

# **DATA LITERACY PARA EL DESARROLLO DE LEARNING ANALYTICS EN ENTORNOS K-12**

AUTORA:  
BELÉN DONATE BEBY

DIRECTORES:  
FRANCISCO JOSÉ GARCÍA-PEÑALVO  
DANIEL AMO FILVÁ

PLAN DE INVESTIGACIÓN  
PROGRAMA DE DOCTORADO FORMACIÓN EN LA SOCIEDAD DEL  
CONOCIMIENTO  
UNIVERSIDAD DE SALAMANCA

FECHA: 13 DE ENERO DE 2022

## INTRODUCCIÓN Y JUSTIFICACIÓN DEL TEMA OBJETO DE ESTUDIO

La temática elegida se centra en un problema presente en los centros educativos de etapas K-12: la alta digitalización de las escuelas en los últimos años y la falta de competencias digitales en el profesorado. Estos cambios radicales en la digitalización educativa se deben a la pandemia del COVID-19 (García-Peñalvo & Corell, 2020; Gil-Fernández et al., 2021; Knopik & Oszwa, 2021) cuya reacción en las instituciones educativas ha sido introducir nuevas prácticas fuertemente apoyadas por herramientas digitales (García-Peñalvo et al., 2020a, 2020b, 2021) y, especialmente, en ecosistemas tecnológicos (García-Holgado & García-Peñalvo, 2013, 2019; Manikas & Hansen, 2013).

Tal y como se recoge en el Acuerdo de 14 de mayo de 2020, de la Conferencia de Educación sobre el marco de referencia de la competencia digital docente: “La educación y la formación juegan un papel determinante: la competencia digital ha pasado a formar parte indispensable de la alfabetización elemental del siglo XXI” (Resolución de 2 de julio de 2020, de la Dirección General de Evaluación y Cooperación Territorial, por la que se publica el Acuerdo de la Conferencia Sectorial de Educación sobre el marco de referencia de la competencia digital docente, 2020). Sin embargo, de acuerdo con (Domínguez Figaredo et al., 2020), estas nuevas prácticas han revelado una falta de habilidades digitales y de alfabetización de datos en los profesores de los mencionados niveles educativos.

En adición, desde una perspectiva pedagógica, en la práctica docente se requiere abordar distintos ritmos de aprendizaje para proporcionar una instrucción adaptada a cada caso. Es por ello por lo que, en entornos digitalizados, las analíticas del aprendizaje podrían ejercer un papel fundamental (Amo & Santiago, 2017). Las analíticas de aprendizaje son definidas como un campo emergente que tiende puentes ente la informática y las ciencias sociales a través del manejo de datos (Siemens, 2012). De este modo, según la Society for Learning Analytics Research (SoLAR, <https://www.solaresearch.org/>) la analítica del aprendizaje es considerada “la medición, la recopilación, el análisis y la presentación de datos sobre los alumnos y sus contextos, con el fin de comprender y optimizar el aprendizaje y los entornos en los que se produce”. Con ello, constituye un área de referencia en el campo de las Tecnologías del Aprendizaje, ya que, combinando diferentes enfoques técnicos y metodológicos en el manejo de datos, puede facilitar la toma de decisiones en el ámbito educativo (García-Peñalvo, 2020). Sin embargo, el uso de las analíticas de aprendizaje implica consideraciones, como la de asegurar la adecuación a las políticas de privacidad y seguridad en el uso de datos. Tal es su relevancia que (Amo et al., 2019) establecen una definición de las LA que lo resalta: “el uso de datos de alumnos en contextos online para su estudio y análisis con técnicas predictivas en un entorno privado y seguro”.

Esta importancia reside en que existe una gran fragilidad en la privacidad, confidencialidad y seguridad durante el tratamiento de los datos educativos en el desarrollo de las LA, cuya resolución viene de la mano del nivel de sistema gestor de bases de datos (Amo et al., 2020).

En esta línea, según Johnson et al. (2016), estas se ajustan a la vigente evaluación por competencias, dado que permiten la recopilación, análisis y uso de datos sobre los

estudiantes y sus contextos, permitiendo una comprensión mayor que optimiza el aprendizaje y los entornos en los que tiene lugar. No obstante, según señala (Ruipérez-Valiente, 2020) para lograr un uso sostenible de las analíticas de aprendizaje, a largo plazo, es necesaria la formación del profesorado de modo que puedan reajustar sus prácticas educativas.

Por todo ello, el objetivo principal de la presente tesis se centra en conocer cuál es el nivel de alfabetización de datos para el desarrollo de analíticas de aprendizaje en las etapas K-12 en territorio español (Donate et al., 2021). Concretamente, este objetivo se centrará en la perspectiva del profesorado, tratando de promover su autoevaluación en la alfabetización de datos. Por otro lado, cabe resaltar que dicha delimitación geográfica se justifica por las características particulares del sistema educativo español. Además, se considera que la población española posee la suficiente amplitud y alcance para ser abordada desde la presente tesis. Sin embargo, se debe señalar como un aspecto esencial, la continuación de esta tesis a través de trabajos que amplíen la práctica a otros territorios.

## HIPÓTESIS DE TRABAJO Y PRINCIPALES OBJETIVOS A ALCANZAR

### **Objetivo general**

- Conocer e impulsar el estado de la cuestión de la alfabetización de datos en las etapas educativas K-12 de España.

### **Objetivos específicos**

- Desarrollar un instrumento que permita conocer el estado empírico de la cuestión de la alfabetización de datos en las etapas educativas K-12 en territorio español.
- Aplicar el instrumento de autoevaluación del nivel de alfabetización de datos en centros educativos de las etapas K-12 en territorio español.
- Publicar los resultados obtenidos en un observatorio digital de la analítica del aprendizaje inicialmente para etapas K-12 del territorio español.
- Promover la autoidentificación de necesidades formativas para el desarrollo de las analíticas de aprendizaje en estos niveles.
- Generar un itinerario formativo que responda a las necesidades encontradas en el profesorado para la alfabetización de datos

### **Preguntas de investigación**

- ¿Cuál es el estado de la cuestión de las analíticas de aprendizaje en las etapas K-12 de territorio español?
- ¿Qué instrumentos de autoevaluación se utilizan actualmente para la medida del estado de la alfabetización de datos?
- ¿Es posible generar un instrumento para evaluar la alfabetización de datos en entornos K-12?
- ¿Es posible acompañar a los resultados de instrumento de acciones formativas para alcanzar niveles competenciales superiores en alfabetización de datos?

## METODOLOGÍA

Para la consecución de los objetivos y dar respuesta a las preguntas de investigación se propone una metodología de investigación mixta de tres fases:

Una primera fase consistente en ejecutar una Revisión Sistemática de la Literatura del estado del arte bajo una metodología documental descriptiva de enfoque cualitativa-cuantitativa a partir de fuentes de información secundarias. Para situar el estado de la cuestión, se realiza una primera aproximación a través de un Mapeo Sistemático de la Literatura y, en extensión, una Revisión Sistemática de la Literatura. Los objetivos de ambas aproximaciones se centran, respectivamente, en explorar la investigación actual sobre las analíticas de aprendizaje en los entornos K-12 del territorio español. Esto permite obtener una visión global del estado de la cuestión en dichas etapas, desde la consolidación del concepto de las analíticas del aprendizaje, hasta la actualidad. La delimitación geográfica puede suponer una limitación en la transferencia de los resultados. Sin embargo, como extensión de la presente tesis, se plantea continuar el trabajo con otras poblaciones de diferentes países. En última instancia, cabe resaltar que la realización de la revisión del estado del arte estará supeditado a las consideraciones expresadas por (García-Holgado et al., 2020; Kitchenham & Charters, 2007).

Una segunda fase consistente en elaborar un instrumento de (auto)evaluación de la alfabetización de datos bajo una metodología documental con enfoque cualitativo. Para el diseño del instrumento de evaluación, se llevará a cabo la selección de trabajos sobre dicha temática que pudieran constituir su marco de referencia. En esta línea, se ha realizado un primer análisis documental sobre los trabajos y cuestionarios existentes en materia de alfabetización de datos educativos. En ellos se observa que la mayoría se concentran en la perspectiva docente, con contribuciones procedentes de la Educación Superior, mostrando interés por desarrollar un movimiento educativo basado en la evidencia a través de datos cuantitativos (Calvani et al., 2012; Slavin, 2002). Estos autores tratan de evaluar, desde un enfoque ético y crítico, el manejo de datos para la práctica, la información y la analítica del aprendizaje (Raffaghelli, 2018). Como base de desarrollo del cuestionario, se ha elegido la revisión bibliográfica realizada por Maybee & Zilinski (2015), tomándose como referencia las siguientes categorías de competencias docentes para el manejo de datos: conciencia y comprensión de datos, acceso a los mismos, compromiso en la interpretación y toma de decisiones, gestión y planificación, comunicación y uso ético y preservación o almacenamiento.

A su vez, se establecen los conocimientos o actividades que definen dichas competencias, de acuerdo a los marcos de referencia propuestos por Juliana E. Raffaghelli (2018) y Mandinach & Jimerson (2016). Para garantizar la validez de los bancos de ítems construidos, así como el contenido de los mismos (Sireci, 1998), se recurrirá a la consulta de expertos para juzgar de manera más precisa la redacción y planteamiento de los ítems para la población de interés; así como a la revisión exhaustiva de las fuentes bibliográficas. Tal y como señala Muñiz Fernández (2010), los cuestionarios requieren indicadores del grado de precisión con el que miden, por lo que garantizar la validez y fiabilidad de los mismos constituye un aspecto fundamental.

Una tercera fase consistente en la ejecución del instrumento desarrollado en la segunda fase bajo una metodología descriptiva con un enfoque cuantitativo. La muestra poblacional incluirá roles educativos de matiz directivo, docente y discente. La ejecución anónima del instrumento permitirá la recolección de datos y su posterior análisis estadístico mediante programas especializados. Posteriormente, se plantea el lanzamiento

y administración del cuestionario VESB4LA, sobre el consentimiento en el uso, manejo y recopilación de datos personales de los estudiantes de Educación Secundaria y Bachillerato, en Learning Analytics.

Los resultados permitirán elaborar un observatorio de datos en cuanto a la alfabetización y manejo de datos en etapas educativas K-12 dentro del territorio español. La relación de los resultados del análisis con las necesidades formativas permitirá la realización de itinerarios formativos relativos a los ítems del cuestionario, culminando con la posibilidad de desarrollar un estudio piloto en el que se lleve a la práctica, lo aprendido en el manejo de analíticas de aprendizaje.

### Cuestiones éticas

La tesis defiende una investigación con recogida de datos y tratamiento de los mismos. Por ello, con el fin de proteger la identidad del alumnado y profesorado implicado, se evitarán posibles conflictos éticos y legales relacionados con la LOPD/RGPD. Del mismo modo, el desarrollo de la tesis estará supeditada a un código ético de investigación como es el desarrollado por la British Educational Research Association (British Educational Research Association (BERA), 2019).

### MEDIOS Y RECURSOS MATERIALES DISPONIBLES

Este trabajo se desarrolla en el programa de Doctorado: Formación en la Sociedad del Conocimiento (García-Peñalvo, 2014, 2015, 2017, 2021) **14/1/22 10:53:00**, siendo su portal la principal herramienta de comunicación y visibilidad de los avances (García-Holgado et al., , 2015; García-Peñalvo et al., 2019b). Para la compleción de los objetivos marcados en esta tesis doctoral se dispone del soporte:

- Del grupo de investigación GRIAL (García-Peñalvo et al., 2019a) dirigido por el Catedrático, Francisco José García Peñalvo, director principal de esta tesis.
- Del grupo de investigación GRETEL del que el Dr. Daniel Amo forma parte, director de esta tesis.

### PLANIFICACIÓN TEMPORAL

#### Plan Tesis Doctoral

<b>1ª FASE 2021-2022</b>	Inicio interacción	07/2021
	Realización SMR	08/2021
	Diseño Plan de Investigación	09/2021-11/2021
	Realización SLR y análisis de los marcos e instrumentos	01/2022-03/2022
<b>2ª FASE 2022-2023</b>	Desarrollo cuestionario Data Literacy educativo específico para entornos K-12	04/2022 - 05/2022

	Validación cuestionario	06/2022 - 07/2022
	Implementación cuestionario	09/2022 - 02/2023
3ª FASE 2023-2024	Análisis de resultados, diseño de acción formativa	03/2023 – 08/2023
	Implementación de la acción	09/2023 - 12/2023
4ª FASE	Evaluación de la acción, elaboración del documento final	01/2024 - 05/2024

### Plan de publicaciones:

La programación de las publicaciones se ha establecido de modo que figure, al menos, un congreso internacional y una publicación en revista anualmente.

Ponencia Mapeo Sistemático de la Literatura en Conferencia Internacional CV21	09/2021
Ponencia “An instrument for self-assessment of data literacy at the micro, meso, and macro educational levels” en Conferencia HCII	2022
Ponencia SLR en Conferencia Internacional	
Publicación artículo SLR en revista	
Ponencia en Conferencia Internacional: instrumento de evaluación Data Literacy	2023
Publicación en revista: Instrumento evaluación de Data Literacy/ programa formativo	2023

Ponencia en Conferencia Internacional: presentación programa formativo	
Ponencia en Conferencia Internacional: implementación programa formativo	2024
Publicación en revista: implementación del programa formativo	

## REFERENCIAS

- Amo, D., Alier, M., García-Peñalvo, F. J., Fonseca, D., & Casañ, M. J. (2020). Privacidad, seguridad y legalidad en soluciones educativas basadas en Blockchain: Una Revisión Sistemática de la Literatura. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(2), 213-236. <https://doi.org/10.5944/ried.23.2.26388>
- Amo, D., Navarro, J., Casañ, M. J., García-Peñalvo, F. J., Alier, M., & Fonseca, D. (2019). *Evaluación de la importancia de la ética, privacidad y seguridad en los estudios de Learning Analytics, en el marco de las conferencias LAK (COMPON-2019-CINAIC-0073)*. Article COMPON-2019-CINAIC-0073. <https://doi.org/10.26754/CINAIC.2019.0073>
- Amo, D., & Santiago, R. (2017). *Learning Analytics: La narración del aprendizaje a través de los datos*. Editorial UOC.
- British Educational Research Association (BERA). (2019). *Guía ética para la investigación educativa (4ª ed.)* British Educational Research Association (BERA). <https://bit.ly/35ZT8v1>
- Calvani, A., Fini, A., Ranieri, M., & Picci, P. (2012). Are young generations in secondary school digitally competent? A study on Italian teenagers. *Computers & Education*, 58(2), 797-807.
- Domínguez Figaredo, D., Reich, J., & Ruipérez-Valiente, J. A. (2020). Analítica del aprendizaje y educación basada en datos: Un campo en expansión. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(2), 33. <https://doi.org/10.5944/ried.23.2.27105>
- Donate, B., García-Peñalvo, F. J., & Amo, D. (2021). Learning Analytics in K-12 Spanish education: A systematic mapping study. *2021 XI International Conference on Virtual Campus (JICV)*, 1-4. <https://doi.org/10.1109/JICV53222.2021.9600315>
- García-Holgado, A., & García-Peñalvo, F. J. (2013). Análisis de integración de soluciones basadas en software como servicio para la implantación de ecosistemas tecnológicos corporativos. In J. Cruz Benito, A. García Holgado, S. García Sánchez, D. Hernández Alfageme, M. Navarro Cáceres, & R. Vega Ruiz (Eds.), *Avances en Informática y Automática. Séptimo Workshop* (pp. 55-72). Departamento de Informática y Automática de la Universidad de Salamanca.
- García-Holgado, A., & García-Peñalvo, F. J. (2019). Validation of the learning ecosystem metamodel using transformation rules. *Future Generation Computer Systems*, 91, 300-310. <https://doi.org/10.1016/j.future.2018.09.011>
- García-Holgado, A., García-Peñalvo, F. J., & Rodríguez-Conde, M. J. (2015). Definition

- of a technological ecosystem for scientific knowledge management in a PhD Programme. In G. R. Alves & M. C. Felgueiras (Eds.), *Proceedings of the Third International Conference on Technological Ecosystems for Enhancing Multiculturality (TEEM'15) (Porto, Portugal, October 7-9, 2015)* (pp. 695-700). ACM. <https://doi.org/10.1145/2808580.2808686>
- García-Holgado, A., Marcos Pablos, S., & García Peñalvo, F. J. (2020). Guidelines for performing systematic research projects reviews. *International Journal of Interactive Multimedia and Artificial Intelligence*, 6(2), 9.
- García-Peñalvo, F. J. (2014). Formación en la sociedad del conocimiento, un programa de doctorado con una perspectiva interdisciplinar. *Teoría de la Educación. Educación y Cultura en la Sociedad de la Información*, 15(1), 4-9.
- García-Peñalvo, F. J. (2015). Engineering contributions to a Knowledge Society multicultural perspective. *IEEE Revista Iberoamericana de Tecnologías del Aprendizaje (IEEE RITA)*, 10(1), 17-18. <https://doi.org/10.1109/RITA.2015.2391371>
- García-Peñalvo, F. J. (2017). *Education in the Knowledge Society PhD Programme. 2017 Kick-off Meeting* Seminarios del Programa de Doctorado en Formación en la Sociedad del Conocimiento (16 de noviembre de 2017), Salamanca, España. <https://goo.gl/bJ5qKd>
- García-Peñalvo, F. J. (2020). Learning Analytics as a Breakthrough in Educational Improvement. In D. Burgos (Ed.), *Radical Solutions and Learning Analytics: Personalised Learning and Teaching Through Big Data* (pp. 1-15). Springer Singapore. [https://doi.org/10.1007/978-981-15-4526-9\\_1](https://doi.org/10.1007/978-981-15-4526-9_1)
- García-Peñalvo, F. J. (2021). *Sesión de inicio (kick-off meeting) del Programa de Doctorado Formación en la Sociedad del Conocimiento: Curso 2021-2022* Seminarios del Programa de Doctorado en Formación en la Sociedad del Conocimiento (2 de noviembre de 2021), Salamanca, España. <https://bit.ly/3CDtMlr>
- García-Peñalvo, F. J., & Corell, A. (2020). La COVID-19: ¿enzima de la transformación digital de la docencia o reflejo de una crisis metodológica y competencial en la educación superior? *Campus Virtuales*, 9(2), 83-98.
- García-Peñalvo, F. J., Corell, A., Abella-García, V., & Grande, M. (2020a). Online assessment in higher education in the time of COVID-19. *Education in the Knowledge Society*, 21. <https://doi.org/10.14201/eks.23013>
- 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. En D. Burgos, A. Tlili, & A. Tabacco (Eds.), *Radical Solutions for Education in a Crisis Context: COVID-19 as an Opportunity for Global Learning* (pp. 85-98). Springer. [https://doi.org/10.1007/978-981-15-7869-4\\_6](https://doi.org/10.1007/978-981-15-7869-4_6)
- García-Peñalvo, F. J., Rivero-Ortega, R., Rodríguez-Conde, M. J., & Rodríguez-García, N. (2020b). The institutional decisions to support remote learning and teaching during the COVID-19 pandemic. In *Proceedings of the 2020 X International Conference on Virtual Campus (JICV) (Tetouan, Marrooc, 3-5 December 2020)*. IEEE. <https://doi.org/10.1109/JICV51605.2020.9375683>
- García-Peñalvo, F. J., Rodríguez Conde, M. J., Therón Sánchez, R., García Holgado, A., Benito Santos, A., & Martínez Abad, F. (2019a). Grupo GRIAL. *IE Comunicaciones. Revista Iberoamericana de Informática Educativa*, 30, 33-48.
- García-Peñalvo, F. J., Rodríguez-Conde, M. J., Verdugo-Castro, S., & García-Holgado, A. (2019b). Portal del Programa de Doctorado Formación en la Sociedad del



- Conocimiento. Reconocida con el I Premio de Buena Práctica en Calidad en la modalidad de Gestión. In A. Durán Ayago, N. Franco Pardo, & C. Frade Martínez (Eds.), *Buenas Prácticas en Calidad de la Universidad de Salamanca: Recopilación de las I Jornadas. REPOSITORIO DE BUENAS PRÁCTICAS (Recibidas desde marzo a septiembre de 2019)* (pp. 39-40). Ediciones Universidad de Salamanca. <https://doi.org/10.14201/0AQ02843940>
- Gil-Fernández, R., León-Gómez, A., & Calderón, D. (2021). Influence of COVID on the educational use of Social Media by students of Teaching Degrees. *Education in the Knowledge Society, 2021, vol. 22, p. e23623*.
- Johnson, L., Becker, S. A., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *NMC Horizon Report: 2016 Higher Education Edition* (pp. 1-50). The New Media Consortium. <https://www.learntechlib.org/p/171478/>
- Kitchenham, B., & Charters, S. (2007). *Guidelines for performing systematic literature reviews in software engineering*.
- Knopik, T., & Oszwa, U. (2021). E-cooperative problem solving as a strategy for learning mathematics during the COVID-19 pandemic. *Education in the Knowledge Society (EKS), 22, e25176-e25176*.
- Mandinach, E. B., & Jimerson, J. B. (2016). Teachers learning how to use data: A synthesis of the issues and what is known. *Teaching and Teacher Education, 60, 452-457*.
- Manikas, K., & Hansen, K. M. (2013). Software ecosystems – A systematic literature review. *Journal of Systems and Software, 86(5), 1294-1306*. <https://doi.org/10.1016/j.jss.2012.12.026>
- Maybee, C., & Zilinski, L. (2015). *Data informed learning: A next phase data literacy framework for higher education*.
- Muñiz Fernández, J. (2010). Las teorías de los tests: Teoría clásica y teoría de respuesta a los ítems. *Papeles del Psicólogo: Revista del Colegio Oficial de Psicólogos*.
- Ruipérez-Valiente, J. A. (2020). El Proceso de Implementación de Analíticas de Aprendizaje. *RIED. Revista Iberoamericana de Educación a Distancia, 23(2), 85*. <https://doi.org/10.5944/ried.23.2.26283>
- Raffaghelli, J. E. (2018). Educators' Data Literacy. *MEDIA EDUCATION: STUDI, RICERCHE, BUONE PRATICHE, Teacher Education&Training on ICT between Europe and Latin America, 91-109*. <https://doi.org/10.4399/97888255210238>
- Siemens, G. (2012). Learning analytics: Envisioning a research discipline and a domain of practice. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge, 4-8*. <https://doi.org/10.1145/2330601.2330605>
- Sireci, S. G. (1998). The construct of content validity. *Social indicators research, 45(1), 83-117*.
- Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational researcher, 31(7), 15-21*.
- Society for Learning Analytics Research & Available from <https://www.solaresearch.org/about/>. (s. f.). Society for Learning Analytics Research (SoLAR). Recuperado 13 de enero de 2022, de <https://www.solaresearch.org/about/>