insight into the purposes of education, how AI tools align with or constrain goals beyond performance, and how agency operates within each educational purpose.	Expository method, collaborative-task design, scenario-based learning, guided critical reflection	Attention, exploration, differentiation, introspection, perspective-taking, insight, integration	Course's Canva platform; Task Diary. M1.S2.TD.EducationPurposes: https://tinyurl.com/5drus3rh ; Character.AI platform	Formative group discussion	35'
Description: In this activity, trainer presents Biesta's framework using the Canva slides and facilitates reflection on how common AI tools (e.g., data dashboards, personalised platforms) serve qualification. Teachers are invited to connect these patterns to previous learning on types of agency (subjective, intersubjective, collective, distributed). After this, teachers are randomly assigned into small groups of three and given one of Biesta's educational purposes: qualification, socialisation, or subjectification. Each group receives a pre-written context from the Task Diary sheet "M1.S2.TD.EducationPurposes", featuring a fictional student facing a dilemma linked to their assigned educational purpose. Based on this context, they will develop a prompt, which will be used to initiate a live, 5-minute conversation with a pre-set Character AI figure (e.g., a school psychologist) via the Character.AI platform. After engaging in the simulated dialogue, groups reflect on how the AI responded, whether it aligned with or expanded the educational purpose, supported the student's agency, and addressed the emotional landscape of the situation.							

MODULE 1. COLLECTIVE AGENCY. SESSION 3. BEYOND THE ALGORITHM: HUMAN STORIES IN AI EDUCATION

MAIN GOALS

- To strengthen teachers' capacity to critically analyse the role of education in the context of AI, by exploring the sociotechnical and human dimensions of algorithmic systems.
 - To enable teachers to engage with ethical, emotional, and political questions surrounding AI, encouraging dialogue and meaningful pedagogical choices.
 - To foster a collective sense of responsibility and possibility in shaping how technology is imagined, integrated, interpreted, and narrated within educational spaces.
- * Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.*

SUGGESTED RESOURCES FOR FURTHER EXPLORATION: Chabert, J.-L., & Martzloff, J.-C. (1999). *A history of algorithms: From the pebble to the microchip* (D. Stump, Trans.). Springer. (Book) Pasquinelli, M. (2021, January 14). *A social history of algorithms* [Video]. Fine Arts Museums of San Francisco. <https://www.youtube.com/watch?v=YcjC1aDlPnw> (Video)

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from Session 2	1.AGENCY SNAPSHOTS	To quickly activate prior learning and prepare teachers for a deeper exploration by connecting the concept of agency to real-world classroom moments.	Paired discussion, prompt-based reflection	Recall, exploration, synthesis, exploration	NA	Informal observation of engagement and relevance of examples shared in pairs	15'
	Description: To begin the session, participants will engage in a brief individual reflection activity designed to reactivate their understanding of the four dimensions of agency (subjective, intersubjective, collective, and distributed) through scenarios related to AI in education. The trainer will read four short statements aloud, each reflecting a different type of agency, and participants will silently consider which dimension each scenario best represents. For example: a teacher reworks AI-generated student feedback to account for emotional and cultural context (subjective); two teachers discuss how to interpret and adjust AI feedback on student writing (intersubjective); a school team of students, teachers, and parents co-develop ethical guidelines for AI use (collective); and a teacher blends AI-generated lesson content with her own ideas to suit her class context (distributed). No discussion or sharing is required, keeping the activity low-pressure while priming participants for deeper exploration. Participants may also briefly share or discuss their thoughts aloud, if they wish, to compare interpretations and spark dialogue.						
Development: The hidden side of Data Curation	2. DECODING THE LAYERS: UNDERSTANDING ALGORITHMS AS CULTURAL PRACTICE	To explore algorithms not merely as technical tools, but as sociotechnical constructs shaped by human decisions, values, and histories, so that teachers can assume their role as narrators, mediators, and decision-makers, and exercise their collective agency in shaping how these systems are integrated into pedagogical practice.	Visual-supported exposition, conceptual exploration, historical-contextual reflection, facilitated discussion	Exploration, abstraction	Course's Canva platform, link to Youtube: https://www.youtube.com/watch?v=z11pYvaLctY	Informal observation of engagement and relevance of examples shared in pairs	30'

	<p>Description: This session guides participants through a layered exploration of algorithms, beginning with how AI and machine learning are commonly explained, and then revealing the deeper cultural and historical roots behind them. So, participants begin by reviewing visual slides on Canva that explain key terms, such as AI, Machine Learning, Deep Learning, and Generative AI, in clear, accessible terms with examples. Then, participants watch a short video on Dürer’s magic square, a historical example of structured, symbolic thinking. This provokes reflection on how algorithmic logic has deep cultural and intellectual roots that predate modern computing. The trainer presents a brief historical insight on magic squares as proto-algorithms, drawing on Chabert’s work to show how algorithms are not culturally neutral, but part of a long tradition of human attempts to structure knowledge, encode patterns, and express symbolic meaning. The session concludes by drawing connections between these insights and the classroom: how can we teach algorithms in a way that values transdisciplinarity, diverse intellectual traditions, and critical awareness? Participants are encouraged to consider how this perspective could be integrated into their own teaching.</p>						
	<p>3. THE HIDDEN SIDE OF DATA CURATION AND DATASET LABELLING</p>	<p>To enable teachers to engage with the human cost of AI development, encouraging them to reflect on their role in fostering awareness, empathy, and critical dialogue within educational contexts.</p>	<p>Visual-supported exposition, questions</p>	<p>Cognitive-emotional dissonance, ethical reasoning</p>	<p>Course’s Canva platform</p>	<p>Informal assessment through oral reflection</p>	<p>30’</p>
	<p>Description: This task invites participants to explore the often unseen human labour behind the development of AI systems. Beginning with a guided reading of the TIME article on Kenyan data labellers working for a subcontractor of OpenAI, participants will learn about the emotional toll, working conditions, and ethical questions that may surround dataset labelling. After reading, the trainer unpacks some implications: “<i>What does it mean to make AI ‘safer’ for people?</i>”; “<i>Who bears the cost of that safety?</i>”; “<i>How do human choices, labour, and biases shape seemingly ‘objective’ AI outputs?</i>”. Additionally, the task will address the environmental concerns associated with AI, such as the energy consumption of training large models and the carbon footprint of data centres. Participants will explore how AI’s environmental impact intersects with human labour, bias, and ethical considerations, prompting a broader reflection on the responsibility of AI developers and users in both social and ecological contexts.</p>						
<p>Closing: The Hidden Human Cost of Data Curation</p>	<p>4. HASHTAG REFLECTION: THE HIDDEN HUMAN COST</p>	<p>To enable teachers to compromise with their role in fostering dialogue within educational contexts through the choice of what is more meaningful to them.</p>	<p>Individual reflective expression, creative synthesis, collaborative visualisation</p>	<p>Exploration, Introspection, Integration</p>	<p>Mentimeter (Word Cloud function) – hashtags (https://tinyurl.com/34bzsy8)</p>	<p>NA</p>	<p>15’</p>
	<p>Description: To conclude the exploration of AI dataset labelling and its ethical dimensions, participants will engage in a brief creative reflection activity. Working in pairs, participants will craft a hashtag (e.g., #BehindTheAlgorithm, #HiddenLabour) that captures a key message, insight, or tension they believe should be more widely known about the hidden labour behind AI systems. Hashtags will be submitted to a Mentimeter word cloud (https://tinyurl.com/34bzsy8), creating a visual and discursive summary of the group’s insights. The activity will culminate in a group discussion based on the generated word cloud.</p>						

Table 1. Suggested resources for further exploration, activity design, and the subject to explore AI and participatory design for teacher education

MAIN GOALS

- To enable teachers to engage in transdisciplinary collective agency by examining how algorithmic decisions are made and audited in educational settings through real-world cases.
 - To reflect on how participatory approaches to AI in education can foster student agency and accountability, while supporting teachers in their roles as critical interpreters and ethical mediators of AI technologies in schools.
 - To engage teachers in exploring how AI tools can be thoughtfully integrated into lesson design in ways that support subject learning objectives, foster collective agency at a school level, and promote critical reflection on the role of automation in education.
- * Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.

SUGGESTED RESOURCES FOR FURTHER EXPLORATION: Zumeta, L., Basabe, N., Wlodarczyk, A., Bobowik, M., & Páez, D. (2016). Shared flow and positive collective gatherings. *Anales de Psicología*, 32(3), 717–727. <https://doi.org/10.6018/analesps.32.3.261651> (**Paper**. A concise summary of the scientific article is presented as a poster on the course's Canva platform.)
 Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education*. Macmillan. (**Book**)

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from session 3 Development: Algorithmic Accountability in Education and Collective Agency in action	1. RECONNECT	To help participants integrate and transfer key ideas from the last session by reflecting on their meaning and restating them in their own terms and experiences.	Pair dialogue	Recap, exploration, differentiation, insight, synthesis, transfer	NA	NA	5'
	2. AUDITING ALGORITHMS FOR EDUCATION ETHICS	To introduce the concept and relevance of algorithmic auditing in education, and how it can be promoted as part of collective agency through transdisciplinary collaboration.	Visual-supported exposition	Attention, Abstraction	Course's Canva platform	Informal questioning	10'
	3. CASE STUDY: JOHANNA AND AI SURVEILLANCE	To critically explore the ethical risks of decontextualised data in school surveillance AI, using a real-world-inspired case to evaluate bias, harm, and responsibility.	Case study analysis, guided group questioning	Perspective-taking, Exploration, Abstraction, Differentiation, Synthesis	Course's Canva platform	Group-discussion-based reflection	25'
		Description: This activity introduces participants to the concept of algorithmic auditing in educational AI systems. It outlines common risks including bias, data gaps, and unfair proxies such as ZIP codes or mental health records. Participants will be exposed to core terms like "impact assessment" and "ethical compliance", and will reflect on the importance of educator involvement in identifying risks and ensuring that AI aligns with ethical principles in education.					
		Description: This task presents participants with a real-world-inspired case involving "Johanna", a student flagged by an AI system as a risk based on contextual data (ZIP code, mental health, creative writing). Participants will be asked to take on the role of school-based auditors tasked with evaluating the fairness, transparency, and accountability of the system that flagged Johanna. In small groups, participants will: identify which aspects of the AI system failed (e.g., lack of context, poor indicators, absence of human oversight). Propose 3–5 ethical criteria or "audit questions" that any AI system in education should meet in order to align with collective agency and education ethics. Reflect on how educators, students, and communities could participate in ongoing algorithmic review and decision-making in schools.					

	4. QUIZ REFLECTION: PARTICIPATORY AI IN SCHOOLS	To consolidate participants' learning on how to foster agency when using algorithms and digital data in schools, through engagement with an AI-powered educational tool (Wayground, formerly Quizizz).	Interactive exploration, individual reflection	Recall, abstraction, evaluation	Course's Canva platform, Wayground (formerly Quizizz) link	Self-assessment quiz	10'
Description: Participants explore the AI-powered learning tool Wayground (formerly Quizizz), which uses automatic grading and AI-generated questions. Teachers will be informed that the quiz was originally generated by AI, although it has been substantially edited to ensure substantive quality. As they complete the short quiz, they will be asked to apply their conceptual understanding of key session themes, including agency, algorithmic ethics, and participatory design, to classroom-based scenarios. The quiz prompts participants to reflect on concrete educational situations, such as designing grading rubrics or managing group dynamics.							
	5. REFLECTION ON COLLECTIVE AGENCY MODULE	To synthesise key principles of collective agency and reflect on their application within democratic educational practices.	Guided discussion, concept integration, reflective dialogue	Synthesis, abstraction, meta-reflection	Course's Canva platform – posters; Task Diary Sheet M1.S4.TD.Poster	Oral reflection	10'
Description: This discussion revisits key themes of collective agency explored throughout the module, including shared decision-making, differentiated roles, and active participation. Participants reflect on how these concepts apply in their own contexts and how they can be strengthened through school practices that promote student voice and mutual accountability. As part of this reflection, teachers will choose the poster that best represents their understanding of collective agency and that they would display in their classroom. They will record their choice and reflections in Task Diary Sheet “M1.S4.TD.Poster”.							
Development and Closing: Designing for Collective Agency with AI Tools in Inquiry-Based Session Planning	6. ROLE-PLAY: CURRICULAR INFUSION FOR COLLECTIVE AGENCY	To encourage teachers to reflect on how student agency in inquiry-based tasks supports deeper engagement and meaning-making compared to automated approaches.	Scenario-based learning, collaborative session plan design, role-play	Exploration, abstraction, recall, differentiation, synthesis, perspective-taking, insight, analysis	Course's Canva platform, Elicit tool, example case description	Presentation of adapted activity and peer assessment	30'
Description: This role-play activity invites teachers to adapt an example of curricular infusion by linking subject-specific goals to broader social aims through the use of AI tools. Using a science lesson plan (presented on Canva) that incorporates Elicit as a starting point, participants work in groups to redesign the activity for their own disciplines, applying their own lesson planning templates. The original example is grounded in a science objective: helping students learn the scientific method. It contrasts two approaches to using Elicit: (1) In one scenario, students collaboratively generate research questions. (2) In the other, students rely on Elicit to generate the questions for them. Teachers will role-play both scenarios, divided into two groups, to reflect on how the automation of reflective and inquiry-based tasks can impact student engagement, reduce opportunities for agency, and limit deeper meaning-making. The activity encourages critical consideration of the trade-offs between efficiency and pedagogical depth, particularly in fostering critical thinking and ownership of learning. Each group will then present their adapted lesson design, highlighting how it promotes collective agency and responsible AI use in the classroom. The session concludes with a reflection on their design choices and how these influence the student learning experience, especially in terms of engagement, understanding, and purposeful inquiry.							

MODULE 2. SUBJECTIVE AGENCY. SESSION 5. INVISIBLE SCRIPTS: AI, STUDENT SELVES, AND AGENCY

MAIN GOALS

- To develop teacher agency as co-auditors of AI systems, developing shared language and criteria for agency-respecting AIED use in schools.
- To support teachers analysing how different AIED applications and features support or hinder the development of student agency.

* Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.

SUGGESTED RESOURCES FOR FURTHER EXPLORATION: Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science*, 1, 164–180. <https://doi.org/10.1111/j.1745-6916.2006.00011.x> (Paper).

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from session 4	1. RECONNECT	To help participants integrate and transfer key ideas from the last session by reflecting on their meaning and restating them in their own terms and experiences.	Pair dialogue	Recap, exploration, differentiation, insight, synthesis, transfer	NA	NA	5'
	<p>Description: In this opening activity, participants revisit the previous session. They reflect in pairs on the following questions: “<i>What challenged your assumptions?</i>”; “<i>What feels difficult to transfer into your own context?</i>”.</p>						
Development: Surveillance, Subjective Agency & the AIED Classroom	2. SAFETY OR SURVEILLANCE?	To develop critical awareness of how surveillance shapes identity and trust in learning environments, affecting key personal dimensions of subjective agency such as intentionality, forethought, self-regulation, and self-reflection.	Critical reading, assigned-position debate, meta-reflection	Attention, synthesis, introspection, cognitive-emotional dissonance, perspective-taking, insight, integration	Course’s Canva platform; Online publication: https://tinyurl.com/bdkvmzhh ; https://www.randomlists.com/team-generator	Quality of reasoning; contributions to the shared whiteboard	45'
	<p>Description: This activity begins with a short reading of the article “Schools use AI to monitor kids, hoping to prevent violence. Our investigation found security risks” (https://tinyurl.com/bdkvmzhh), either aloud by the facilitator or silently by participants, followed by an open discussion around the question: “<i>What stood out to you from this article?</i>”. Teachers are then divided into two groups (or smaller subgroups if needed) using an online team generator (https://www.randomlists.com/team-generator). Group A is tasked with defending the use of surveillance tools for student safety, wellbeing, and school security, while Group B will argue against their use, highlighting risks to student emotional experience, privacy, autonomy, and freedom of thought. Each group has five minutes to prepare their arguments, drawing on personal experiences and ideas, the article, and broader ethical concerns. The debate follows, with each group presenting their views and engaging in a round of clarifications and refutations. A shared digital or physical whiteboard is used to record key points. The activity concludes with a reflection phase, in which the facilitator presents visual posters on data mining, predictive manipulation, and the psychological impact of surveillance. Participants reflect on these materials using questions such as: “<i>How do these ideas reflect or challenge the arguments we just heard?</i>”; “<i>Where should the line be drawn between care and control in education?</i>”. Finally they add insights to the shared board, particularly in relation to the development of subjective agency.</p>						
	3. SUBJECTIVE AGENCY IN THE AGE OF AI	To deepen understanding of the concept of subjective agency and how its development can be fostered.	Short-conceptual recap	Exploration, abstraction, synthesis	Course’s Canva platform, Animaker video integrated into the platform	Group-discussion-based reflection	10'

	<p>Description: Building on the previous “Safety or Surveillance?” debate, this activity shifts focus from ethical tension to developmental impact. Participants explore how surveillance technologies affect subjective agency, using Bandura’s four key dimensions (intentionality, forethought, self-regulation, and self-reflection) as a framework. At this stage, the facilitator will synthesise some key takeaways, visually (through an Animaker video created for that purpose) and verbally, to support shared understanding and analysis. These insights summarise how constant surveillance can interfere in the development of subjective agency.</p>					
<p>Development and Closing</p>	<p>4. AIED FEATURES AND AGENCY</p>	<p>To develop educators’ capacity to critically assess digital tools with AI beyond efficiency or engagement.</p>	<p>Table-based analysis in pairs, critical dialogue, whole-group synthesis</p>	<p>Recall, abstraction, introspection, perspective-taking, differentiation, insight</p>	<p>Course’s Canva platform, M2.S5.TD.AITable</p>	<p>Output table 30’</p>
<p>Description: In this activity, participants work in pairs to analyse the potential psychological and developmental impacts of various AI-driven features used in education. Using a table that lists six common AI features (e.g., adaptive pacing, automated feedback, predictive analytics), each pair will evaluate how each feature could support or hinder the four dimensions of subjective agency as defined by Bandura: intentionality, forethought, self-regulation, and self-reflection (cf. Task Diary sheet “M2.S5.TD.AITable”). Participants will draw on prior input from the session (presentations, case studies, and discussions) and are encouraged to consider both risks and opportunities. The activity concludes with a short group debrief, in which insights are shared and patterns are identified (e.g., “Which features most frequently reduce agency? ”; “Which ones might be reclaimed or redesigned to support it? ”).</p>						

MODULE 2. SUBJECTIVE AGENCY. SESSION 6. RECLAIMING AGENCY AND MEANING IN AI-CLASSROOM INTERACTIONS

MAIN GOALS

- To foster connection and ownership in AI-mediated learning environments by exploring the educational implications of how students may feel when interacting with systems they do not fully understand or cannot influence.
 - To support teachers in designing classroom experiences that enhance subjective agency through choice, goal alignment, and feedback fluency.
- * Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.

SUGGESTED RESOURCES FOR FURTHER EXPLORATION: Schreiner, M. R., Neszemlyi, B., Schwarz, K. A., & Kunde, W. (2025). Goals rather than predictions determine the sense of agency. *iScience*, 112583. <https://doi.org/10.1016/j.isci.2025.112583>; Frith, C. D. (2014). Action, agency and responsibility. *Neuropsychologia*, 55, 137–142. <https://doi.org/10.1016/j.neuropsychologia.2013.09.007> (**Papers**).

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from session 5	1. RECONNECT	To help participants integrate and transfer key ideas from the last session by reflecting on their meaning and restating them in their own terms and experiences.	Pair dialogue, whole-group share-out	Recap, exploration, differentiation, insight, synthesis, transfer	NA	NA	10'
	<p>Description: In this opening activity, participants revisit the previous session. First, they reflect in pairs on the following question: “<i>Did the previous session surface any tensions between your values, available tools, and practices?</i>”. Then, in a process of “recaptioning”, they restate a key concept, takeaway, or metaphor from the session using their own words or images, which they can share with the whole group.</p>						
Development: Feeling Control in Data-Driven Classrooms	2. EMBODIED EXPERIENCE: “BALL VS. BALLOON”	To allow participants to experience the difference between fluent, intentional action and disrupted, unpredictable feedback, forming the basis for analysing agency in AI-enhanced learning.	Embodied paired simulation, guided reflection	Sensorimotor awareness, introspection, perspective-taking, conceptual abstraction, emotional-cognitive integration, transfer	One foam ball (medium) per pair; one balloon per pair (extra balloons and balls on hand in case of pops or drops); Course’s Canva platform	Group-discussion-based reflection	15'
	<p>Description: In this activity, participants work in pairs, with one taking the role of Initiator and the other as Responder. In the first round, using a foam ball, the Initiator throws the ball to the Responder, who catches and returns it in a smooth, rhythmic pattern for 6 to 8 exchanges. This establishes a sense of predictable, controlled interaction. In the second round, the ball is replaced with a balloon. The Initiator throws the balloon in the same way, but due to its unpredictable floating and wobbling, the outcome may feel delayed or misaligned with their intention. After both rounds, pairs reflect on the experience, considering the following questions: “<i>When did you feel more in control?</i>”; “<i>Did the object’s behaviour match your intention?</i>”; “<i>What did it feel like to act when outcomes were unpredictable?</i>”. These questions help explore the relationship between feedback, control, and the subjective sense of agency. These reflections are then expanded in a short plenary debrief. The facilitator links the experience to the research of Schreiner and Frith on the sense of agency (cf. “Suggested resources for further exploration”), which argue that our feeling of control over actions is not a direct perception, but rather an inference based on cues such as fluency, timing, and outcome consistency. When feedback is unpredictable or feels misaligned, even when the action is technically</p>						

FROM ETHICS TO AGENCY: PARTICIPATORY DESIGN OF A TEACHER TRAINING COURSE FOR AI IN EDUCATION

	successful, our subjective sense of agency decreases. This principle is foundational to understanding how students experience AI-based learning tools: even highly functional systems can reduce agency if they lack fluency, transparency, or a clear connection between intention and result.						
Development: Agency in Practice	3. PROMPTING AS A PRACTICE OF AUTHORSHIP	To engage participants in designing purposeful AI prompts that foster authorship and meaningful learning, highlighting the impact of prompt structure on agency, intention, and outcome.	Small group design challenge, structured reflection, peer feedback, group discussion	Exploration, analysis, differentiation, synthesis	Course's Canva platform, prompt design framework (7 guiding elements), ChatGPT or Gemini access,	Peer review of prompts using guiding questions	30'
	Description: In this activity, participants explore prompting as an intentional, communicative act that shapes not just outputs from AI, but also learning outcomes and student agency. Working in small groups, each team selects a curricular topic from their disciplinary or interdisciplinary field. Their task is to collaboratively design an educational prompt using the 7 guiding elements (Context, Instruction, Questioning, Sequencing, Iteration, Tone, Example). The goal is to construct a task that fosters a sense of control, interest, engagement with both the content and the task, and motivation in students, enabling them to feel genuinely involved in the learning process, even if the output is generated by AI. Each group drafts a structured prompt using the framework and enters it into a generative AI tool (e.g., ChatGPT or Gemini). They then evaluate the response and reflect on whether the AI's output aligns with their intention. Groups exchange prompts with others, offering constructive feedback guided by questions such as: "How might the prompt better support authorship and meaningful engagement?". To close, the whole group discusses the pedagogical implications of prompt design, particularly how subjective agency can be supported or constrained by how tasks are framed. A guided reflection helps participants identify benefits and drawbacks of co-authoring with AI, and strategies to foster learner ownership, intentionality, and deeper engagement.						
	4. CO-CONSTRUCTING CRITERIA FOR AGENCY	To explore how subjective agency can be enhanced when students define how they wish their AI-assisted work to be evaluated, based on personally meaningful criteria, before any comparison or judgment by peers occurs.	Reflective simulation, individual writing task, group discussion and guided debrief	Introspection, metacognition, perspective-taking, cognitive-emotional dissonance, integration	Course's Canva platform	Group-discussion-based reflection	35'
	Description: This activity builds on the previous prompt design challenge, where participants created a structured prompt and used AI to generate a piece of content. Now, they shift perspective and simulate a classroom practice focused on enhancing subjective agency through self-defined evaluation. Participants imagine themselves as students who have just completed a written task using AI support. Before any external assessment or peer comparison occurs, each participant privately selects four evaluation criteria from a suggested list (i.e., clarity, creativity, coherence, originality, ethical use of AI, relevance to the task, and integration of AI-generated and personal content) They then write a brief reflective note explaining why these criteria are meaningful to them and how they reflect their intention for the imagined task. This step encourages participants to consider how authorship and agency are supported when learners define what success means for them. Next, participants share their selected criteria in small groups, opening discussion around alternative assessment practices that centre learner perspective and intention. This sharing stage promotes critical dialogue about how students might reclaim a sense of ownership in AI-enhanced tasks. To close, a guided debrief in pairs invites reflection on questions such as: "How did it feel to choose your own criteria before receiving feedback?"; "Did defining your own success shift how you related to the task or the tool?"; "What does this suggest about how students might reclaim authorship when working with AI tools?". Pairs then share their conclusions with the whole group, contributing to a richer repertoire of assessment practices that can be considered and applied in the classroom.						
Closing: Wrap-up - Inter-session task	5. AGENCY JOURNAL SNAPSHOT	To consolidate insights from the session by inviting participants to reflect on their personal learning, practical applications, and next steps to support agency.	Individual written reflection, optional sharing	Metacognition, introspection, self-regulation, integration, differentiation	Course's Canva platform, M2.S6.TD.AgencySnapshot	Quality and relevance of individual takeaways	0' (Inter-session task)
	Description: In this final wrap-up activity, each participant completes the Task Diary sheet titled "M2.S6.TD.AgencySnapshot" and is invited to spend 10 minutes reflecting and responding in writing. The questions are: "What surprised me about the Co-Constructing Criteria task, either about my own choices or others'?" ; "Where in my practice could I make space for students to define success on their own terms?"; "What is one small action I could take this month to support subjective agency in my classroom, particularly when using AI tools?". Participants may keep their responses private or, if they wish, share a key takeaway or personal commitment with the group.						

MODULE 3. INTERSUBJECTIVE AGENCY. SESSION 7. INTERSUBJECTIVE AGENCY: SOCIALISATION AND TOGETHERNESS

MAIN GOALS

- To support teachers in understanding how intersubjective agency is challenged in AI-mediated educational environments, particularly in how system design influences relationships, participation, and peer dynamics.
- To deepen teachers' understanding of how belonging, shared meaning-making, and situational freedoms are central to students' socialisation processes, and how these can be shaped by the integration of AI technologies in schools.
- To support teachers in recognising how students' sense of authorship and responsibility shifts in AI-mediated learning, and to explore strategies for reinforcing ethical reflection and fostering meaningful distributed agency in classrooms.
- To promote a meaningful approach to teaching and learning by incorporating AI technologies in ways that are relevant to how agency is experienced and how educational experiences are transferred and made meaningful, while reinforcing collective teacher agency through the creation of a community of practice that extends beyond the course.

* Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.

SUGGESTED RESOURCES FOR FURTHER EXPLORATION: OBHI, S., S., & HALL P. (2011). SENSE OF AGENCY IN JOINT ACTION: INFLUENCE OF HUMAN AND COMPUTER CO-ACTORS. EXPERIMENTAL BRAIN RESEARCH, 211(3-4), 663–670. [HTTPS://DOI.ORG/10.1007/S00221-011-2662-7](https://doi.org/10.1007/s00221-011-2662-7) (PAPER)

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from session 6	2. RECONNECT	To help participants integrate and transfer key ideas from the last session by reflecting on their meaning and restating them in their own terms and experiences.	Pair dialogue, whole-group share-out	Recap, exploration, differentiation, insight, synthesis, transfer	NA	NA	5'
	Description: In this opening activity, participants revisit the previous session. They are invited to reflect in pairs on the following question: <i>“What idea from the previous session would you like to explore further or apply in practice?”</i> .						
Development: Collaborative Design	2. ALGORITHMS FOR CONNECTION	To help participants explore how belonging, meaning-making, and moral development can be fostered through design, ethical imagination, and collaborative group negotiation.	Visuals and metaphors (e.g., piano duet); small group collaborative design (3 groups, each with a distinct focus); group presentations and reflection	Exploration, perspective-taking, differentiation	Course's Canva platform	Group self-assessment; peer feedback	25'
	Description: To launch this activity, the trainer begins by introducing the concept of intersubjective agency, using the piano duet metaphor to illustrate how agency is not a solitary function but a relational process involving synchronisation, anticipation, and mutual responsiveness. This is followed by a discussion of how relational agency emerges in education, through co-action, mutual recognition, and negotiated meaning, and why it is central to belonging, meaning-making, and moral development. The trainer then presents the role of AI in relational spaces, highlighting both its potential to scaffold collaboration and its risks (e.g., limiting perspective-taking, reinforcing premature consensus, and reducing expressive freedom). With this conceptual grounding, participants are divided into three small groups, each assigned to work on one key dimension of						

	<p>relational agency: Group 1. Belonging. Designing systems that foster inclusion, recognition, and emotional safety; Group 2. Meaning-Making. Designing systems that support dialogue, co-construction, and reflection; Group 3: Moral Development. Designing systems that promote fairness, ethical reasoning, and perspective-taking. Each group is challenged to design a simple rule-based system or algorithm that embodies their assigned relational value set. They must include feedback loops (“pause and check” moments) and a set of game-like rules that reward behaviours aligned with their relational focus. In this round, participants are also asked to consider the teacher as a key relational figure, incorporating rules or dynamics that reward meaningful engagement with educators (e.g., listening and questioning a teacher’s view, co-defining goals with them). Each group assigns a “value guardian” to monitor alignment with the values throughout the process and prompt periodic reflection. The challenge culminates in a brief sharing session where designs are compared not in terms of completion, but in how well they embodied their relational goals. Peer feedback is exchanged between groups. Some examples include: “Belonging. Focus: emotional inclusion, affirmation, identity expression. ‘If a team member’s perspective is linked to a lived experience and the group validates its relevance → +4 belonging points’”; “Meaning-Making. Focus: negotiation of shared understanding, curiosity, conceptual integration. ‘If the team revises their original idea after a peer challenges the logic with a ‘what if’ question → +4 meaning-flex points’”; “Moral Development. Focus: ethical reasoning, fairness, perspective-taking, responsibility. ‘If a participant defends a marginalised voice or view, even if unpopular → +5 integrity badge’”.</p>						
<p>Development: Agency, GenAI, Authorship</p>	<p>3. “WHO ACTED HERE?”</p>	<p>To explore how the perception of who acts, human or AI, affects one's sense of agency, and how generative AI influences our understanding of authorship, and responsibility, while identifying constructive approaches to support ethical decision-making.</p>	<p>Animated video presentation, group discussion, role-play in rotating pairs (student/teacher), scenario-based reflection and dialogue</p>	<p>Introspection, exploration, perspective-taking, cognitive-emotional dissonance, insight, synthesis, integration</p>	<p>Course’s Canva platform, Animaker video integrated into the platform</p>	<p>Observation of depth and nuance in role-play and discussion</p>	<p>30’</p>
<p>Description: To open the activity, the trainer presents a short animated video created with Animaker and integrated into the course platform. The video illustrates how a person’s sense of agency shifts depending on whether they believe they initiated an action or whether it was driven by an external system. This concept is based on Obhi’s research on the perception of voluntary action, which suggests that the more we attribute an outcome to an external agent, such as AI, the less agency we feel, even if we were the ones who technically caused it. The video sets the stage for a discussion about agency as both a social and mental interpretation. Key themes include the impact of automated systems on responsibility, implications for academic integrity, the dilution of personal responsibility, weakened ethical thinking, and a reduced capacity for reflective engagement. Participants then break into small groups and rotate through four role-pair configurations, each involving a student and a teacher. The scenario involves a student who used generative AI to complete a school task. The four configurations are as follows: Scenario 1: Full Copy, Negative Response; Scenario 2: Full Copy, Constructive Response; Scenario 3: Critical Use, Negative Response; Scenario 4: Critical Use, Constructive Response.</p>							
<p>Closing: Wrap-up - Inter-session task</p>	<p>4. A Pedagogical Design Task (Shaping Agency with AI)</p>	<p>To enable teachers to apply their understanding of various layers of agency and educational purposes than can be activated in a practical, context-specific manner within their own teaching context.</p>	<p>Individual design thinking, conceptual application, reflective practice.</p>	<p>Transfer, exploration, perspective-taking, differentiation, application, synthesis, integration</p>	<p>Teachers' subject matter programme, chosen AI tools, personal notes/Task Diaries from previous sessions; Form: Module 3. Peer Evaluation of Draft Activity (7th-8th Session)</p>	<p>Activity alignment with pedagogical goals and the incorporation of feedback to refine the design.</p>	<p>30’</p>

			<p>https://tinyurl.com/mukde5nh: Form: Module 3. Evaluation of Draft Activity (8th Session) https://tinyurl.com/5n8ux59s</p>	
<p>Description: Teachers are invited to choose a specific subject matter and a theme from their programme or curriculum and design an activity incorporating an available AI tool. The main focus of this design task is to ensure that teachers explore the different layers of agency (subjective, intersubjective, collective, or distributed) and educational purposes (qualification, socialisation, or subjectification) they intend to achieve. Teachers will be encouraged to create activities that explore takeaways that can also be transferred to the home environment, fostering conversations between students and their families. During the session, it will be decided which teacher each participant will be responsible for evaluating. After refining their session plans at home, teachers will send their activity plans to the designated teacher/pair for feedback. The peer evaluator will assess the session plan based on criteria such as curriculum infusion of agency goals, meaningfulness of the activity, potential of the chosen AI tool, and feasibility (cf., Form: “Module 3. Peer Evaluation of Draft Activity (7th-8th Session)”). After receiving feedback, teachers will refine and finalise their drafts outside of the session. They will have three weeks (before the 8th session) to implement the activity in their classrooms and evaluate its relevance, before the 8th session. The evaluation will focus on learning goals, relationships covered, AI ethical-agency competencies achieved, and student feedback. Teachers will reflect on the activity's impact, considering whether it met the intended goals and how it influenced student engagement and agency (cf., Form: “Module 3. Evaluation of Draft Activity (8th Session)”).</p>				

MODULE 4. DISTRIBUTED AGENCY. SESSION 8. INTEGRATION AND FUTURE PRAXIS

MAIN GOALS

- To facilitate the integration, transfer, and generalisation of learning through exploring diverse classroom experiences, emphasising insights in ethically integrating AI and fostering distributed agency.
- To establish a collaborative "Community of Practice" among educators for ongoing peer support, resource sharing, and continued professional development in agency and ethics with AIED.
- To empower teachers to articulate their role and future steps in shaping AI in education, reinforcing a sense of collective agency and ethical responsibility beyond the course.

** Teachers will have the opportunity to reflect on this using resources powered by AI, exploring both their advantages and limitations in a pedagogically integrated, purposeful, and meaningful way.*

SUGGESTED RESOURCES FOR FURTHER EXPLORATION: OBHL, S., S., & HALL P. (2011). SENSE OF AGENCY IN JOINT ACTION: INFLUENCE OF HUMAN AND COMPUTER CO-ACTORS. EXPERIMENTAL BRAIN RESEARCH, 211(3-4), 663–670. [HTTPS://DOI.ORG/10.1007/S00221-011-2662-7](https://doi.org/10.1007/s00221-011-2662-7) (PAPER)

THEMATIC STAGES	ACTIVITY	SPECIFIC OBJECTIVES	METHODOLOGIES	LEARNING PROCESSES	RESOURCES	ASSESSMENT	DURATION
Introduction: bridge from session 6	3. CAPSTONE AND TRANSFER: BUILDING A COMMUNITY FOR COLLECTIVE AGENCY	<ul style="list-style-type: none"> • To foster a deeper understanding of the practical implications of integrating AI for agency and ethical development in diverse educational contexts. • To provide a forum for teachers to reinforce their sense of belonging to a community by presenting their implemented activities and their self-evaluations in a secure space. 	Teacher presentations (structured sharing of implemented activities), whole-group discussion and facilitated debate, collaborative decision-making for community of practice setup.	Application, synthesis, critical reflection, communication, peer learning, collaborative problem-solving, planning, commitment	NA	NA	5'
<p>Description: Participants will briefly reconnect, bridging from the previous session's learning to the subsequent implementation period. They will be invited to share a "headline" or a very brief, high-level takeaway from their experience implementing their activities over the past three weeks. The trainer will then briefly check for any immediate, overarching questions or technical issues encountered during the implementation phase. Each teacher will present their "Designing for Agency: AI in Your Classroom" activity, succinctly covering: the specific subject matter and theme, the chosen AI tool, the targeted layer of agency and educational purpose, a summary of their findings, and a brief personal reflection on how it went. The emphasis will be on sharing lessons learned through the lived experience of putting theory into practice. Following the presentations, the facilitator will lead a general debate, encouraging teachers to synthesise the collective insights and identify common themes, successful strategies, and recurring challenges across the diverse activity implementations. The discussion will be guided by core questions: <i>"Based on your reflections and observations from others' experiences, what refinements would you suggest for similar AI-integrated activities in the future?"</i>; <i>"What new, surprising, or significant insights have you gained about meaningfully integrating AI to foster distributed agency and achieve specific educational purposes?"</i>; <i>"Considering our collective journey in this course, what are your next steps for professional learning in this area, both individually and as a community?"</i>. To ensure sustained professional development and ongoing collaboration beyond the formal course structure, the group will collaboratively decide on a preferred digital tool or platform to serve as a shared environment for their "Community of Practice". This platform will be established as a resource for sharing lesson plans, refining pedagogical practices, and integrating meaningful resources that foster collective, distributed agency. Discussions will include initial forms for engagement, preferred communication channels, and potential first collaborative projects or topics of shared interest to launch their ongoing journey as a learning community in AIED, agency, and educational ethics.</p>							

APPENDIX F

EXTENDED ABSTRACT IN SPANISH

This appendix presents an extended abstract in Spanish.



VNiVERSIDAD
D SALAMANCA

**DE LA ÉTICA A LA AGENCIA:
DISEÑO PARTICIPATIVO DE UN
CURSO DE FORMACIÓN DOCENTE
SOBRE IA EN LA EDUCACIÓN**

Programa de Doctorado

Formación en la Sociedad del Conocimiento

TESIS DOCTORAL

ANA ISABEL MOUTA COSTA

DIRECTORAS

EVA MARÍA TORRECILLA SÁNCHEZ

ANA MARÍA PINTO LLORENTE

Resumen

La Inteligencia Artificial está siendo cada vez más integrada en entornos educativos, pero sus implicaciones éticas y su impacto sobre la agencia pedagógica siguen siendo insuficientemente explorados. Esta tesis investiga los desafíos éticos y las preocupaciones relacionadas con la agencia en la educación mediante un proceso de investigación basada en el *Educational Design Research*, con el objetivo de desarrollar un curso de formación para profesores de Educación Primaria y Educación Secundaria diseñado a partir de sus propias opiniones.

El estudio comienza con una revisión sistemática de literatura entre 2011 y 2022, realizada según las directrices PRISMA, que mapea el estado actual de la investigación sobre la IA en la educación. Esta fase identifica vacíos significativos en los marcos éticos, en la orientación específica para docentes y en la preservación de la agencia educativa. Sobre esta base, la investigación adopta una metodología de *futures studies*, utilizando el método Delphi para construir de manera colectiva ocho escenarios futuros. Estos escenarios exploran los imaginarios sociotécnicos que configuran las implicaciones pedagógicas de la IA, incluyendo cuestiones de equidad, evaluación, voz estudiantil y autonomía profesional.

Las fases posteriores de la investigación involucran a formadores de docentes mediante grupos focales iterativos con el fin de explorar cómo la IA transforma las dinámicas de la agencia subjetiva, intersubjetiva y colectiva en contextos educativos. Los hallazgos revelan una necesidad urgente de superar las narrativas tecnosolucionistas dominantes y, en su lugar, apoyar a los docentes en la recuperación de su rol como agentes éticos y relacionales. Estas ideas informan el diseño conjunto de un curso de desarrollo profesional que integra prácticas de aprendizaje dialógicas, experienciales y reflexivas. El curso se aloja en una plataforma personalizada basada en Canva y se estructura en torno a un marco tridimensional de agencia educativa, ofreciendo a los docentes herramientas conceptuales y prácticas para involucrarse críticamente con la IA.

Al poner en primer plano las dimensiones simbólicas, relacionales y éticas de la educación, esta tesis sostiene que la integración responsable de la IA no solo debe ser técnicamente sólida, sino también coherente con los fines fundamentales de la educación: subjetivación, cualificación y socialización. Se propone que la sostenibilidad de la agencia docente requiere una atención especial a la preservación y al cuidado del léxico educativo, un léxico que sostenga la complejidad, la apertura, el discernimiento ético, así como el deseo y la

memoria frente a las presiones algorítmicas. Es a través de este deseo que pueden emerger nuevos imaginarios de los sistemas sociotécnicos y de ecosistemas educativos más integrales.

Esta tesis ofrece cuatro aportes principales. En primer lugar, un mapeo comprensivo sobre la IA en la educación y la ética. En segundo lugar, un *toolkit* de dilemas éticos aplicados a la IA en la educación. En tercer lugar, un marco conceptual de agencia en contextos educativos mediados por IA. Y, en cuarto lugar, un curso de formación docente contextualizado y centrado en la agencia. En conjunto, estos resultados constituyen una contribución teórica y empíricamente fundamentada a los esfuerzos académicos y profesionales que buscan cultivar entornos educativos en los cuales las decisiones sobre el uso de la IA, y las condiciones bajo las cuales se implementa, se construyan colectivamente mediante procesos dialógicos y participativos que sostengan el propósito educativo, la agencia humana y el horizonte democrático de la escuela. Esta investigación contrarresta las tendencias despolitizadoras y desprofesionalizadoras de los modelos tecnocráticos al apoyar al profesorado en su compromiso crítico con la Inteligencia Artificial, en su resistencia a la automatización irreflexiva y en el cuestionamiento de la normativización algorítmica.

Palabras clave: Educación, Inteligencia Artificial, Inteligencia Artificial en Educación, Agencia Colectiva, Sentido de Agencia, Formación Docente, Investigación de Diseño Educativo, *Futures Studies*

1. Introducción

1.1. La Inteligencia Artificial como Sistema Sociotécnico

Esta tesis parte del reconocimiento de que la Inteligencia Artificial (IA) no puede ser entendida exclusivamente como un avance tecnológico, sino como un sistema sociotécnico profundamente enraizado en la cultura, la imaginación colectiva y las estructuras sociales. Se destacan los imaginarios como elementos centrales en la configuración de la IA: representaciones culturales que moldean no solo el diseño tecnológico, sino también la inversión pública, las políticas y los beneficiarios del desarrollo tecnológico.

La noción de inteligencia en IA se distancia de una simple imitación de la cognición humana. Según autores como Pasquinelli (2023) y Malabou (2021), la inteligencia se concibe como un fenómeno plástico y colectivo, influido por contextos sociopolíticos, históricos y tecnológicos. En este sentido, la IA no “muestra” la inteligencia humana, sino que cristaliza formas colectivas de inteligencia y su desarrollo está siempre mediado por relaciones de poder.

1.2. La Inteligencia Artificial en la Educación

La IA en la educación (IAED) ha experimentado un crecimiento significativo desde la década de 1970, con una aceleración reciente en áreas como tutorías inteligentes, asistentes virtuales, gestión escolar y aprendizaje personalizado. Estas tecnologías prometen reducir la carga docente, identificar habilidades específicas y potencial a los estudiantes, mejorar el acceso a una educación de calidad y expandir el acceso a recursos educativos personalizados, especialmente en contextos vulnerables (Luckin, 2016; Miao, 2021; UNESCO, 2022). Sin embargo, también plantean desafíos éticos y pedagógicos significativos, especialmente en relación con la autonomía de estudiantes y docentes, la privacidad, y la equidad.

1.3. Desafíos del Uso de la IA en Entornos Educativos

Utilizando el marco de Biesta (2013) sobre la finalidad educativa (calificación, socialización y subjetivación), se identifican riesgos asociados con la IAED. Entre ellos, la estandarización del aprendizaje, la reducción de la agencia del estudiante, la opacidad de los sistemas, los sesgos culturales y la ausencia de explicabilidad. La dependencia de sistemas automatizados puede debilitar el juicio moral, reducir la autonomía docente y afectar negativamente tanto al desarrollo ético como al sentido de pertenencia y al universo compartido de significación. Asimismo, se cuestiona una visión meramente legalista de la privacidad, que descuida dimensiones existenciales como el anonimato, la intimidad y la posibilidad de no ser observado, aspectos que la escuela también debería proteger.

1.4. La Ética de la IAED en la Política Educativa y la Formación Docente

Aunque existen marcos éticos generales para la IA (Dignum, 2021), su aplicación específica al ámbito educativo sigue siendo limitada y fragmentaria (Holmes et al., 2019; Hrastinski et al., 2019). Las iniciativas de formación para docentes han tendido a centrarse en aspectos técnicos, dejando de lado la dimensión humana, pedagógica y andragógica. A partir de 2022, con la masificación de la Inteligencia Artificial Generativa (GenIA), surgieron preocupaciones crecientes respecto a la necesidad urgente de formar al profesorado en competencias vinculadas a la alfabetización en IA (e.g., AI4T, 2024). Diversos investigadores han destacado la necesidad urgente de una reforma curricular que fomente el pensamiento crítico y capacite a las comunidades educativas para aprovechar las oportunidades que ofrecen estas tecnologías, al mismo tiempo que aborden sus posibles riesgos (García-Peñalvo, 2023).

1.5. El Giro Conceptual: de la Ética a la Agencia

La tesis propone un desplazamiento conceptual desde la noción de "ética" entendida como marco normativo, hacia la de "agencia" como condición previa o requisito fundamental para la posibilidad misma de la ética. Este cambio responde tanto a la evolución de las tecnologías, especialmente la aparición de la GenIA como actor autónomo, como a una reflexión crítica sobre las limitaciones de los marcos éticos existentes. Se identifican limitaciones principales en la forma en que la ética ha sido operacionalizada o aplicada en el ámbito educativo:

- Ambigüedad epistemológica con listas de principios sin base teórica clara.
- Fragmentación conceptual derivada de competencias definidas de forma imprecisa y escasamente articuladas entre sí.
- Desconexión con la implementación, al mantenerse al margen del contexto educativo, sin traducirse en criterios operativos útiles para docentes y directivos en su labor cotidiana; y sin atender al papel de la agencia docente y estudiantil en el debate sobre IA y educación, priorizando enfoques prescriptivos en lugar de formativos.

Ante estos límites, la tesis defiende que sólo una ética dialógica puede fundamentar el diálogo agentivo en la escuela. A diferencia de las éticas normativas unidireccionales, la ética dialógica permite construir normas a través del discurso participado, fundamentado en consensos racionales, justificaciones procesuales y apertura al disenso. Sin embargo, para que esta ética pueda concretarse en el ámbito educativo, es necesario contar con un corpus conceptual operativo que pueda traducirse en una competencia básica para la formación del

juicio ético y la acción situada. En ese marco, la tesis recupera y amplía el concepto de agencia (Archer, 2014, 2015; Bandura, 2001), proponiendo una estructura triádica, resultado de la investigación desarrollada a lo largo de la tesis, que redefine la formación docente desde una perspectiva ética:

- Agencia subjetiva, entendida como la capacidad individual para la intencionalidad, la planificación, la autorregulación y la reflexividad.
- Agencia intersubjetiva, entendida como la construcción de sentido a través del diálogo, la pertenencia y la interacción.
- Agencia colectiva, que implica el poder de la reflexividad compartida para negociar y transformar estructuras institucionales y sociotécnicas.

Este enfoque sostiene que fortalecer la agencia no es solo deseable pedagógicamente, sino necesario éticamente. Solo quienes poseen agencia pueden actuar de manera ética: deliberar, integrar, resistir y transformar. Por tanto, desarrollar la agencia es una condición para que cualquier marco ético sea realmente significativo y aplicable en la educación mediada por IA. En este contexto, también se invita a repensar sobre la tecnología no como una herramienta neutra, sino como una fuerza autónoma relativa. Ya no está determinada únicamente por antiguos programas científicos, políticos o metafísicos. La tecnología contemporánea, especialmente la IA, actúa a través de una creatividad social sectorial, generando conexiones entre dominios diversos y provocando desplazamientos simbólicos y funcionales que ya no responden a valores heredados (Latour, 2005). Su capacidad transformadora no solo da cuenta de su no neutralidad, sino que también revela su potencia para dinamizar cuerpos, relaciones y subjetividades. En este nuevo paisaje, los seres humanos dejan de ser portadores pasivos de fuerzas para convertirse ellos mismos en fuerzas activas.

Por ello, priorizar una ética dialógica centrada en la agencia permite situar la IA no como una amenaza que deba ser contenida, ni como una solución técnica inevitable, sino como un fenómeno que requiere ser experimentado críticamente, desde el compromiso colectivo, la responsabilidad compartida y una imaginación emancipadora. Solo a través de estos lentes significativos será posible que la IA en educación tenga un sentido pedagógico, humano y transformador.

1.6. Diseño de la Investigación

La investigación adopta un enfoque metodológico basado en el *Educational Design Research* (EDR), que combina investigación con diseño iterativo de soluciones educativas.

Este enfoque se considera especialmente adecuado para contextos de cambio rápido y alta incertidumbre (Kelly, 2006), como el de la IAED. Se sostiene que comenzar con métodos experimentales en torno a aplicaciones formativas que utilizan IA sería prematuro e incluso éticamente problemático, dado que muchas de sus implicaciones pedagógicas, sociales y afectivas aún no se comprenden del todo.

El EDR permite una indagación ética performativa, en la que se explora no solo “qué funciona”, sino también “qué importa” y “para quién”. La investigación se desarrolla mediante ciclos de diseño, implementación, evaluación y rediseño con la participación iterativa de actores educativos.

1.7. Preguntas de Investigación

A partir del análisis conceptual y de los desafíos pedagógicos identificados, la tesis se articula en torno a tres preguntas de investigación:

- ¿Cuáles son las consideraciones éticas asociadas con la integración de sistemas de IA en contextos educativos?
- ¿Cómo puede co-diseñarse un marco conceptual sobre IA, ética y agencia en educación que oriente prácticas pedagógicas centradas en la agencia humana?
- ¿Cómo puede diseñarse e implementarse un curso de formación docente sobre ética e IA, utilizando un enfoque de *Educational Design Research*?

1.8. Desarrollo de la Investigación

La investigación se desarrolló en tres fases interrelacionadas que articularon la exploración teórica, el trabajo empírico y el diseño formativo como parte de un mismo proceso investigativo (ver Figura 1). En una primera etapa, se llevó a cabo una Revisión Sistemática de Literatura (RSL), 2011–2022, guiada por los criterios PRISMA, con el objetivo de mapear el estado del arte en torno a la ética de la IAED.

En una segunda fase, se profundizó en la dimensión empírica mediante cuatro estudios complementarios. Primero, un estudio Delphi con expertos permitió construir ocho escenarios hipotéticos que no solo anticipan posibles impactos de la IA, sino que también abren espacios para imaginar futuros educativos alternativos, inclusivos y éticamente plurales. Luego, a través de grupos focales con formadores de docentes, se examinaron las tensiones éticas, simbólicas y profesionales emergentes en torno a la IA, especialmente en relación con la noción de agencia. A partir de estos hallazgos, se delinearón principios orientadores para el diseño de

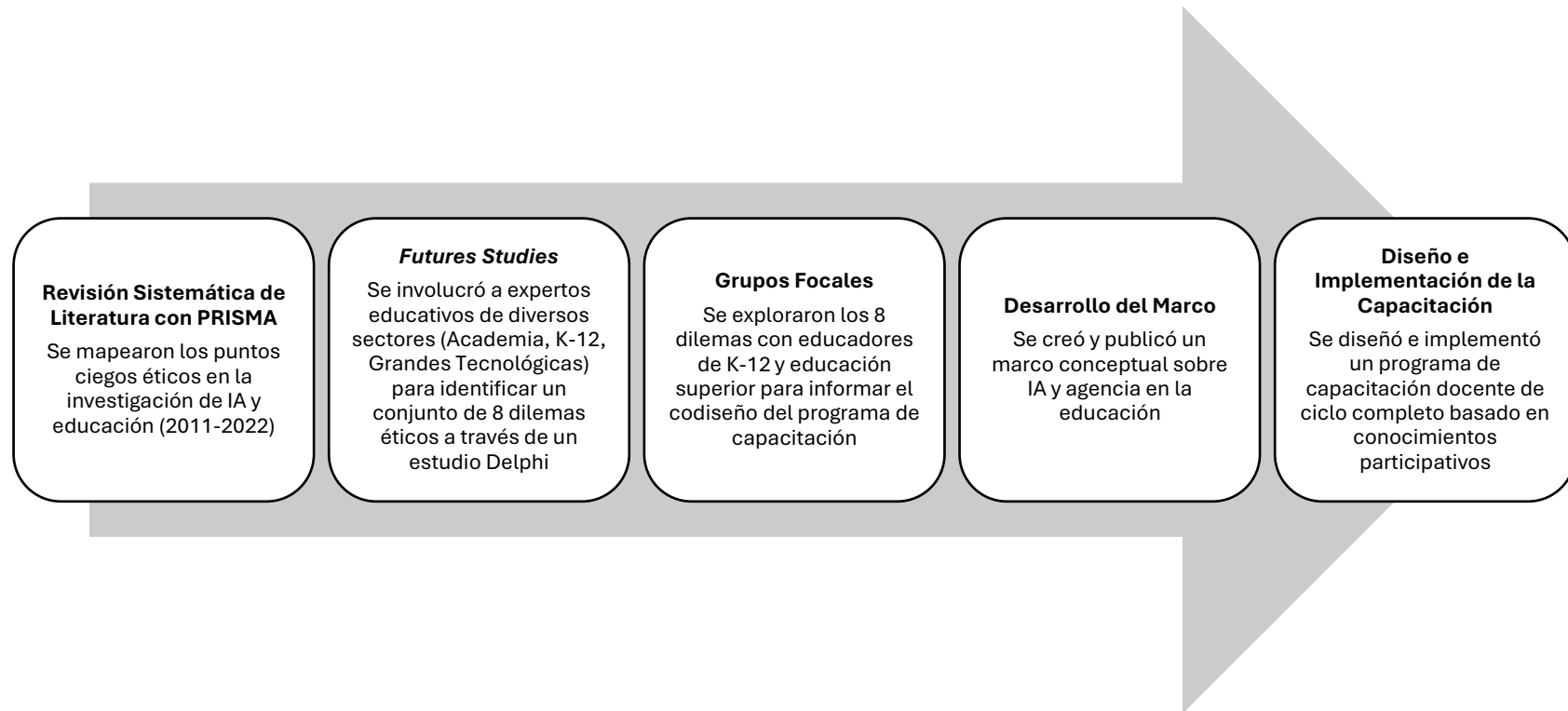
DE LA ÉTICA A LA AGENCIA: DISEÑO PARTICIPATIVO DE UN CURSO DE FORMACIÓN DOCENTE
SOBRE IA EN LA EDUCACIÓN

intervenciones formativas sensibles a las dimensiones subjetiva, intersubjetiva y colectiva de la agencia educativa.

En la tercera fase, se diseñó e implementó el curso de formación docente *De la automatización a la autonomía: agencia educativa con IA*, concebido como respuesta concreta a los desafíos identificados.

Figura 1

Panorama de la Trayectoria de Investigación EDR



2. Estado del Arte: Revisión Sistemática de Literatura sobre Ética e IA en la Educación

La RSL realizada explora las dimensiones éticas presentes en la investigación en torno de las tecnologías de IAED entre 2011 y 2022. Proporciona un análisis exhaustivo de diversas aplicaciones y objetivos con el fin de identificar deficiencias en la literatura existente y ofrecer perspectivas para repensar la ética de la educación en la era de la IA. La revisión también examina cómo se han abordado las diferencias culturales, la inclusión y las emociones junto con los esfuerzos de desarrollo de capacidades, las directrices y los marcos para el uso ético de la IAED. El estudio destaca el papel fundamental que desempeñan los educadores para garantizar el uso ético de la IA en la educación y busca estimular debates éticos que reconozcan la IA como una herramienta no neutral, fortaleciendo las discusiones más amplias sobre la ética de la educación.

2.1. Metodología

La RSL siguió un enfoque estructurado, respondiendo a criterios de elegibilidad predefinidos.

2.1.1. Fase de Planificación

El objetivo del estudio fue comprender cómo se ha abordado la ética en los estudios sobre IAED. Para ello, se utilizó el método PICOC, estructurando la revisión en torno a estudios de IAED y ética, con un enfoque en modelos conceptuales, marcos y programas en contextos educativos diversos. Se plantearon cuatro preguntas de mapeo centradas en la evolución de la producción científica desde 2011, autores clave, países más productivos y áreas temáticas predominantes. Además, se formularon cinco preguntas de investigación para un análisis más profundo: tecnologías y desafíos generales (PI1), inclusión y diversidad cultural (PI2), monitoreo emocional (PI3), desarrollo de capacidades éticas (PI4), y principios y marcos regulatorios (PI5). La estrategia de búsqueda incluyó palabras clave amplias aplicadas en Scopus y Web of Science, considerando publicaciones entre 2011 y 2022 en inglés, francés, portugués y español.

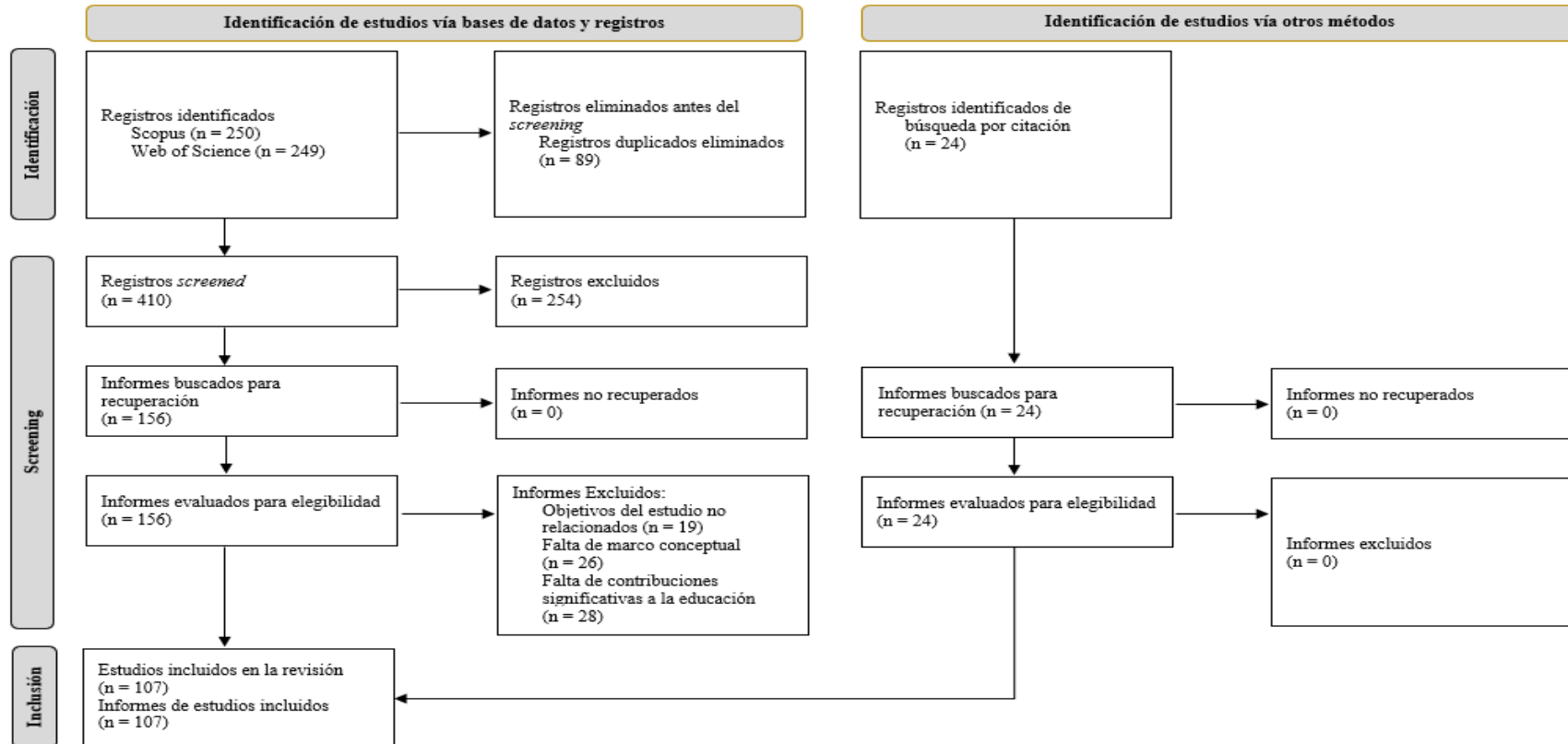
2.1.2. Fase de Realización y Análisis

Como se ilustra en la Figura 2, se recuperaron inicialmente 250 publicaciones de Scopus y 249 de Web of Science. Tras la eliminación de duplicados, se filtraron 410 artículos a partir del título y el resumen. Un segundo filtrado preliminar redujo el número a 156 contribuciones, a las que se aplicó una lista de verificación de evaluación de calidad basada en cinco criterios, lo que permitió incluir 83 artículos. A estos se sumaron 24 artículos adicionales seleccionados como primeras o segundas referencias relevantes para las preguntas de investigación,

especialmente aquellos que ofrecían una comprensión más profunda o ejemplos complementarios. Así, se conformó una muestra final de 107 artículos.

Figura 2

Diagrama de Flujo PRISMA 2020 para el Proceso de RSL



Nota. Adaptado de “The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews”, por M. J. Page, J. E. McKenzie, P. M. Bossuyt, et al., 2021, BMJ, 372, n71 (<https://doi.org/10.1136/bmj.n71>). Copyright 2021 por BMJ Publishing Group.

2.2. Resultados

2.2.1. Análisis de Mapeo

Los hallazgos del análisis de mapeo muestran que, entre 2011 y 2018, las publicaciones sobre ética de la IA en la educación fueron inestables, reflejando una etapa fundacional, mientras que a partir de 2019 se observa un crecimiento exponencial, con un pico en 2021. W. Holmes fue el autor más prolífico, seguido por I. Tuomi, y otros como V. Dignum o R. Luckin con dos publicaciones cada uno. El Reino Unido (47 %) y Finlandia (20 %) lideran la producción, seguidos por España, Suecia y Australia. Aunque EE.UU. domina en publicaciones generales sobre IAED, la discusión ética está encabezada por Europa. Las áreas más representadas fueron informática, ingeniería, ciencias sociales y educación (80 %), lo que evidencia un enfoque técnico que requiere ampliarse hacia disciplinas humanísticas para abordar adecuadamente los desafíos éticos de la IAED.

2.2.2. Análisis de PIs

Los estudios revisados muestran que el 36 % se enfocan en tecnologías de IAED y sus riesgos éticos, abordando aplicaciones desde la Educación Primaria hasta la Educación Superior, con propósitos como el aprendizaje personalizado, la metacognición o la inclusión, pero con desafíos como la vigilancia emocional, la pérdida de agencia y la privacidad. Un 17 % de los artículos abordan cuestiones culturales y de inclusión, revelando sesgos algorítmicos, baja adaptabilidad a contextos no occidentales y dificultades de accesibilidad. Solo el 4 % se centra en el monitoreo emocional, advirtiendo sobre la manipulación afectiva, la vigilancia emocional y un diseño que ignora emociones complejas. En cuanto a la formación en ética, el 41 % de los estudios se refiere a este tema, pero predominan en carreras técnicas, con escasa presencia en la Educación Básica y falta de evaluación del impacto pedagógico. Finalmente, el 8 % analiza marcos y principios, destacando instrumentos como el *Beijing Consensus on AI in Education* (UNESCO, 2019) y *The Ethical Framework for AI in Education* (The Institute for Ethical AI in Education, 2021), subrayando la necesidad de enfoques contextualizados, terminología clara y perspectivas anticipatorias.

2.3. Discusión y Conclusiones

La introducción de nuevas tecnologías como la IAED ofrece una oportunidad para reconsiderar no solo sus consecuencias éticas específicas, sino también la ética más amplia de la educación en sí misma. Se identifican varios puntos ciegos significativos en la literatura sobre ética e IAED. En primer lugar, existe una escasa incorporación de marcos filosóficos o psicológicos del desarrollo moral, lo que limita la capacidad de evaluar críticamente los impactos de estas tecnologías. Asimismo, se omite casi por completo la discusión sobre el

transhumanismo y las implicaciones de la IA en las habilidades cognitivas o la transformación de relaciones en entornos híbridos. Son escasas las perspectivas críticas que analizan los límites éticos de la concepción de inteligencia implícita en los diseños de la IAED, predominando una visión reduccionista centrada en enfoques conductistas y cognitivistas, que dejan de lado las dimensiones afectivas, relacionales y sociales. Además, el énfasis en la optimización puede erosionar el valor del pensamiento reflexivo necesario para un desarrollo humano integral. También se observa una insuficiente atención a los problemas de equidad relacionados con el acceso desigual a los beneficios de la IA. Por último, solo el 4 % de los estudios revisados aborda los dilemas éticos vinculados a la gestión emocional realizada por estos sistemas.

La implementación ética de la IA en la educación exige un enfoque integral, participativo y dialógico. Se recomienda involucrar a los docentes mediante grupos de discusión, diseñar programas piloto basados en evidencia y alinear los contenidos y metodologías formativas con principios socioconstructivistas. La formación debe abordar aspectos éticos específicos de la IAED, perspectivas interdisciplinarias, implicaciones para la agencia y el bienestar estudiantil, así como el rol de la IA en pedagogía, evaluación y aprendizaje. Se sugiere emplear dinámicas como el *role-play* con estudiantes y fortalecer prácticas de autoevaluación y coevaluación como ejercicios de autoría. En cuanto al desarrollo de capacidades y brechas políticas, se identifican desafíos significativos: la formación ética en IAED es prácticamente inexistente, limitada a enfoques técnicos y profesionales, sin herramientas para evaluar prácticas pedagógicas con IA. Incluso en contextos donde la IA ya se utiliza en el aula, la formación sigue siendo asistemática, no intencionada y basada en materiales de baja calidad. Aunque existen marcos éticos, muchos presentan vacíos o no son adecuados para el ámbito educativo. Además, los estudiantes suelen recibir una exposición limitada al uso ético de la IA, ya que la enseñanza de la ética se enfoca mayormente en marcos deontológicos o profesionales, dejando de lado enfoques más dialógicos y orientados al cuidado. Por ello, se propone la infusión curricular de estos temas para cerrar esta brecha.

3. Investigación Empírica I: *Futures Studies*

Este capítulo detalla la primera fase de la investigación empírica de la tesis, que investiga los desafíos éticos que plantea la IAED. El estudio empleó el Método Delphi para involucrar a expertos en una consulta estructurada, lo que llevó al desarrollo y refinamiento de ocho escenarios futuros hipotéticos (Hsu & Sandford, 2007; Okoli & Pawlowski, 2004).

3.1. Introducción

El estudio enfatiza la importancia de la participación de los actores sociales en el diseño de proyectos que forman parte de sus vidas. Dada la novedad del tema y la escasa evaluación

de impacto, la implicación activa de los actores educativos resulta crucial. La investigación propone integrar la ética con enfoques participativos, deliberativos y de múltiples partes interesadas (Dignum, 2019; Floridi & Cowls, 2019), bajo la premisa de que el público puede realizar una evaluación éticamente informada sobre las nuevas tecnologías (Jasanoff & Kim, 2015). Este enfoque reconoce el derecho humano a participar en las decisiones públicas que afectan a la vida de las personas, y busca que las perspectivas éticas de los actores educativos configuren escenarios futuros para informar el diseño de un kit de herramientas de desarrollo profesional continuo para educadores.

Ante el carácter incipiente de la teoría sobre ética en la educación con IAED, esta investigación incorpora metodologías de *futures studies* para anticipar, monitorear y abordar los desafíos éticos que estas tecnologías pueden plantear. Se reconoce que las percepciones de los participantes sobre el presente se transforman al explorar distintos tipos de futuros imaginados (Miller & Tuomi, 2022; van der Heijden, 2005).

Técnicas cualitativas como el Método Delphi y la construcción de escenarios resultan clave no para predecir el futuro, sino para ofrecer perspectivas alternativas ante la incertidumbre, involucrando a las personas en una agencia colectiva orientada a pensar y transformar sus entornos. Estas metodologías son especialmente pertinentes para generar ideas en temas inexplorados o controvertidos (Okoli & Pawlowski, 2004; Hsu & Sandford, 2007). El Método Delphi, en particular, es una técnica estructurada de consulta a expertos, basada en cuestionarios iterativos y retroalimentación anónima para alcanzar consensos informados.

3.2. Metodología

El estudio tuvo como objetivo investigar los desafíos éticos de la IAED desde la perspectiva de actores educativos, utilizando el Método Delphi para desarrollar herramientas de formación docente que promuevan sensibilidad ética, pensamiento crítico y agencia profesional. Participaron dieciocho expertos internacionales con amplia trayectoria en EdTech, seleccionados por su experiencia en investigación, implementación, evaluación, desarrollo y formulación de políticas.

El estudio Delphi se llevó a cabo en tres iteraciones sucesivas. En la primera, dieciocho expertos respondieron un cuestionario de ocho ítems sobre la intersección entre la IAED y la ética. En la segunda ronda, 12 participantes (67 %) revisaron los resultados anteriores y priorizaron determinadas categorías. En la tercera iteración, se analizaron comentarios para evaluar y rediseñar ocho escenarios hipotéticos considerando su plausibilidad, coherencia y claridad.

3.3. Resultados

Se identificaron factores clave como tendencias e incertidumbres críticas (económicas, políticas, sociales y tecnológicas), incluyendo la gamificación de la vida, la influencia de la IA en la expresión emocional y el papel creciente de las corporaciones en definir el discurso educativo. Como resultado, se generó una lista final de escenarios futuros que abordan temas como la necesidad de evidencia de impacto, la equidad, la autonomía estudiantil, la privacidad, la transparencia, la participación informada, el diseño responsable de tecnologías y los riesgos asociados a formas automatizadas de evaluación. Estos escenarios están disponibles como recurso formativo para docentes en: https://drive.google.com/file/d/1o6ToayuZ80Knj4R6QsWBbL7NW0D0aAm7/view?usp=share_link.

3.4. Discusión y Conclusiones

La ética de la IAED debe centrarse en la ética de la educación en primer lugar, ya que cada nueva tecnología representa una oportunidad para repensar los fundamentos éticos del acto educativo. Los sistemas de IA ofrecen a las escuelas la posibilidad de definir un terreno ético común. La mayoría de las preocupaciones éticas identificadas se concentraron en la dimensión relacional, subrayando la relevancia de preservar la interacción humana en contextos mediados por IAED. Los escenarios elaborados en este estudio enfatizan valores compartidos, la participación activa de los actores educativos y una conciencia crítica sobre las disparidades en el acceso a la tecnología y la calidad pedagógica. Se destaca la necesidad de contar con evidencia sólida que respalde los beneficios de la IA en el aprendizaje, así como el riesgo de una creciente brecha entre escuelas públicas y privadas, entre regiones geográficas y entre grupos socioeconómicos.

Asimismo, se resalta que los estudiantes y otros actores del ámbito educativo deben comprender las dependencias sociotécnicas que atraviesan estas tecnologías y sus implicaciones a fin de participar de manera informada en la formulación de políticas escolares sobre IA. El estudio también advierte sobre los riesgos de una adopción acrítica de estas tecnologías, aludiendo a la noción de “vigilancia” propuesta por Zuboff (2019) como uno de los peligros de la automatización educativa. La educación se concibe como una práctica ética que debe generar espacios para la interacción humana empática. Sin embargo, los sistemas de IA pueden debilitar la autonomía estudiantil, restringir oportunidades de interacción significativa y reducir el currículo a contenidos fácilmente automatizables, afectando el desarrollo de la inventiva, la autoeficacia y la autorregulación. Además, se señala el riesgo de

una “desprofesionalización moral” del profesorado, cuando las decisiones pedagógicas se delegan a sistemas automatizados sin la debida reflexión crítica.

En este contexto, se denuncia la mercantilización de la educación, donde la lógica del servicio puede imponerse a la protección de la privacidad. En la era del dataísmo, la preocupación ética gira en torno a la posibilidad de transformar a las personas en entidades completamente medibles y gestionables mediante datos (Williamson, 2022). Esto exige fomentar espacios colaborativos entre docentes y estudiantes para imaginar y construir futuros educativos deseables, así como para analizar críticamente lo que se ha denominado la "gramática onto-epistémica del dataísmo".

Aunque se reconocen los beneficios potenciales de la IAED, como la evaluación en tiempo real y nuevas formas de seguimiento del aprendizaje, su uso debe apoyarse en acuerdos explícitos. El estudio destaca la importancia de formar al profesorado como guía reflexiva, capaz de equilibrar su rol pedagógico con la promoción de la agencia estudiantil y un compromiso crítico e inventivo con el mundo. La IA, al actuar como agente interactivo, puede generar tensiones en la distribución de agencia entre estudiantes, docentes e instituciones. Si bien puede ampliar la autonomía, también existe el riesgo de que limite opciones educativas y fomente decisiones automatizadas o culturas de vigilancia que restrinjan el desarrollo individual y colectivo.

4. Investigación Empírica II: Colectivos Participativos

4.1. Introducción

El Desarrollo Profesional Continuo (DPC) es ampliamente reconocido por su papel en el fortalecimiento de las prácticas pedagógicas y en la mejora del rendimiento estudiantil (Abakah et al., 2022; Darling-Hammond et al., 2017; UNESCO, 2015). Sin embargo, existen críticas persistentes sobre su efectividad limitada, particularmente cuando se basa en modelos tradicionales de carácter uniforme o jerárquico (OECD, 2021; Holmes et al., 2022). Estas críticas destacan la falta de adaptación a las necesidades diversas del profesorado, lo que genera propuestas de formación poco relevantes.

En el contexto de la IAED, se plantea la necesidad de un enfoque transformador del DPC, centrado en los desafíos éticos que estas tecnologías suponen. Esto implica formar a los docentes para que puedan participar activamente en la toma de decisiones, fomentar una cultura de agencia pedagógica y asegurar una implementación significativa y ética de la IAED.

Por ello, la formación docente en ética e IA no debe limitarse a la adquisición técnica, sino orientarse al desarrollo de competencias críticas que permitan a los educadores comprender los algoritmos, sus dependencias sociotécnicas y cuestionar la supuesta neutralidad tecnológica.

En esta línea, Fawns (2022) propone la idea de una "pedagogía entrelazada", que refuerza la necesidad de integrar la dimensión ética en la práctica pedagógica, al concebir la tecnología como parte inseparable de un sistema educativo más amplio, donde lo pedagógico, lo ético, el contenido y los intereses sociales se interrelacionan de manera constante. Esta perspectiva fortalece el rol del profesorado como agente activo en la toma de decisiones sobre cómo, para quién y con qué valores se implementan las tecnologías en la educación.

4.2. Metodología

Se trabajó con cuatro grupos integrados por 19 participantes de cinco países, seleccionados por muestreo por conveniencia y bola de nieve. Todos eran profesionales involucrados en la formación docente, con experiencia en educación y alta competencia en español.

Cada participante recibió dos escenarios 48 horas antes de la sesión. Las reuniones fueron virtuales, grabadas con consentimiento informado y transcritas usando una herramienta automática (Trint), luego validadas manualmente. El guion de moderación incluyó preguntas abiertas sobre IA, actitudes, ética, agencia y diseño de programas formativos.

El análisis de los datos fue temático, utilizando un enfoque mixto: deductivo (con un manual de codificación preliminar) e inductivo (para capturar temas emergentes). La codificación se realizó en NVivo. Se aplicaron dos pruebas de confiabilidad: test-retest (con una correlación de Pearson de 0,86) y análisis intercodificador (acuerdo del 81%, Kappa de 0,83). Se construyó una matriz estructural con categorías como actitudes hacia la IA, oportunidades, desafíos éticos y formación docente. Se analizaron únicamente los fragmentos pertinentes a la pregunta de investigación relativa a los programas de formación ética en IA dirigidos a docentes de Educación Primaria y Secundaria.

Esta metodología permitió asegurar rigurosidad, diversidad de perspectivas y una comprensión profundizada sobre cómo estructurar programas formativos en ética e IAED, centrados en la agencia del profesorado.

4.3. Resultados

Las discusiones de los grupos focales ofrecieron orientaciones sobre cómo estructurar y desarrollar programas de formación en ética de la IAED. Los participantes coincidieron en que la prioridad principal debe ser fomentar un pensamiento ético ampliado, no como un tema aislado, sino como base del proceso formativo. Esta perspectiva ética se vinculó estrechamente con la necesidad de actualizar conocimientos y habilidades, reconociendo la naturaleza cambiante de la IA y la importancia de que los docentes comprendan sus aspectos técnicos para poder evaluarla críticamente.

A medida que los docentes se enfrentan a los desafíos éticos y técnicos de la IAED, también se ven impulsados a cuestionar los paradigmas educativos existentes y a redefinir sus roles profesionales. La integración de estas tecnologías no debe reproducir modelos pedagógicos obsoletos, sino fomentar el cuestionamiento de las prácticas de enseñanza y aprendizaje. En este sentido, los docentes fueron vistos como agentes activos que deben tomar decisiones informadas y significativas sobre el uso de la IA. También se mencionó la necesidad de respetar marcos normativos aunque este aspecto fue menos destacado que las preocupaciones éticas y pedagógicas.

En cuanto al diseño de los programas, se destacó la importancia de anticipar posibles resistencias y facilitar transiciones graduales. Para ello, se propusieron formaciones prácticas basadas en experiencias reales, incluyendo dilemas éticos, estudios de caso y ejercicios de simulación. Además, se valoró la creación de entornos colaborativos de aprendizaje, donde docentes y familias dialoguen sobre los desafíos que plantea la IA, así como el fomento de espacios de desarrollo profesional continuo, tanto formal como autónomo.

Algunos participantes expresaron preocupaciones sobre la pertinencia y el momento adecuado para implementar esta formación, especialmente cuando existen otras prioridades formativas urgentes. Esto subrayó la importancia de establecer prioridades claras y adaptar los programas a cada contexto educativo. El análisis reveló que la formación ética en IAED debe ser un componente central del desarrollo profesional docente con capacidad para transformar no solo lo que los docentes saben o cómo actúan, sino también cómo piensan, interpretan el presente y construyen el futuro de la educación.

4.4. Discusión y Conclusiones

Las discusiones de los grupos focales revelaron una comprensión integral de cómo debe organizarse un programa de formación ética en IAED. Los participantes concibieron la ética no como un módulo aislado, sino como el eje transversal de toda la propuesta formativa. Esta visión estructural vincula directamente el pensamiento ético con la necesidad de actualizar conocimientos técnicos y cuestionar las bases del sistema educativo. Asimismo, los formadores subrayaron que el programa debe responder a un enfoque contextualizado, flexible y situado, capaz de adaptarse a las realidades de los centros educativos, los niveles de acceso a tecnología y las trayectorias del profesorado.

Dentro de estas dimensiones estructurales, se enfatizó también la interdependencia entre ética, conocimiento técnico, pedagogía crítica y agencia docente. La formación ética debe ayudar a los docentes a evaluar no solo cómo se usan las herramientas de IA, sino también con qué fines, en qué contextos y con qué impactos posibles sobre la equidad, la privacidad o la

autonomía del estudiante. Por lo tanto, las dimensiones estructurantes del programa no se limitan al contenido, sino que incluyen principios orientadores, objetivos pedagógicos y criterios éticos para el diseño y evaluación.

Desde una perspectiva de diseño, se destacó la importancia de prever posibles resistencias al cambio y planificar transiciones graduales hacia la integración significativa de la IA. Los participantes identificaron que estas resistencias pueden originarse tanto en factores emocionales como en la falta de familiaridad con el tema, por lo que recomendaron incorporar actividades que conecten con las experiencias previas de los docentes. El diseño del programa debe incluir tiempo para la reflexión, fomentar la colaboración entre pares y promover el diálogo con actores clave como las familias.

Además, los formadores alertaron sobre la necesidad de establecer prioridades realistas: la formación en ética de la IA no debe imponerse sin considerar otras necesidades formativas urgentes, como la inclusión. Por tanto, un diseño sensible debe evaluar el momento adecuado para su implementación y garantizar que haya condiciones mínimas, como el acceso a tecnologías relevantes, que permitan que la formación tenga sentido y favorezca una reflexión situada a partir del *tinkering*. También se sugirió que estas temáticas se introduzcan de forma progresiva en la formación inicial docente.

En cuanto a las estrategias para implementar el programa, los participantes subrayaron la necesidad de enfoques prácticos, experienciales y participativos. Propusieron incorporar estudios de caso, dilemas éticos reales, simulaciones y ejercicios de reflexión colectiva que permitan a los docentes vivenciar la toma de decisiones éticas y desarrollar su juicio profesional.

Se valoró especialmente el aprendizaje entre pares y el desarrollo profesional informal como componentes claves. La formación no debe limitarse a cursos teóricos, sino ofrecer espacios continuos de intercambio y actualización permanente. La combinación de métodos prácticos y reflexión crítica facilita que los docentes pasen de ser consumidores de tecnología a convertirse en co-creadores de usos pedagógicos éticamente informados. En este sentido, la implementación debe estar alineada con una pedagogía que priorice el sentido, la agencia y la autonomía.

El estudio también abordó críticamente la "racionalidad de la técnica": la tendencia a priorizar la eficiencia y la automatización sobre dimensiones pedagógicas y éticas. Se señala que los programas de DPC deben instigar a los docentes para desafiar esta lógica, fomentando una postura crítica y reflexiva hacia la tecnología. Esto implica no solo saber usar herramientas de IA, sino tener el criterio profesional para decidir cuándo y por qué usarlas, o incluso cuándo

no usarlas. En este contexto, la agencia pedagógica se convierte en una respuesta crítica frente al riesgo de desprofesionalización y de adopción pasiva de soluciones tecnológicas.

Entre las limitaciones del estudio se encuentran el número de participantes y su enfoque en los niveles educativos de Educación Primaria y Secundaria. Futuros estudios podrían examinar la implementación concreta y los efectos de programas de formación basados en estos hallazgos, así como adaptar este enfoque a otros contextos educativos.

5. Investigación Empírica III: La Naturaleza Cambiante de la Agencia en la Educación con IA

5.1. Introducción

Este capítulo presenta un análisis secundario de los datos de grupos focales previos, con el objetivo de comprender de manera más profunda cómo se conceptualiza y discute la agencia por parte de los educadores en el contexto de la integración de la IAED. Los hallazgos de las fases de investigación anteriores, que involucraron a formadores de docentes en consideraciones éticas y estrategias pedagógicas para la IAED, identificaron áreas clave para abordar en los programas de formación docente. Este análisis refinado de la agencia es clave para orientar el diseño del programa de formación docente al poner en el centro una ética que se expresa a través de la agencia en contextos educativos.

El capítulo profundiza en cómo las tecnologías de IA están afectando la agencia de los docentes, examinando tanto las oportunidades como los desafíos que surgen. Destaca que la IA puede apoyar la agencia de los docentes al automatizar tareas rutinarias, liberar tiempo para el trabajo pedagógico significativo y ofrecer nuevas herramientas para la adaptación y la evaluación. Sin embargo, también subraya los riesgos de que la IA socave la agencia al imponer metodologías estandarizadas, limitar la toma de decisiones pedagógicas y promover una cultura de dependencia tecnológica.

En un contexto de creciente automatización y dependencia de sistemas tecnológicos, la noción de agencia está siendo redefinida. Las interacciones con sistemas sociotécnicos, especialmente aquellos basados en IA, plantean nuevos retos y posibilidades tanto para la agencia individual como colectiva. Más recientemente, el debate sobre la agencia se ha ampliado para considerar su carácter distribuido, cuestionando la idea de que se trata de una propiedad exclusivamente humana o individual. Desde enfoques como los de Deleuze y Guattari (1980), Haraway (1985, 2016) y Latour (2005), se ha propuesto que la agencia emerge en redes complejas donde interactúan entidades humanas y no humanas, como algoritmos, máquinas, tecnologías, objetos, animales o incluso factores ambientales. Esta perspectiva invita a repensar el lugar del docente en relación con la tecnología no como usuario pasivo o simple

operador, sino como parte de un entramado de relaciones dinámicas que configuran, condicionan y también abren posibilidades de acción.

5.2. Metodología

Este estudio adoptó un enfoque cualitativo, utilizando el análisis secundario de los datos de los grupos focales, con el objetivo de comprender cómo los docentes conceptualizan la transformación de la agencia humana en relación con la IAED, basándose en el marco teórico y metodológico de la ética dialógica. Se realizó un análisis secundario de los datos generados en los grupos focales descritos en el capítulo anterior, reorientando el foco hacia la noción de agencia, dado que muchos fragmentos del discurso, aunque no se centraban explícitamente en este concepto, abordaban temas relacionados con los sistemas sociotécnicos y la educación.

Participaron 19 formadores de docentes de cinco países hispanohablantes, seleccionados por muestreo por conveniencia y bola de nieve, con experiencia en la formación inicial y continua del profesorado y una sólida trayectoria en la integración de tecnologías en la educación. Las sesiones se desarrollaron de forma virtual, fueron grabadas, transcritas y posteriormente verificadas por dos investigadores. Para el análisis, se utilizó una estrategia combinada deductiva e inductiva, que incluyó búsquedas específicas de términos relacionados con la agencia (como autonomía, control, toma de decisiones, autodeterminación), lo que permitió identificar cómo los docentes articulan esta noción en el contexto de la IA.

La codificación final fue validada mediante pruebas de confiabilidad, incluyendo un test-retest y un análisis intercodificador, asegurando la consistencia en la interpretación de los datos y permitiendo una comprensión profunda del fenómeno estudiado.

5.3. Resultados

Los resultados de este análisis secundario, presentados en la Tabla 1, profundizaron la comprensión de la compleja relación entre la IA y la agencia en el ámbito educativo. En el plano de la agencia individual, los participantes identificaron el potencial que ofrece la IA. Para los estudiantes, se destacó que estas tecnologías pueden facilitar una mayor adecuación de los ejercicios a los estudiantes, permitiéndoles avanzar a su propio ritmo, elegir estilos de estudio que se ajusten a sus necesidades y participar más activamente en su proceso formativo. En el caso del profesorado, se valoró la posibilidad de delegar tareas rutinarias, como la corrección automática o la gestión de datos administrativos, lo cual liberaría tiempo para centrarse en aspectos más complejos y significativos de la práctica docente, como la interacción personalizada con el alumnado o la creación de propuestas pedagógicas innovadoras. No obstante, también se expresaron preocupaciones sobre los riesgos que estas mismas tecnologías pueden suponer para la agencia individual. Se mencionó, por ejemplo, el peligro de una

desprofesionalización docente si las decisiones pedagógicas clave fueran determinadas por algoritmos, lo cual podría erosionar su autonomía y juicio profesional. En el caso del estudiantado, se advirtió que un uso excesivo o mal diseñado de la IA podría restringir su capacidad para explorar, equivocarse y construir una trayectoria significativa, especialmente si los sistemas solo ofrecen contenidos filtrados por lógicas algorítmicas cerradas.

Desde una perspectiva intersubjetiva, los docentes destacaron el impacto de la IA en las relaciones pedagógicas y los propósitos compartidos en el entorno educativo. Se identificó una tensión creciente en torno al papel del profesorado, cuya función puede verse diluida si la IA asume decisiones fundamentales o redefine las metas del aprendizaje sin un debate colectivo. La erosión del sentido de propósito y la posible reducción del rol docente a una mera supervisión técnica fueron vistas como amenazas al vínculo pedagógico y al compromiso ético con los estudiantes. Al mismo tiempo, se subrayó la necesidad de mantener formas de mediación humana que permitan interpretar los datos y las recomendaciones generadas por los sistemas inteligentes, reafirmando la importancia de una autoridad pedagógica situada, capaz de sostener procesos educativos significativos en diálogo con los valores compartidos de la comunidad escolar.

A nivel de agencia colectiva, la IA fue vista como un recurso que podría optimizar la toma de decisiones institucionales y facilitar la colaboración entre docentes. Se valoró su capacidad para generar análisis de datos útiles para identificar necesidades comunes y diseñar respuestas educativas más informadas y contextualizadas. Además, se destacó el potencial de estas tecnologías para fortalecer redes profesionales mediante el intercambio automatizado y ajustado de recursos, materiales y buenas prácticas. Sin embargo, emergieron también importantes desafíos vinculados a la centralización del poder y la pérdida de control local. Los participantes manifestaron preocupación por el hecho de que muchas soluciones de IA provienen de grandes corporaciones tecnológicas que pueden imponer agendas educativas estandarizadas, a menudo desconectadas de las realidades culturales y pedagógicas específicas de cada comunidad. Esta tensión evidenció la necesidad de mantener la capacidad de las comunidades escolares para definir sus propios fines educativos. En este marco, se destacó una forma particular de agencia colectiva expresada a través de la resistencia. Cuando los docentes perciben que las tecnologías amenazan sus valores o su autonomía profesional, pueden organizarse para cuestionarlas, adaptarlas o incluso rechazar su implementación. Lejos de representar una actitud cautelosa, esta resistencia se interpretó como una estrategia activa para transformar las tecnologías en función de los propósitos pedagógicos y éticos propios, abriendo

así posibilidades para una implementación más justa y significativa de la IA en contextos educativos diversos.

Tabla 1

Diagrama Temático de los Niveles Subjetivo, Intersubjetivo y Colectivo de la Agencia en Educación

NIVEL	DIMENSIÓN	DESCRIPCIÓN
SUBJETIVO	PROCESOS DE TOMA DE DECISIONES	<p>Razonamiento Crítico Las decisiones impulsadas por la IA pueden socavar la capacidad de los estudiantes para razonar sobre variables personalmente significativas.</p> <p>Poder de la Estructura Influencia de factores sociales, culturales y ambientales que dan forma y pueden restringir las elecciones individuales en función de los recursos disponibles y las oportunidades visibles, lo que subraya la necesidad de evaluar críticamente si la tecnología y la IA pueden expandir o restringir estas opciones.</p> <p>Diversidad y Circunscripción de las Elecciones Los sistemas de IA que se basan en rutas de aprendizaje predefinidas pueden restringir la capacidad de los estudiantes para explorar una amplia gama de posibilidades.</p>
	INDIVIDUACIÓN	<p>Diferenciación y Autoría Se espera que los sistemas de IA apoyen a la comunidad educativa explorando las diversas trayectorias de desarrollo y las expresiones personales únicas de cada estudiante (por ejemplo, ritmo de aprendizaje, repertorio conductual, manifestaciones específicas) en lugar de imponer medidas estandarizadas.</p>
	PREVISIÓN, AUTORREGULACIÓN Y AUTORREFLEXIÓN	<p>Propiedades Agénticas Cualidades que permiten a los individuos apropiarse de sus vidas y decisiones, permitiéndoles dar forma activamente a sus experiencias y alcanzar sus metas.</p>
	DESARROLLO MORAL	<p>Desconexión Moral El desapego progresivo de la responsabilidad y la erosión de los estándares éticos ocurren cuando los individuos justifican o racionalizan acciones, a menudo atribuyendo la toma de decisiones a fuerzas externas, como la tecnología o las figuras de autoridad. Esto puede disminuir la responsabilidad personal y el razonamiento moral.</p>
INTERSUBJETIVO	PROPÓSITO Y VALORES	<p>Roles de los Maestros Erosión de los roles y el propósito de los maestros, a medida que la IA comienza a manejar tareas pedagógicas centrales, lo que podría convertir su rol en uno de mera supervisión.</p>
	AUTORIDAD Y ESTATUS PROFESIONAL	<p>Agencia Proxy Los sistemas de IA se consideran herramientas de apoyo, con la expectativa de que los maestros mantengan su rol como agentes proxy, preservando su autonomía y autoridad profesional.</p>
	MODELADO Y CONFIANZA	<p>Andamiaje Proceso de apoyo a través del cual los maestros aprovechan sus propias experiencias y habilidades socioemocionales para guiar a los estudiantes en el desarrollo de competencias críticas.</p>
COLECTIVO		Participación

ESPACIO POLÍTICO	Importancia de un enfoque participativo donde todas las partes interesadas (por ejemplo, estudiantes, maestros, padres) participan en la configuración del papel de la IA en la educación.
	Toma de Decisiones Distribuida
	Llamado a la toma de decisiones colaborativa en la integración de la IA, con los maestros abogando por un papel en la configuración de su implementación, destacando la necesidad de prácticas democráticas a nivel institucional.
	Democracia y Estado de Derecho
	Importancia de la participación, la agencia y la responsabilidad, arraigadas en las necesidades de la comunidad. Promueve una sociedad guiada por marcos institucionales establecidos que protegen las libertades individuales y mejoran el bienestar social.
AGENCIA DISTRIBUIDA	Dinámica de Ensamblajes Sociotécnicos
	Agencia compartida entre actores humanos y tecnológicos crean un contexto único para el aprendizaje y refuerzan la idea de que la experiencia educativa se co-construye, moldeada por la interacción de diversas influencias.

5.4. Discusión y Conclusiones

Una de las preocupaciones más relevantes emergentes del estudio es la desprofesionalización de los docentes ante la automatización creciente de funciones tradicionalmente ligadas al juicio humano, la sensibilidad pedagógica y la toma de decisiones situada. Cuando los sistemas de IA asumen funciones evaluativas, de planificación o de personalización del aprendizaje, se corre el riesgo de que los docentes se transformen en meros ejecutores de instrucciones algorítmicas. Esto reduce su papel a operarios de sistemas cerrados, desplazándolos de su rol como diseñadores, mediadores críticos y facilitadores activos del aprendizaje.

Frente a este escenario, preservar y fortalecer la agencia docente implica más que adquirir competencias técnicas. Requiere el desarrollo de una racionalidad de la técnica, es decir, una comprensión profunda de los fines, límites y supuestos que sustentan el diseño y funcionamiento de los sistemas de IA. Mantener la agencia significa no solo saber usar una herramienta tecnológica, sino tener la autonomía para modificarla, rechazarla o adaptarla a las necesidades pedagógicas y contextuales. En este sentido, los docentes deben ser sujetos activos en la concretización de la tecnología, participando en su definición y aplicación concreta (Simondon, 2008).

En este marco, los programas de desarrollo profesional continuo deben superar el enfoque instrumental. Es crucial que capaciten a los docentes para evaluar críticamente los sistemas de IA, identificar sesgos, comprender implicaciones éticas, y participar en la formulación de políticas y directrices institucionales. Crear comunidades de práctica donde se aborden dilemas reales, se compartan experiencias y se construyan soluciones colectivas es esencial para promover una agencia situada, reflexiva y colaborativa.

La discusión también subraya el peso de los factores culturales, sociales e institucionales en la configuración de la agencia. En muchos contextos educativos, las normas tradicionales y las estructuras jerárquicas limitan la autonomía profesional, dificultando la posibilidad de desafiar o transformar los marcos existentes. Esto puede obstaculizar la capacidad de los docentes para leer críticamente su ecosistema de aprendizaje, adaptarse a nuevas exigencias y ejercer discernimiento ético ante los usos de la IA en educación. En este sentido, se destaca el valor de una ética dialógica, que favorezca el debate abierto, la deliberación colectiva y la construcción de significados compartidos sobre el papel de la IA en la educación.

Desde un punto de vista más estructural, la IA plantea interrogantes profundos sobre los procesos de toma de decisiones, desarrollo moral, subjetividad y trayectoria profesional en contextos educativos. La personalización algorítmica puede alterar la experiencia educativa

común, diluyendo los espacios de sincronía emocional y logro colectivo que constituyen el núcleo formativo de la escuela. A medida que los sistemas de IA reconfiguran relaciones, jerarquías y criterios de evaluación, se reescriben también las condiciones para la participación democrática y la formación ciudadana.

Este estudio ha permitido esbozar un primer modelo para pensar la agencia desde una perspectiva estratificada, subjetiva, intersubjetiva y colectiva, articulada al entramado tecnológico y educativo. Este modelo propone que futuras investigaciones examinen cómo distintos tipos de sistemas de IA (por ejemplo, herramientas co-creativas versus sistemas prescriptivos) impactan de manera diferencial en la agencia de los actores educativos. También sugiere la urgencia de diseñar marcos regulatorios participativos, que garanticen una integración de la IA centrada en la agencia y la justicia educativa.

6. Diseño y Evaluación Preliminar del Curso de DPC para Docentes

6.1. Introducción: Situando el Curso en la Trayectoria de Investigación Doctoral

La creación del curso de DPC presentado en este capítulo no constituye una iniciativa aislada, sino que emerge como culminación de un proceso de investigación largo, estructurado y coherente, enmarcado en una metodología de EDR iterativa. Esta metodología ha permitido una integración rigurosa entre la producción de conocimiento teórico, la recogida de evidencias empíricas y la elaboración de una intervención pedagógica situada. Cada fase previa del proyecto aportó insumos conceptuales y prácticos fundamentales para la construcción del curso, que se configura tanto como una aplicación concreta de los hallazgos obtenidos como un espacio experimental desde el cual poner a prueba los marcos teóricos desarrollados y generar nuevo conocimiento situado sobre la formación docente en tiempos de IA.

No se trata de un curso técnico centrado en el uso instrumental de herramientas, sino de una propuesta que sitúa la reflexión ética, la agencia pedagógica y la co-creación de sentido en el centro del proceso formativo. El objetivo general es que los participantes desarrollen su agencia colectiva para analizar críticamente los efectos de la IA, identificar dilemas éticos complejos, tomar decisiones pedagógicas fundadas y diseñar estrategias formativas éticamente sólidas.

La noción de agencia pedagógica articula la arquitectura conceptual del curso. Este enfoque reconoce que solo los sujetos dotados de agencia son capaces de actuar éticamente, y que fortalecer dicha agencia es una condición necesaria para la autonomía profesional en entornos altamente tecnologizados.

Pedagógicamente, el curso se apoya en un enfoque constructivista, participativo y dialógico. Promueve el aprendizaje activo mediante el uso de problemas reales, fomenta la

deliberación ética en grupos heterogéneos, y sitúa todas las actividades dentro de los marcos contextuales reales de los docentes participantes. A través del Aprendizaje Basado en Problemas, se exploran situaciones ambiguas que permiten generar discusión, tensionar principios éticos y diseñar respuestas creativas desde una perspectiva crítica. La pedagogía dialógica que atraviesa el curso reconoce que no hay respuestas únicas a los dilemas morales que plantea la IA, y que el juicio ético se construye colectivamente. A su vez, la dimensión situada del curso garantiza que los aprendizajes no sean abstractos, sino directamente aplicables a la práctica profesional.

El curso está estructurado en cinco módulos interconectados, que reflejan los hallazgos de las etapas anteriores y han sido diseñados para construir de manera progresiva tanto la comprensión crítica como la capacidad de acción ética de los participantes.

6.2. Método: Desarrollo Detallado de los Módulos del Curso

A través de una estructura modular, el curso busca dotar a los participantes de capacidades para analizar, diseñar e implementar prácticas educativas que refuercen la agencia subjetiva, intersubjetiva y colectiva. El diseño del curso responde a una planificación estratégica que integra objetivos temáticos, metodologías activas, recursos tecnológicos y actividades centradas en la práctica real del profesorado. El enfoque se aleja del tecnocentrismo y se basa en principios de aprendizaje dialógico, exploración reflexiva y apropiación crítica de las tecnologías. Desde el inicio, las sesiones están diseñadas para minimizar posibles resistencias, comenzando con marcos narrativos amplios que sitúan la discusión antes de introducir específicamente los temas relacionados con la IA.

El curso se apoya en diversas herramientas de IA potencialmente adaptables al entorno educativo, no como fines en sí mismos, sino como medios para cuestionar, problematizar y rediseñar las prácticas pedagógicas. Su uso está integrado en experiencias situadas que fomentan la autoría, el pensamiento crítico y el ejercicio de una agencia distribuida.

La arquitectura conceptual del curso se organiza en torno a cuatro módulos, cada uno centrado en una dimensión específica de la agencia educativa, que se presentan a continuación.

• Módulo 1: Agencia Colectiva

Este módulo explora la posición del profesorado dentro del panorama sociotécnico configurado por la IA. Se analiza cómo los imaginarios educativos y las narrativas tecnológicas no solo moldean la organización escolar y las experiencias educativas, sino también la propia capacidad de imaginar alternativas y futuros distintos. Las actividades están orientadas a desarrollar una conciencia crítica sobre el papel colectivo de docentes y estudiantes en la configuración de futuros educativos posibles. En este marco, se

abordan conceptos clave como los ensamblajes sociotécnicos y los fines de la educación, entendidos como calificación, socialización y subjetivación.

• **Módulo 2: Agencia Subjetiva**

Este módulo aborda cómo los sistemas de IA influyen en la percepción que los estudiantes tienen sobre su autonomía, intencionalidad y autorregulación. A través de dilemas éticos situados, los docentes analizan cómo algunas decisiones pedagógicas, mediadas por la tecnología, afectan la capacidad de los estudiantes para actuar con sentido, y cómo estas decisiones pueden conducir a una progresiva desvinculación entre las tareas cognitivas y afectivas y sus correlatos morales, expresivos y creativos. Se promueven prácticas que fortalezcan la construcción del sentido y la participación activa del alumnado.

• **Módulo 3: Agencia Intersubjetiva**

En esta unidad, se privilegia la dimensión relacional del aprendizaje, analizando cómo las interacciones humanas se ven afectadas por la automatización y la hiperpersonalización. Las actividades invitan a reflexionar sobre el valor pedagógico de la incertidumbre, el diálogo abierto y la construcción compartida de significados. Se exploran las consecuencias que los entornos algorítmicos tienen sobre la pertenencia, la reciprocidad y la socialización educativa.

• **Módulo 4: Agencia Distribuida**

En la última sesión, se trabaja la noción de agencia como una práctica compartida entre actores humanos y no humanos. Se promueve el diseño y experimentación de experiencias educativas donde la autoría y la responsabilidad ética sean distribuidas entre docentes, estudiantes, familias, tecnologías e instituciones. El curso culmina con una reflexión colectiva sobre las prácticas implementadas y la creación de una comunidad de práctica orientada a sostener el aprendizaje continuo y el compromiso ético más allá del curso, donde la agencia distribuida se comprende como parte integral de la agencia colectiva.

6.3. Evaluación y Resultados

Durante la fase de evaluación del curso, participaron como evaluadores dos profesores universitarios con experiencia en la formación docente, quienes ya habían intervenido en la fase Delphi del proyecto. Esta continuidad permitió mantener una mirada longitudinal y crítica sobre el proceso de investigación y sobre el diseño del curso.

Su conocimiento previo de los dilemas éticos explorados en la investigación les permitió evaluar si el curso respondía de forma coherente a las problemáticas detectadas inicialmente. A pesar de haber participado en fases anteriores del proyecto, los evaluadores no estuvieron involucrados en el desarrollo final del curso, lo que garantizó una evaluación informada, pero no condicionada, permitiendo así un distanciamiento crítico con conocimiento de contexto.

La fase de evaluación del curso combinó dos estrategias complementarias y entrelazadas: el *walkthrough* formativo y la microevaluación piloto. Esta articulación metodológica respondió al carácter iterativo y situado del enfoque EDR, permitiendo observar tanto el curso en acción como su recepción inicial en contextos reales de formación.

El *walkthrough* se desarrolló en formato mixto, simulando el ritmo real del curso. El primer participante completó las sesiones iniciales de forma presencial, mientras que el resto del recorrido se llevó a cabo a través de intercambios asincrónicos. Sus comentarios fueron incorporados de forma inmediata e iterativa al proceso de rediseño. El segundo participante, por su parte, accedió a la versión revisada del curso, facilitada por el primer evaluador, y ofreció sus observaciones mediante formularios reflexivos, centrado más en su aplicabilidad desde la mirada del docente destinatario. La recolección de datos combinó procedimientos *think-aloud*, observación directa, discusiones facilitadas y diarios de evaluación estructurados bajo cinco dimensiones: primeras impresiones, experiencia de usuario, relevancia pedagógica, impacto potencial en estudiantes y adecuación temporal. Los hallazgos revelaron que el curso mantenía una sólida coherencia pedagógica, con una progresión conceptual clara en torno a las distintas capas de agencia. No obstante, se identificaron algunas desconexiones entre las actividades y sus respectivas discusiones, así como la necesidad de incluir más ejemplos contextualizados y un mayor soporte teórico para aquellas actividades de carácter más inductivo. Estas observaciones fueron abordadas y mejoradas de forma inmediata. Se destacaron como especialmente relevantes las actividades que integraban el uso de herramientas de IA con instancias de reflexión ética.

Posteriormente, la fase de microevaluación se llevó a cabo con diez docentes de Educación Primaria y Secundaria, durante un congreso de desarrollo profesional. La sesión piloto se organizó en dos bloques de dos horas cada uno, sumando un total de cuatro horas, y se centró en cinco actividades clave del curso. A través de un cuestionario mixto, se evaluaron tres dimensiones: coherencia conceptual, potencial formativo y pertinencia profesional. El tema más reconocido fue “Agencia humana e IA: subjetiva, relacional, colectiva”, considerado tanto central como desafiante, lo que validó su relevancia y necesidad de mayor acompañamiento. Las actividades más valoradas fueron aquellas vinculadas con IA generativa, especialmente las

que abordaban autoría y diseño de *prompts*, por su clara conexión con la práctica docente. En términos cualitativos, los participantes calificaron el curso como urgente e indispensable, destacando su aporte al pensamiento crítico, la conciencia ética y la integración reflexiva de tecnologías. Se subrayó además la importancia de preservar la memoria y la reflexión humana frente a la automatización, consolidando la dimensión simbólica y relacional del curso como espacio de construcción de significados. En el plano tecnológico, la experimentación con modelos de lenguaje de gran escala fue valorada por facilitar un *tinkering* pedagógico situado. En el plano de contenidos, conceptos como agencia y subjetivación se comprendieron mejor mediante actividades colaborativas y contextualizadas. En lo relacional, se resaltó la necesidad de mantener vínculos humanos significativos en contextos mediados por IA, reconociendo la diferencia, el testimonio y la memoria como claves éticas. Finalmente, en el dominio procesual, el curso logró activar transiciones afectivas y cognitivas significativas mediante experiencias simbólicas corporales, como el ejercicio con globos y pelotas, que facilitaron una apropiación vivencial del concepto de agencia.

En conjunto, la articulación entre *walkthroughs*, microevaluación y rediseño iterativo no solo reforzó la solidez del curso, sino que confirmó su coherencia con los fundamentos de la investigación, permitiendo que la intervención avance hacia una implementación ampliada, respaldada por evidencias cualitativas de su relevancia formativa.

6.4. Discusión y Conclusiones

Este capítulo representa la culminación del proyecto de *Educational Design Research*, cuyo propósito central ha sido abordar la interacción entre la IA en la educación, la ética y la agencia humana, de un modo que informase el desarrollo de un curso destinado a docentes de Educación Primaria y Secundaria. La metodología de EDR, adoptada de McKenney y Reeves (2018), se despliega a través de tres fases iterativas e interconectadas: análisis y exploración, diseño y construcción, y evaluación y reflexión. Desde su concepción, este proyecto investigador fue estructurado para evolucionar iterativamente desde una exploración teórica rigurosa hacia una aplicación práctica sólidamente fundamentada, asegurando que cada hallazgo informara el siguiente paso.

El enfoque ético, crítico y participativo del curso lo convierte en un prototipo adaptable y escalable, capaz de replicarse en diversos contextos institucionales sin perder coherencia. Su valor radica en que promueve una formación situada, éticamente fundamentada y orientada a la acción, que prepara a los docentes no solo para comprender la IA, sino para intervenir activamente en sus propias prácticas y en los procesos de agencia colectiva en torno a la IA dentro de sus escuelas. Al abordar directamente la amenaza de desprofesionalización, el curso

devuelve al profesorado su papel central como actor reflexivo, responsable y creativo. Al fomentar una racionalidad técnica crítica y al fortalecer la agencia individual y colectiva, se constituye en una herramienta para resignificar el lugar de la tecnología en la educación.

7. Conclusiones Finales: La IA como Desafío Pedagógico y Ético

La presente tesis doctoral aborda la integración de la IA en los ecosistemas educativos como un fenómeno complejo, marcado por ambigüedades éticas, tensiones pedagógicas y la necesidad urgente de repensar colectivamente los fines de la educación. Frente a una narrativa dominante que promueve la adopción técnica y escalable de la IA, esta investigación propone una mirada crítica y centrada en la agencia educativa, entendida como una capacidad distribuida y relacional que debe ser protegida y cultivada en contextos atravesados por tecnologías sociotécnicas. La investigación se desarrolló a través del enfoque de *Educational Design Research*, combinando indagación conceptual, co-diseño participativo y desarrollo práctico de un curso de formación docente para la Educación Primaria y Secundaria, enfocado en ética, IA y agencia.

La primera pregunta de investigación se centró en identificar las consideraciones éticas que surgen al integrar sistemas de IA en contextos educativos. Para ello, se realizó una RSL entre 2011 y 2022, que reveló vacíos significativos en los marcos éticos existentes, tales como ambigüedad epistemológica, fragmentación conceptual, falta de orientación contextual, escasa formación ética docente y la importación acrítica de modelos legales o tecnológicos al ámbito pedagógico. A este diagnóstico se sumó un estudio Delphi con expertos en educación, que permitió construir escenarios futuros y un conjunto de dilemas éticos diseñados para fomentar la reflexión crítica y participativa sobre los impactos de la IA en la educación. Estos estudios demostraron que las cuestiones éticas ligadas a la IA reflejan problemas más amplios sobre el sentido mismo de educar.

La segunda pregunta de investigación buscó co-diseñar un marco conceptual que articulara IA, ética y agencia educativa. Mediante grupos focales con formadores docentes, se exploraron las transformaciones de la agencia provocadas por la interacción con determinadas tecnologías de IA en el campo educativo. El análisis de estas conversaciones permitió identificar un marco de agencia educativa compuesto por tres capas interdependientes: la subjetiva, vinculada con la autoría individual y el juicio ético; la intersubjetiva, relacionada con el diálogo, el reconocimiento mutuo y las dinámicas de aula; y la colectiva, referida a decisiones institucionales, cultura democrática y co-construcción del sentido educativo. La dimensión de la agencia distribuida es intencionalmente capturada dentro de la capa de la agencia colectiva, entendida no como una mera suma de voluntades individuales, sino como una ecología

relacional donde emergen formas de acción compartida, mediadas por tecnologías, instituciones y vínculos simbólicos. Esta elección responde a la necesidad de destacar que, en contextos educativos mediados por IA, la agencia no se disuelve ni se fragmenta en nodos técnicos, sino que se reconfigura en función de las condiciones materiales, afectivas y organizacionales que permiten, o limitan, la toma de decisiones situada y significativa. Incluir la agencia distribuida dentro de la capa colectiva permite abordar con mayor profundidad cómo la acción educativa se entreteje en redes humanas y no humanas, y cómo los docentes, estudiantes, plataformas y algoritmos participan en procesos de co-construcción del sentido pedagógico. Así, la agencia colectiva incluye tanto los modos más tradicionales de participación democrática como las formas más contemporáneas de interdependencia sociotécnica, en las que el poder de actuar se distribuye, pero no necesariamente se equilibra. Esta decisión metodológica y conceptual busca resistir la tentación de crear una nueva entidad ontológica derivada de la distribución de agencias entre humanos y no humanos. En su lugar, propone mantener el anclaje de la agencia en prácticas humanas situadas, subrayando que, aunque las tecnologías participan en la configuración de los entornos educativos, no reemplazan la capacidad humana de interpretar, deliberar y sostener el juicio ético en contextos complejos.

El estudio mostró que los docentes suelen focalizarse en la agencia individual del estudiante, pero tienden a descuidar las dimensiones más amplias de la agencia relacional y colectiva, lo cual plantea riesgos en contextos donde la IA tiende a desplazar funciones profesionales, erosionar el juicio pedagógico y promover imaginarios cerrados sobre el futuro educativo.

La tercera pregunta se orientó al diseño e implementación de un curso de formación profesional para docentes, guiado por el enfoque EDR. El curso, titulado *De la automatización a la autonomía: agencia educativa con IA*, fue desarrollado en colaboración con educadores y articulado en torno al marco de agencia. La arquitectura del curso evitó el uso de sistemas tradicionales de gestión online del aprendizaje, como los LMS convencionales, que suelen privilegiar trayectorias individualizadas, automatizadas y centradas en el rendimiento. En su lugar, se optó por una plataforma más flexible y abierta (Canva). Esta elección no fue meramente técnica, sino profundamente ética y política: buscó resistir el encuadre tecnocrático que convierte la formación docente en un proceso estandarizado y lineal para, en cambio, afirmar una visión del aprendizaje como práctica relacional, colaborativa y reflexiva. Se incorporaron herramientas como Synthesia, Character.AI, Mentimeter y ChatGPT, no como soluciones tecnológicas, sino como medios para explorar (*tinkering*) las implicancias de la IA

en la práctica educativa. A través de sesiones progresivas, actividades experienciales y evaluaciones cualitativas basadas en reflexión personal, retroalimentación entre pares y guía docente, el curso pretende promover una comprensión profunda y situada de la ética y la agencia en tiempos de automatización. Su implementación en formato piloto con un grupo reducido de docentes ofreció resultados prometedores, aunque limitados en alcance y duración.

En síntesis, esta investigación ofrece cuatro contribuciones principales:

1. Un mapeo del campo de la IA en la educación y la ética, basado en la literatura y actualizado frente al auge de la IA generativa.
2. Un conjunto de dilemas éticos, diseñados participativamente como herramienta para el diálogo en entornos escolares y en instancias de toma de decisiones educativas e institucionales sobre IA en educación.
3. Un marco tridimensional de agencia educativa, que articula niveles subjetivos, intersubjetivos y colectivos (y distribuidos) de acción pedagógica.
4. Un curso de formación docente de ocho sesiones, contextualizado y centrado en la agencia, acompañado de una plataforma digital y un *eBook* con materiales didácticos y guías detalladas (disponible en: <https://anamouta.my.canva.site/from-automation-to-autonomy-educational-agency-with-ai>).

La tesis propone reconquistar el espacio simbólico de la educación: recuperar su lexicón ético y su capacidad de hospedar lo inesperado. Ello implica defender espacios educativos donde el anonimato protegido, el disenso, la demora y la formación ética tengan lugar, resistiendo la presión por la visibilidad constante y la acción automática.

La investigación reconoce sus limitaciones: una evaluación del curso en contexto reducido y con participantes autoseleccionados, la imposibilidad de medir impactos a largo plazo y la necesidad de avanzar hacia una diferenciación más fina entre tipos específicos de sistemas de IA, atendiendo a cómo sus configuraciones técnicas y sus modos de integración en la práctica educativa afectan de forma diferenciada las dimensiones subjetiva, intersubjetiva y colectiva de la agencia. A pesar de estos límites, la tesis abre múltiples caminos futuros. Entre ellos, estudios longitudinales sobre el impacto sostenido en la práctica docente, implementación en contextos más diversos, desarrollo de estrategias de escalabilidad, investigación sobre la perspectiva estudiantil y una actualización constante del curso conforme evolucionan las tecnologías y los desafíos educativos.

Este trabajo representa una contribución teórica, metodológica y práctica a los debates contemporáneos sobre IA y educación. Invita a docentes, investigadores y responsables de política educativa a no centrarse únicamente en “cómo integrar la IA”, sino a preguntarse para

qué, en qué condiciones y en nombre de qué fines educativos se hace. Solo así será posible construir colectivamente ecologías escolares donde la agencia colectiva, el juicio pedagógico y el deseo de imaginar futuros educativos prosperen.

Referencias

- Abakah, E., Widin, J., & Ameyaw, E. K. (2022). Continuing professional development (CPD) practices among basic school teachers in the Central Region of Ghana. *SAGE Open*, 12(2). <https://doi.org/10.1177/21582440221094597>
- AI4T. (2024a). *Artificial intelligence for and by teachers*. <https://www.ai4t.eu/>
- Archer, M. S. (2014). *Structure, agency and the internal conversation*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139087315>
- Archer, M. S. (2015). How agency is transformed in the course of social transformation: Don't forget the double morphogenesis. In M. S. Archer (Ed.), *Generative mechanisms transforming late modernity* (pp. 135–157). Springer. https://doi.org/10.1007/978-3-319-10788-6_6
- Bandura, A. (2001). Social cognitive theory: An agentic perspective. *Annual Review of Psychology*, 52, 1–26. <https://doi.org/10.1146/annurev.psych.52.1.1>
- Biesta, G. (2013). *Beautiful Risk of Education*. Routledge. <https://doi.org/10.4324/9781315635866>
- Darling-Hammond, L., Hyler, M. E., & Gardner, M. (2017). *Effective teacher professional development*. Learning Policy Institute.
- Deleuze, G., & Guattari, F. (1980). *A thousand plateaus: Capitalism and schizophrenia*. University of Minnesota Press.
- Dignum, V. (2019). *Responsible Artificial Intelligence: How to Develop and Use AI in a Responsible Way*. Springer. <https://doi.org/10.1007/978-3-030-30371-6>
- Dignum, V. (2021). The role and challenges of education for responsible AI. *London Review of Education*, 19(1), Article 1. <https://doi.org/10.14324/LRE.19.1.01>
- Fawns, T. (2022). An entangled pedagogy: Looking beyond the pedagogy–technology dichotomy. *Postdigital Science and Education*, 4, 711–728. <https://doi.org/10.1007/s42438-022-00302-7>
- Floridi, L., & Cowls, J. (2019). A unified framework of five principles for AI in society. *Harvard Data Science Review*, 1(1). <https://doi.org/10.1162/99608f92.8cd550d1>
- García-Peñalvo, F. J., & Vázquez-Ingelmo, A. (2023). What do we mean by GenAI? A systematic mapping of the evolution, trends, and techniques involved in generative AI. *International Journal of Interactive Multimedia and Artificial Intelligence*, 8(4), 7-16. <https://doi.org/10.9781/ijimai.2023.07.006>

- Haraway, D. (1985). A cyborg manifesto: Science, technology, and socialist-feminism in the late twentieth century. In *Simians, cyborgs, and women: The reinvention of nature* (pp. 149–181). Routledge.
- Haraway, D. J. (2016). *Staying with the trouble: Making kin in the Chthulucene*. Duke University Press. <https://doi.org/10.1515/9780822373780>
- Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.
- Holmes, W., Persson, J., Chounta, I.-A., Wasson, B., & Dimitrova, V. (2022). *Artificial intelligence and education: A critical view through the lens of human rights, democracy and the rule of law*. Council of Europe Publishing. <https://rm.coe.int/artificial-intelligence-and-education-a-critical-view-through-the-lens/1680a886bd>
- Hrastinski, S., Olofsson, A. D., Arkenback, C., Ekström, S., Ericsson, E., Fransson, G., Jaldemark, J., Ryberg, T., Öberg, L.-M., Fuentes, A., Gustafsson, U., Humble, N., Mozelius, P., Sundgren, M., & Utterberg, M. (2019). Critical imaginaries and reflections on artificial intelligence and robots in postdigital K-12 education. *Postdigital Science and Education*, 1(2), 427–445. <https://doi.org/10.1007/s42438-019-00046-x>
- Hsu, C.-C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research & Evaluation*, 12(10), 1–8. <https://doi.org/10.7275/pdz9-th90>
- Jasanoff, S., & Kim, S.-H. (Eds.). (2015). *Dreamscapes of modernity: Socio-technical imaginaries and the fabrication of power*. University of Chicago Press.
- Kelly, A. E. (2006). Quality criteria for design research: Evidence and commitments. In J. van den Akker, K. Gravemeijer, S. McKenney & N. Nieveen (Eds.), *Educational design research* (pp. 107–118). Routledge.
- Latour, B. (2005). *Reassembling the social: An introduction to actor-network-theory*. Oxford University Press.
- Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.
- Malabou, C. (2021). *Plasticity and intelligence: Between philosophy and the neurosciences*. Polity Press.
- McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.). Routledge.

- Miao, F., Holmes, W., Huang, R., & Zhang, H. (2021). *AI and education: Guidance for policy-makers*. United Nations Educational, Scientific and Cultural Organization.
<https://unesdoc.unesco.org/ark:/48223/pf0000376709>
- Miller, R., & Tuomi, I. (2022). Making the futures of AI in education: Why and how imagining the future matters. *European Journal of Education*, 57(4), 537–541.
<https://doi.org/10.1111/ejed.12529>
- OECD. (2021). *OECD Digital Education Outlook 2021: Pushing the frontiers with artificial intelligence, blockchain and robots*. OECD Publishing.
<https://doi.org/10.1787/589b283f-en>
- Okoli, C., & Pawlowski, S. D. (2004). The Delphi method as a research tool: An example, design considerations and applications. *Information & Management*, 42(1), 15–29.
<https://doi.org/10.1016/j.im.2003.11.002>
- Pasquinelli, M. (2023). *The eye of the master: A social history of artificial intelligence*. Verso Books.
- Simondon, G. (2008). *Imagination et invention (1965–1966)*. La Transparence.
- The Institute for Ethical AI in Education. (2021). *The ethical framework for AI in education*.
<https://www.buckingham.ac.uk/wp-content/uploads/2021/03/The-Institute-for-Ethical-AI-in-Education-The-Ethical-Framework-for-AI-in-Education.pdf>
- UNESCO. (2015). *SDG 4 indicators*. <https://www.education-progress.org/en/indicators>
- van der Heijden, K. (2005). *Scenarios: The art of strategic conversation* (2nd ed.). Wiley.
- UNESCO. (2019). *Beijing consensus on artificial intelligence and education*.
<https://unesdoc.unesco.org/ark:/48223/pf0000368303>
- UNESCO. (2022). *K–12 AI curricula: A mapping of government-endorsed AI curricula*.
<https://unesdoc.unesco.org/ark:/48223/pf0000380602>
- Williamson, B. (2022). Education datafication and AI: Governance, ethics, and equity. *Learning, Media and Technology*, 47(1), 1–15.
<https://doi.org/10.1080/17439884.2021.2003404>
- Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. PublicAffairs.