

Evaluación de la importancia de la ética, privacidad y seguridad en los estudios de Learning Analytics, en el marco de las conferencias LAK

Evaluation of the importance of ethics, privacy and security in Learning Analytics studies, under the LAK conferences

Daniel Amo¹, Marc Alier², David Fonseca³, Francisco José García-Peñalvo⁴, María José Casañ², Joan Navarro¹
daniel.amo@salle.url.edu, marc.alier@upc.edu, david.fonseca@salle.url.edu, fgarcia@usal.es, mjasany@essi.upc.edu, joan.navarro@salle.url.edu

¹Departament d'Enginyeria
La Salle, Universitat Ramon Llull
Barcelona, Spain

²Departament Eng. Serveis i Sistemes d'Informació
Universitat Politècnica de Catalunya
Barcelona, Spain

³Departament d'Arquitectura
La Salle, Universitat Ramon Llull
Barcelona, Spain

⁴Departamento de Informática y Automática
Universidad de Salamanca
Salamanca, Spain

Resumen- Los orígenes del Learning Analytics son difusos y rodeados por un intenso debate sobre su definición. Este debate se sitúa en el ámbito académico y comunidad científica. Además, se pretende identificar el alcance, retos y barreras en la aplicación de análisis de datos educativos. El presente artículo corresponde a una investigación más amplia, centrada en el tratamiento y gestión de datos educativos. Hemos conducido una Literature Review aplicando Text Analytics sobre los títulos, resúmenes y autores de los artículos publicados en todas las conferencias LAK (Learning Analytics & Knowledge, periodo 2011-2019). El objetivo del análisis de texto es doble. Por una parte, averiguar si en los congresos LAK existe debate alrededor de la temática "privacidad y seguridad en el tratamiento y uso de datos educativos en Learning Analytics". Por otra parte, dar una aproximación del nivel de profundidad y aportar nuevas direcciones de investigación, si así fuera necesario. El resultado refleja una amplia tendencia en los congresos LAK de hacer un tratamiento informatizado, predictivo y masivo de datos educativos para ilustrar casos de estudio, marcos teóricos y propuestas de enfoque. Muy pocos artículos presentados se concentran en ética y/o privacidad (pero con un alto impacto científico), y ninguno en seguridad.

Palabras clave: *Learning Analytics, Security, Privacy, Digital Identity.*

Abstract- The origins of Learning Analytics are diffuse and surrounded by an intense debate about its definition. This debate is located in the academic and scientific community. In addition, it aims to identify the scope, challenges and barriers in the practice of the definition This article corresponds to a broader investigation, focused on the treatment and management of educational data. We have conducted a Literature Review applying Text Analytics on the titles, abstracts and authors of articles published in all LAK conferences (Learning Analytics & Knowledge, period 2011-2019). The objective of text analysis is twofold. On the one hand, to find out if at the LAK congresses there is debate on the subject "privacy and security in the treatment and use of educational data in Learning Analytics". On the other hand, give an approximation of the level of depth and provide new research directions, if necessary. The result reflects a broad tendency in the

LAK congresses to make a computerized, predictive and massive treatment of educational data to illustrate case studies, theoretical frameworks and proposals for approach. Very few articles presented focus on ethics and/or privacy (but with a high scientific impact), and none on security.

Keywords: *Learning Analytics, Security, Privacy, Digital Identity.*

1. INTRODUCCIÓN

Learning Analytics (LA) ha tenido un gran impacto e interés en la comunidad científica desde su introducción en el 2010 hasta la actualidad (Figura 1). Este interés ha ido acompañado de un gran volumen de investigaciones, cuyos autores han intentado aplicar, definir, sus usos y explicar aportaciones posibles de LA y sus distintas definiciones. A pesar de los esfuerzos y definiciones de referencia generadas por distintos autores, aún no hay un consenso total y definitivo sobre en qué consiste LA exactamente. LA aún está en proceso de madurez.

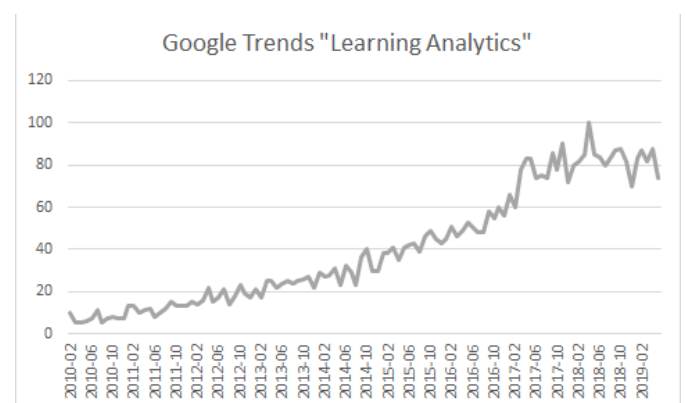


Figura 1: LA en Google Trends, 01 de junio del 2019.

Octubre 9-11, 2019, Madrid, ESPAÑA

V Congreso Internacional sobre Aprendizaje, Innovación y Competitividad (CINAIC 2019)

En los últimos 10 años la investigación en LA ha crecido. No ha sido así su transferencia a la práctica en educación, cuyo impacto ha sido mucho más lento (Aznarte, Hidalgo, Rubió, & Ruipérez, 2019). La amplia adopción de LA en educación supondría superar una serie de retos, como la adopción por parte de los docentes de la competencia digital y *data literacy* (Mandinach & Gummer, 2013).

La preocupación por la privacidad, seguridad y buen uso de los datos educativos recolectados en los procesos de Learning Analytics no suscitó interés hasta años después de su introducción. Algunos autores entienden que el uso de datos de alumnos es un asunto sensible que requiere de una atención ética. Más adelante, aspectos como la ética, privacidad o seguridad de los datos entraron en debate provocando que distintos autores empezaran a investigar a su alrededor (Drachsler & Greller, 2016; Hoel & Chen, 2016).

Nuestro interés pivota principalmente en la privacidad y seguridad de datos educativos acerca de los estudiantes, a menudo menores. Es una cuestión importante y sensible que requiere de un acercamiento en profundidad. LA precisa la recolección de datos de las interacciones de los alumnos para después analizarlos con el objeto de extraer *insights*. En algunos casos se están manipulando datos de menores, hecho que supone una sensibilidad adicional. Los últimos sucesos en cuanto a malos usos de datos personales (Herold, 2014) y las publicaciones de nuevos reglamentos por parte de autoridades gubernamentales (European Parliament & Council of the European Union, 2016) ponen de manifiesto la necesidad de regular por ley el uso de datos educativos en herramientas de LA. Por lo tanto, es fundamental asegurar la privacidad y seguridad en procesos de LA.

Consideramos que los congresos LAK son de referencia para la comunidad académica, puesto que es donde se origina de la discusión inicial sobre la práctica y validez de Learning Analytics. En consecuencia, el objetivo de este trabajo es averiguar si en los congresos LAK existe debate alrededor de la temática “privacidad y seguridad en el tratamiento y uso datos educativos en Learning Analytics”, así como dar una aproximación del nivel de profundidad y aportar nuevas direcciones de investigación, si así fuera necesario.

La estructura de este trabajo se organiza en cinco secciones. Una primera a modo de introducción. Una segunda en la que exponemos el contexto de la investigación. Una tercera donde describimos la metodología utilizada para conseguir los resultados. Una cuarta en la que exponemos los resultados conseguidos, y una quinta a modo de fin de trabajo en la que exponemos las conclusiones extraídas.

2. CONTEXTO

El concepto de Learning Analytics aún está en proceso de definición. La primera referencia en Wikipedia data del 23 de Agosto del 2010 (“Learning Analytics,” 2010). Define Learning Analytics como “el uso de datos y modelos para predecir el progreso y rendimiento de los alumnos, y la habilidad de actuar con esa información”. Esta definición tiene una influencia directa del Educational Data Mining, y en realidad surge de la página web de nextgenlearning.com (Next Generation, 2010), datada en 11 de Julio del 2010. La entrada creada en Wikipedia es muy relevante, puesto que es resultado del grupo de discusión que abre George Siemens para debatir

sobre este nuevo concepto de análisis de datos educativos (“Learning Analytics GoogleGroups,” 2010).

Siemens -que aboga por el conectivismo- es el primero en entrar en debate académico para definir Learning Analytics desde una perspectiva pedagógica (Siemens, 2010). A las definiciones de entonces les añade datos más allá del entorno virtual de aprendizaje (EVA), como pueden ser redes sociales o blogs personales. Describe la analítica del aprendizaje como “the use of intelligent data, learner-produced data, and analysis models to discover information and social connections, and to predict and advise on learning” (Siemens, 2010). La definición de Siemens es más amplia en el sentido de que se adapta al sistema educativo y lo complementa. Su descripción, en sus propias palabras, “es menos limpia, pero no intenta modificar el sistema educativo”, sino utilizar los resultados analíticos para mejorarlo.

La definición de EDUCASE se orienta al uso exclusivo de la minería de datos. En cambio, la definición de Siemens es más conectivista y orientada a ampliar el espectro de Learning Analytics. Como la amplia más allá de considerar aspectos analíticos, Siemens indica que LA concierne a otros conceptos como acciones, definición del currículum, adaptación y personalización, predicción y establecimiento de competencias (Figura 2).

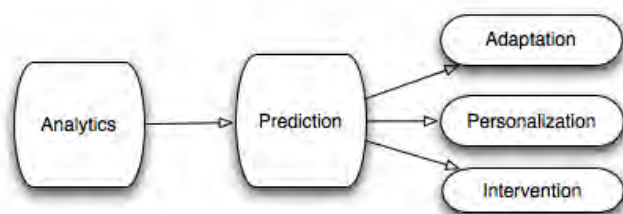


Figura 2: Objetivos de uso de los datos en Learning Analytics

Nótese aún el predominio del campo de minería de datos en el sentido que Siemens explica Learning Analytics como un proceso puramente predictivo.

Siemens establece el 2010 como el año en el que el concepto Learning Analytics toma presencia en el mundo educativo. Al mismo tiempo podemos apuntar a Siemens como el máximo impulsor de Learning Analytics, tras su trayectoria de cursos conducidos alrededor de esta temática (Admin, 2010), de congresos LAK y charlas, pero sobretudo por la constitución de SOLAR (Society for Learning Analytics Research) (SoLAR, 2019), en el que fue Presidente Fundador entre el 2013 y el 2015, y donde ejerce actualmente de miembro activo.

Los congresos LAK tienen sus orígenes en el 2011, año en el que se efectuó el primer encuentro en Banff, Canadá. Estos congresos tienen una alta relevancia, puesto que los resultados académicos de la comunidad de investigadores LAK representan una muestra muy significativa de las tendencias en la analítica del aprendizaje. En cada encuentro se presentan investigaciones que ponen de manifiesto en qué ámbitos hay interés de usar el análisis de datos educativos. Por consiguiente, toma fuerza analizar estos encuentros para conocer la evolución de Learning Analytics, así como los pasados, presentes y futuros campos de interés y tendencias en sus posibles aplicaciones y usos.

3. DESCRIPCIÓN

En este manuscrito presentamos un primer análisis textual de los títulos de artículos en distintos congresos LAK. Para realizar el análisis textual hemos recolectado todos los títulos de los artículos presentados a lo largo de los 9 años que llevan realizándose los congresos. Los *proceedings* de cada uno de los congresos están publicados en la ACM Digital Library (Athabasca University, Technology Enhanced Knowledge Research Institute, University of Queensland, Centre for Educational Innovation and Technology, Association for Computing Machinery., & ACM Digital Library., 2011). Mediante un proceso de *data scraping* descargamos los títulos, los resúmenes y los autores de cada uno de los títulos. Usamos expresiones regulares para extraer los datos.

- Para extraer títulos usamos la expresión regular: $(?:<td.*>)(.*)?(?:</td>)$
- Para extraer autores utilizamos la expresión regular: $(?:<a.*author.*>)(.*)?(?:)$
- Para extraer resúmenes utilizamos la expresión regular: $(?:<span.*<p>)(.*)?(?:</p>)$

El siguiente paso fue desarrollar dos algoritmos, uno para análisis de los títulos y otro para los autores de estos.

A. Análisis de títulos

Para el análisis textual de los títulos y resúmenes creamos un proceso en Python usando la librería Natural Language Toolkit (NLTK Project, 2019). NLTK permite realizar análisis de textos más allá de un simple conteo de palabras para conocer aspectos como gramática o complejidad léxica.

El proceso que llevamos a cabo consistió en los siguientes 3 pasos:

1. Seleccionar palabras de más de 3 letras: En este paso eliminamos palabras de baja significancia y que podían alterar drásticamente el resultado analítico.
2. Seleccionar nombres, verbos y adjetivos: En este paso seleccionamos las palabras de alta significancia y que aportaban valor al resultado analítico.
3. Agregar y ordenar: En este paso realizamos un conteo de las palabras resultantes de aplicar los dos filtros anteriores. Finalmente realizamos una ordenación descendiente para colocar las palabras más citadas en primer lugar.

A pesar de utilizar técnicas automáticas de conteo de palabras de una gramática específica, el análisis de los títulos nos ofrece la oportunidad de conocer qué aspectos son relevantes en cada uno de los LAK. En la sección de resultados mostramos las distintas tendencias detectadas y las visualizaciones que dan soporte a la discusión.

B. Análisis de autores

El análisis de autores nos permite conocer el interés e impacto de los congresos LAK para la comunidad científica. A la vez nos permite detectar los autores referentes en las distintas temáticas de la analítica de datos educativos.

Para el análisis de autores realizamos un primer *data scraping* dando resultado una serie de archivos de texto con el listado de los autores partícipes en cada uno de los congresos. Seguidamente desarrollamos una serie de *shell scripts* y sintaxis Bash. Despreciamos Python y tomamos un cambio de rumbo en el lenguaje de desarrollo debido a la facilidad, plasticidad y alta capacidad de la sintaxis Bash y *shell scripts* para trabajar con distintos ficheros de textos y extracción de patrones mediante expresiones regulares.

4. RESULTADOS

Mediante un análisis de los títulos extraemos los intereses de la comunidad científica partícipe de los congresos LAK, la evolución de la comprensión de LA y otras cuestiones de impacto.

Para explicar los resultados hemos dividido su exposición cinco partes. En la primera mostramos la limpieza de datos que llevamos a cabo para eliminar interferencias en las conclusiones. En la segunda exponemos por cada LAK las palabras más citadas según los títulos presentados. En la tercera exponemos el impacto de los autores y su repercusión en la fluctuación de un gran número de palabras. En la cuarta exponemos una definición genérica de LA a partir del análisis de las palabras de cada LAK. Finalizamos la exposición de los resultados mostrando la preocupación en relación con la ética, privacidad y seguridad en LA.

A. Limpieza de datos

La Figura 3 muestra una *small multiples* de la nube de palabras generada por los términos más citados en los títulos y resúmenes de los artículos publicados desde el LAK11 al LAK19.

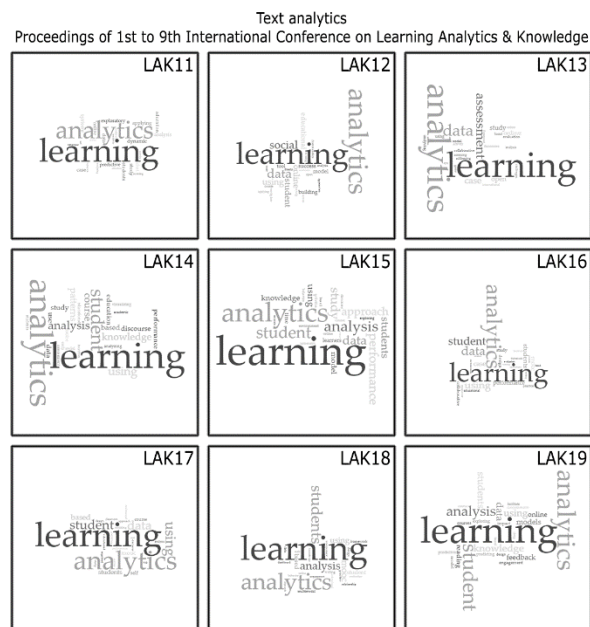


Figura 3: *Small multiples* de los términos más repetidos en los títulos LAK11-19

Visualmente se perciben dos palabras destacadas como son “learning” y “analytics”. Por consiguiente, y para asegurar resultados relevantes en el análisis textual de los títulos,

justificamos eliminar tanto las palabras “learning” y “analytics” como sus conjugaciones.

B. Palabras relevantes en los LAK11-19

La siguiente tabla muestra los términos más citados en las primeras 5 posiciones por cada LAK. En algunas filas aparecen más de cinco posiciones. Hemos considerado remarcar la posición de aquellas palabras relacionadas con aspectos éticos, de privacidad y seguridad en estos congresos.

Tabla 1

Resultados del análisis de texto de los artículos LAK 11-19

Congreso	Número de citas y términos
LAK11	(2) systems, predictive, online, education, data, case, study, exploratory, towards, dynamic, students , applying (1) learnometrics, metrics, objects, attention, visualization, recommendation, networks, ethical...
LAK12	(9) data (8) online, student , using (5) educational (4) building, model, tool, success (3) communities, meets, mining, workplace, open, case, study, course
LAK13	(7) assessment, data (5) case (4) online, study (3) evaluation, social, collaborative, using, analyzing, open (2) addressing, learner, issues, topic, models, discussions, pedagogical, intervention, student , multimodal, workshop...
LAK14	(1) ethical... (10) student (6) using (5) course, patterns, system (4) knowledge, study, performance, education (3) assessment, educational, success, data, visualizing, early, students , social, analyzing, discourse (1) ethical
LAK15	(12) student (9) performance, study (8) students , approach (7) data, using (2) privacy (1) ethical
LAK16	(16) data (15) student (14) using (9) students (8) case, mooc, performance (2) privacy (1) ethical
LAK17	(20) using (18) student (16) data (10) students (9) mooc (2) privacy
LAK18	(11) students (6) mooc, using (4) practice, design, frameworks, writing, student , multimodal

(3) knowledge, scale, educational, environment, behavior, test, dashboard, study, evaluation, online, research, strategies, data...

(2) evaluating, retention, moocs, assessment, teachers, model, feedback, tool, method...

(1) **ethics**

LAK19

(13) **student**

(6) using, **students**, knowledge

(5) data, reading, models, online, feedback

(4) predicting, exploring, predictive, engagement, model, impact

(3) time, study, course, facilitate, deep, trajectories, comprehension, moocs, design...

(1) **privacy**

C. Fluctuación en la participación

Comprobamos que en cada LAK hay una serie de palabras que se repiten y marcan un hilo argumental homogéneo. En cambio, aparecen muchas otras palabras que aportan contexto adicional al pivotar alrededor de las que se mantienen constantes.

La fluctuación en la participación a lo largo de los nueve congresos LAK explica la fluctuación en la aparición de las palabras poco constantes. La Figura 4 muestra autores y artículos presentados en cada LAK. Se puede comprobar un crecimiento constante de la participación, aunque en el LAK18 disminuye de forma drástica a poco más de la mitad. Por lo tanto, una gran cantidad de autores demuestra que existen muchos focos de interés alrededor de LA.

#Papers & #Authors

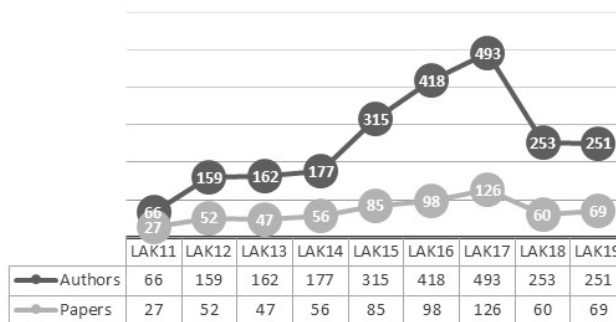


Figura 4: Número de artículos y autores presentados a lo largo de los distintos LAK11-19

D. Definición genérica de Learning Analytics

Cada uno de los LAK presenta una serie de palabras que consideramos importantes, aquellas que tienen más citas, y otras que consideramos que explican el contexto, aquellas que se agrupan en puntuaciones más bajas y que complementan las que tienen más citas. Desde esta perspectiva desarrollamos el análisis de los términos reflejados en la Tabla 1.

Puede parecer una obviedad decir que el foco de interés de los distintos congresos LAK sea los estudiantes. No obstante, los resultados del análisis textual demuestran que no es así en un inicio, donde los primeros LAK se centraban en otros aspectos a mejorar o crear.

Como primer resultado detectamos un contexto de investigación inicial de campos muy diversos, con los que se procura definir unos primeros marcos, modelos, técnicas y

recomendaciones en el uso de LA. Este contexto teórico se refina y evoluciona durante su aplicación a lo largo de los últimos cinco congresos, en los que se percibe una madurez y constancia con foco en los estudiantes o el uso explícito de sus datos:

- LAK11: De los 27 artículos presentados extraemos un conjunto de 12 palabras que se citan 2 veces y un conjunto de 101 palabras que se citan 1 sola vez. Puede que por ser el primer congreso, el abasto de los distintos artículos sea aún muy amplio. Hay demasiadas palabras al mismo nivel que demuestra que el concepto de LA o el foco de investigación aún está por definir.
- LAK12: De los 52 artículos presentados extraemos conjuntos de palabras que apuntan a un foco concreto de investigación como es el estudiante o “el uso de datos de estudiantes en contextos online para su estudio y análisis”.
- LAK13: De los 47 artículos presentados extraemos conjuntos de palabras que difuminan el foco del congreso anterior. Disminuye el interés sobre el estudiante, aunque se mantiene el interés en entornos online. No es hasta este tercer congreso que la palabra “feedback”, aspecto esencial en LA, aparece en alguno de los títulos de los artículos.
- LAK14: De los 56 artículos presentados extraemos conjuntos de palabras como “student” o “students” que vuelven a definir un interés directo sobre el estudiante. La importancia de analizar el estudiante, sus datos o su contexto directo, ya extraída en el LAK12, se mantiene a lo largo de los siguientes congresos.
- LAK15: De los 85 artículos presentados extraemos conjuntos de palabras que siguen demostrando el interés hacia el uso de datos y análisis de los estudiantes. La privacidad como tema central toma una ligera relevancia que se mantiene por los dos siguientes congresos.
- LAK16: De los 98 artículos presentados extraemos conjuntos de palabras que confirman de forma definitiva el primer interés de LA surgido del LAK12: “el uso de datos de estudiantes en contextos online”. Palabras pivotantes como “MOOC” toman relevancia en este congreso.
- LAK17: De los 126 artículos presentados extraemos conjuntos de palabras que siguen el mismo interés del anterior congreso. Los MOOCs se posicionan también como relevantes en esta edición. Aspectos de privacidad aparecen en dos de los artículos presentados.
- LAK18: De los 60 presentados extraemos conjuntos de palabras que mantienen el mismo interés en cuanto al estudio de estudiantes. No obstante, las palabras citadas en los niveles 3 y 2 muestran otra vez un gran amplio interés de campos heterogéneos. En cuestiones de privacidad se vuelve a perder foco directo, aunque vuelve a aparecer el foco perdido en la ética durante el congreso anterior.

- LAK19: De los 69 artículos presentados extraemos conjuntos de palabras que mantienen el mismo interés de los anteriores congresos. Se muestra en este congreso un claro matiz predictivo de la LA mediante la cita directa de palabras como “predicting” o “predictive” en 11 de 69 artículos presentados (15%).

Mediante los resultados extraídos de cada uno de los LAK nos atrevemos a formular una definición genérica de LA como el “uso de datos de alumnos en contextos online para su estudio y análisis con técnicas predictivas en un entorno privado y seguro”. Esta definición es muy difusa y no se puede concretar en el primer LAK11, puesto que aún se está en modo exploración de sus posibilidades. Es a partir del LAK14 donde se normaliza la aparición de los términos que dan como resultado nuestra descripción: “students”, “using”, “data”, “online”, “mooc”, “study” y “predictive”. Aunque los términos privacidad y seguridad tienen poco o muy poco impacto en los resultados, consideramos que la descripción debe contenerlas si queremos abordar la problemática presentada en la introducción.

Otras palabras complementan nuestra definición genérica dotándola de contexto y propósito. En algunos artículos la palabra “feedback” denota el interés de mejora de la retroalimentación requerida para el alumnado mediante el uso de LA, en otros la palabra “dashboard” demuestra usos efectivos de visualizaciones concretas mediante el uso de LA, e incluso en otros se anula la cualidad “online” cuando se presentan conceptos como Multimodal Learning Analytics en entornos físicos de aula.

E. Ética, privacidad y seguridad

No hemos encontrado un foco suficiente de interés en los títulos como para considerar una fuerte relevancia de aspectos éticos y privacidad en los congresos LAK. De los términos relacionados con “privacy” se detectan solo en 7 títulos publicados desde el 2015 al 2019. De los términos relacionados con “ethic” se detectan solo en 6 títulos desde el 2012 al 2018.

A pesar de estos bajos resultados, y mediante la extracción y análisis de los resúmenes, sí detectamos una preocupación generalizada en cuestiones éticas y de privacidad a lo largo de los congresos LAK. Además, podemos decir que hay una clara preocupación desde la cúpula, puesto que los que se pronuncian en los artículos son autores de peso dentro de la comunidad LAK como son Siemens, Pardo o Drachslers. Por consiguiente, consideramos que la inclusión de privacidad debe estar presente en nuestra definición genérica.

Términos relacionados con “security” no aparecen mencionados en ninguno de los títulos de los congresos LAK11-19. Solo aparece la palabra “security” en el resumen de un artículo del LAK18 cuyo foco es el uso de Blockchain. Destacamos que la comunidad científica, incluso artículos publicados por autores de este trabajo (Amo, Fonseca, Alier, García-Peñalvo, & Casañ, 2019), apunta a Blockchain como una tecnología emergente no adecuada para asegurar una adecuada privacidad y seguridad de los datos sensibles de alumnos. Por consiguiente, llegamos a la conclusión de que la seguridad no es aspecto de interés para los investigadores LAK.

5. CONCLUSIONES

En el presente trabajo exponemos los orígenes de LA, sus representantes iniciales, las organizaciones que ahora son referencia y el debate aún abierto para establecer una definición consensuada y definitiva. El debate abierto demuestra que LA es aún un campo científico para explorar y matizar. Además, exponemos por qué los congresos LAK son de máximo peso en este campo científico y por lo tanto, de interés de estudio.

Exponemos también la problemática y retos en relación con la privacidad y seguridad en procesos LA. Los datos de alumnos son sensibles, en algunos casos de menores, donde una serie de leyes generan un marco claro de recolección y tratamiento. El problema de la privacidad y seguridad en los procesos LA es el verdadero motivo que nos empuja a analizar los congresos LAK. El objetivo final es comprender el nivel de importancia que tiene el contexto de la privacidad y seguridad en los congresos LAK.

Hemos realizado un Literature Review aplicando Text Analytics sobre los títulos, resúmenes y autores de los artículos publicados en todas las conferencias LAK. El análisis de los títulos nos ha permitido comprender los distintos focos de interés a lo largo de los congresos LAK. Los resultados muestran un contexto de investigación inicial de campos muy diversos, con los que se procura definir unos primeros marcos, modelos, técnicas y recomendaciones en el uso de LA. Este contexto teórico inicial se refina y evoluciona a lo largo de los restantes congresos, en los que se percibe un foco cada vez más fuerte en los estudiantes o uso explícito de sus datos.

Los resultados también nos han permitido generar una definición genérica. Junto a la necesidad de asegurar una privacidad y seguridad adecuada a la ley, establecemos la definición genérica de LA como el “uso de datos de alumnos en contextos online para su estudio y análisis con técnicas predictivas en un entorno privado y seguro”.

En términos de seguridad, nuestros resultados muestran que la no es de interés en los congresos LAK. Solamente un artículo de los más de 620 publicados cita aspectos de seguridad. Además, el foco del artículo es Blockchain, una tecnología muy criticada en la comunidad científica. Por consiguiente, creemos que es necesario un debate y una posición contundente dentro de los congresos LAK para defender la seguridad de datos educativos en procesos LA.

6. AGRADECIMIENTOS

Con el soporte de la Secretaria d'Universitats i Recerca del Departament d'Empresa i Coneixement de la Generalitat de Catalunya con la ayuda 2017 SGR 934.

REFERENCIAS

Admin. (2010). Welcome to the open online course on Learning & Knowledge Analytics. Retrieved April 27, 2019, from <https://www.learninganalytics.net/page/4/>

Amo, D., Fonseca, D., Alier, M., García-Peñalvo, F. J., & Casañ, M. J. (2019). Personal Data Broker Instead of Blockchain for Students' Data Privacy Assurance (pp. 371–380). Springer, Cham. https://doi.org/10.1007/978-3-030-16187-3_36

Athabasca University. Technology Enhanced Knowledge Research Institute, P., University of Queensland. Centre for Educational Innovation and Technology, G.,

Association for Computing Machinery., G., & ACM Digital Library., D. (2011). *LAK '11 : proceedings of the 1st International Conference on Learning Analytics and Knowledge, February 27 - March 1, 2011, Banff, Alberta, Canada*. ACM. Retrieved from <https://dl.acm.org/citation.cfm?id=2090116>

Aznarte, J. L., Hidalgo, R., Rubió, E., & Ruipérez, J. A. (2019). Canal UNED - Inteligencia Artificial y la personalización del aprendizaje (Mesa redonda). Retrieved June 27, 2019, from <https://canal.uned.es/video/5caf043ba3eeb095388b4567>

Drachler, H., & Greller, W. (2016). Privacy and analytics: it's a DELICATE issue a checklist for trusted learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 89–98).

European Parliament, & Council of the European Union. (2016). Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) (GDPR). Retrieved June 27, 2019, from <https://eur-lex.europa.eu/legal-content/en/TXT/?uri=CELEX%3A32016R0679>

Herold, B. (2014). inBloom to Shut Down Amid Growing Data-Privacy Concerns. Retrieved June 21, 2018, from http://blogs.edweek.org/edweek/DigitalEducation/2014/04/inbloom_to_shut_down_amid_growing_data_privacy_concerns.html

Hoel, T., & Chen, W. (2016). *Implications of the European Data Protection Regulations for Learning Analytics Design*. Retrieved from http://ec.europa.eu/justice/data-protection/reform/index_en.htm

Learning Analytics. (2010). Retrieved April 2, 2019, from https://en.wikipedia.org/wiki/Learning_analytics

Learning Analytics GoogleGroups. (2010). Retrieved April 27, 2019, from <https://groups.google.com/forum/#!forum/learninganalytics>

Mandinach, E. B., & Gummer, E. S. (2013). A Systemic View of Implementing Data Literacy in Educator Preparation. *Educational Researcher*, 42(1), 30–37. <https://doi.org/10.3102/0013189X12459803>

Next Generation. (2010). Learning Analytics. Retrieved April 27, 2019, from <http://web.archive.org/web/20100711081424/http://www.nextgenlearning.com/the-challenges/learning-analytics>

NLTK Project. (2019). Natural Language Toolkit — NLTK 3.4.3 documentation. Retrieved April 27, 2019, from <https://www.nltk.org/>

Siemens, G. (2010). What are Learning Analytics? Retrieved April 27, 2019, from <http://web.archive.org/web/20100831015646/http://www.elearnspace.org/blog/2010/08/25/what-are-learning-analytics/>

SoLAR. (2019). About SoLAR. Retrieved April 27, 2019, from <https://solaresearch.org/about/>