

# El Proceso de Implementación de Analíticas de Aprendizaje

## (The Implementation Process of Learning Analytics)

José A. Ruipérez-Valiente  
*Universidad de Murcia, UMU (España)*

DOI: <http://dx.doi.org/10.5944/ried.23.1.26283>

### **Cómo referenciar este artículo:**

Ruipérez-Valiente, J. A. (2020). El Proceso de Implementación de Analíticas de Aprendizaje. *RIED. Revista Iberoamericana de Educación a Distancia*, 23(2), pp. 85-101. doi: <http://dx.doi.org/10.5944/ried.23.1.26283>

### **Resumen**

Con el despegue de la popularidad del área de analítica de aprendizaje durante la última década, numerosas investigaciones han surgido y la opinión pública se ha hecho eco de esta tendencia. Sin embargo, la realidad es que el impacto que ha tenido en la práctica ha sido bastante bajo, y se está produciendo poca transferencia a las instituciones educativas. Una de las posibles causas es la elevada complejidad del campo, y que no existan procesos claros; por ello, en este trabajo, se propone un pragmático proceso de implementación de analíticas de aprendizaje en cinco etapas: 1) entornos de aprendizaje, 2) recolección de datos en crudo, 3) manipulación de datos e ingeniería de características, 4) análisis y modelos y 5) aplicación educativa. Además, se revisan una serie de factores transversales que afectan esta implementación, como la tecnología, ciencias del aprendizaje, privacidad, instituciones y políticas educacionales. El proceso que se detalla puede resultar de utilidad para investigadores, analistas de datos educacionales, educadores e instituciones educativas que busquen introducirse en el área. Alcanzar el verdadero potencial de las analíticas de aprendizaje requerirá de estrecha colaboración y conversación entre todos los actores involucrados en su desarrollo, que permita su implementación de forma sistemática y productiva.

*Palabras clave:* analítica de aprendizaje; minería de datos educacionales; metodología; tecnología educativa; educación informada por datos; ciencia de datos.

### **Abstract**

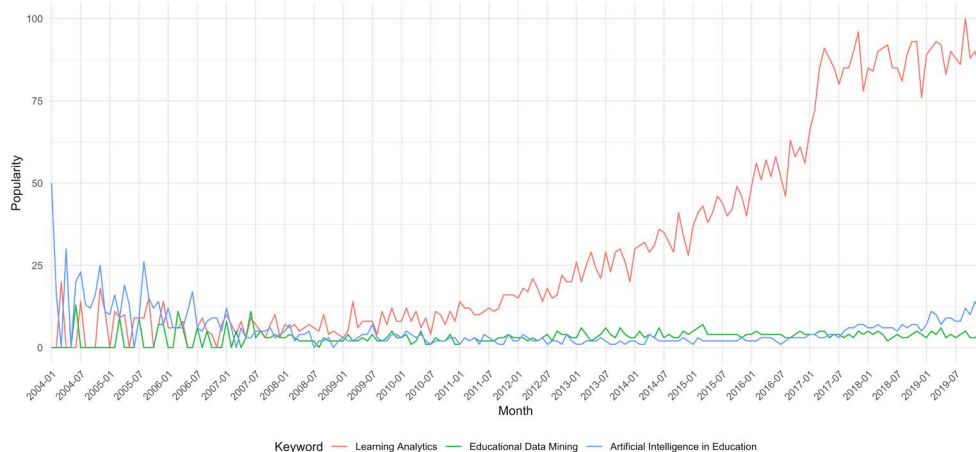
With the popularity takeoff of the learning analytics area during the last decade, numerous research studies have emerged and public opinion has echoed this trend as well. However, the fact is that the impact the field has had in practice has been quite limited, and there has been

little transfer to educational institutions. One of the possible causes is the high complexity of the field, and that there are no clear implementation processes; therefore, in this work, we propose a pragmatic implementation process of learning analytics in five stages: 1) learning environments, 2) raw data capture, 3) data tidying and feature engineering, 4) analysis and modelling and 5) educational application. In addition, we also review a series of transverse factors that affect this implementation, like technology, learning sciences, privacy, institutions, and educational policies. The detailed process can be helpful for researchers, educational data analysts, teachers and educational institutions that are looking to start working in this area. Achieving the true potential of learning analytics will require close collaboration and conversation between all the actors involved in their development, which might eventually lead to the desired systematic and productive implementation.

*Keywords:* learning analytics; educational data mining; methodology; educational technologies; data-driven education; data science.

El área de analítica de aprendizaje (*learning analytics*, LA) ha despegado durante la última década situándose como una de las comunidades de mayor crecimiento dentro del área de investigación educacional (Gasevic, Dawson, Mirriahi y Long, 2015). Un importante reconocimiento de la madurez a la que ha llegado el área, es que actualmente en 2020 la conferencia más importante del área, Learning Analytics and Knowledge, está situada en las métricas de Google Scholar del área de tecnologías educativas<sup>1</sup> en sexta posición y como única conferencia entre las 20 fuentes más citadas. En la primera edición de esta conferencia (LAK'11), el campo fue definido de forma amplia como “la medición, recopilación, análisis e informe 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”<sup>2</sup>. Esta definición y los objetivos del área están bastante relacionados con otros campos que se encuentran asentados por más años, como la minería de datos educacionales (*educational data mining*, EDM) y la inteligencia artificial en la educación (*artificial intelligence in education*, AIED). Estos campos hermanos han buscado hacer un avance conjunto, pero focalizándose en partes diferentes del proceso, y las comunidades han buscado que los investigadores colaboren fuertemente (Siemens y Baker, 2012). Este interés y popularidad han trascendido de la investigación a la comunidad de practicantes y educadores. En la figura 1 se muestra la popularidad de 0 a 100 obtenida a través de Google Trends de estas tres palabras clave a lo largo de los años. Se puede observar el destacable crecimiento de *learning analytics* desde que la comunidad empezó a formarse en 2011. Se puede ver también como las otras dos palabras clave tenían más presencia inicialmente, pero han sido un poco eclipsadas en los últimos años teniendo un crecimiento en la popularidad en búsquedas mucho más modesto.

Figura 1. Popularidad temporal de 0 a 100 de las palabras clave en Google Trends.



La gran cantidad de trabajos publicados en el área han motivado a realizar numerosos estudios que han revisado parte de la literatura (Aldowah, Al-Samarraie y Fauzy, 2019; Daniel, 2012; Leitner, Khalil y Ebner, 2017; Viberg, Hatakka, Bälter y Mavroudi, 2018). Estas revisiones han descrito el crecimiento del área, cuáles son los métodos más comúnmente utilizados, así como las aplicaciones educativas que tienen dichos estudios con mayor frecuencia; la mayoría de los trabajos coinciden en el elevado potencial del área para la transformación del sector educativo, si se afronta correctamente, ya que en numerosas ocasiones se ve como los estudios pueden perder el foco de centrarse en mejorar el aprendizaje de los estudiantes (Gašević, Dawson y Siemens, 2015). Sin embargo, también se coincide en que la mayoría de los estudios se centran en casos de estudio independientes y tratando de entender el proceso de aprendizaje de los estudiantes a muy baja escala, y que la transferencia a la práctica y la implementación sistemática de analíticas de aprendizaje por parte de instituciones ha sido muy escasa (Viberg et al., 2018).

Una de las principales barreras para la implementación de forma sistemática por parte de las instituciones y los educadores, es la alta multidisciplinariedad y complejidad del campo (Ferguson, 2012). Por ello, en este trabajo se tiene como objetivo el proporcionar una guía de los pasos a seguir para implementar analíticas de aprendizaje, y que otros factores contextuales y transversales a estos pasos hay que tener fuertemente en cuenta en estos proyectos. Mientras que ya ha habido trabajos que intentan modelar o proponer marcos de referencia del área, estos se centran más en una visión investigadora, y no en una visión más práctica y de implementación, haciendo que sean difícilmente accesibles a nivel de conocimientos por educadores o instituciones. Por ello, en esta guía se presenta de forma práctica y pragmática

los pasos y elementos a tener en cuenta durante el proceso de implementación de analíticas de aprendizaje.

## TRABAJOS RELACIONADOS

Ha habido diversos investigadores que han propuesto marcos de referencia, o hablan sobre los diferentes actores involucrados en las analíticas de aprendizaje. Quizás la primera iniciativa fue por parte de Campbell y Oblinger en la que se proponía un modelo de analíticas académicas basado en cinco etapas: capturar, reportar, predecir, actuar y refinar (Campbell y Oblinger, 2007). Este modelo estaba pensado para funcionar como un motor para la toma de decisiones o la guía de acciones, se empezaba haciendo una captura de datos relacionados con los estudiantes, se reportaba la información obtenida a través de dichos datos, se podían usar los datos para predecir el futuro, lo cual permitía actuar en base a dicha predicción, y cerrar el ciclo mediante el refinamiento del sistema educacional.

Posteriormente con el auge del campo en 2012, surgieron nuevas propuestas para intentar darle un marco común a los trabajos que iban apareciendo en la literatura. Una de ellas fue el trabajo de Clow, en el que define el ciclo de las analíticas de aprendizaje en cuatro pasos: estudiantes, datos, métricas e intervenciones (Clow, 2012). El proceso en este caso empieza centrado en los estudiantes, que generan una serie de datos en los entornos de aprendizaje, estos datos se pueden procesar para generar métricas o analíticas que proporcionan información sobre el proceso de aprendizaje, y esta información puede ser usada por los diferentes actores para actuar en consecuencia y mejorar el contexto educacional, finalmente cerrando el ciclo de nuevo con los estudiantes. Chatti y sus colegas propusieron un modelo de referencia de la analítica de aprendizaje basado en cuatro bloques (Chatti, Dyckhoff, Schroeder y Thüs, 2012). El primero son los datos educacionales recogidos y los entornos donde está sucediendo el aprendizaje. El segundo son los actores que intervienen en este aprendizaje, incluyendo no solo a los estudiantes y profesores, sino también a tutores, instituciones, administradores de sistemas o investigadores. El tercero hace referencia a los objetivos específicos que tiene la aplicación de analítica de aprendizaje en función de los actores involucrados. Por último, el cuarto hace referencia a los métodos que hay que implementar para la consecución de dichos objetivos. Finalmente, el modelo propuesto por Greller y Drachsler se centra en las dimensiones a tener en cuenta en el campo de analítica de aprendizaje (Greller y Drachsler, 2012). Las dimensiones críticas que se describen en este trabajo son con respecto a los objetivos propuestos en el proyecto de analíticas de aprendizaje, los datos educacionales recopilados, los instrumentos que se van a utilizar (tecnología, algoritmos, teorías...) y los actores que intervienen en este proyecto; por último, incluyen dos dimensiones que no se habían discutido hasta ahora, como son las limitaciones internas que puedan existir debido a competencias a la aceptación

de estas analíticas, así como las restricciones externas que puedan haber por convenciones o normas.

Muchos de estos trabajos han propuesto marcos de referencia de alta calidad y complejidad, a través de los cuales se pueden modelar las diferentes dimensiones que influyen en proyectos de analítica de aprendizaje. Sin embargo, a pesar de su valor, en ocasiones pueden resultar para nuevos investigadores, educadores o instituciones complicados de entender, de cara a cuáles son los pasos a seguir para desarrollar este tipo de proyectos. En este trabajo, nos centramos de una manera más pragmática en el proceso de implementación de las analíticas de aprendizaje desde un punto de vista más técnico y centrado en el usuario y aplicación. Para ello presentamos el proceso de una forma cronológica, indicando cuáles son las preguntas clave en cada etapa y los elementos transversales que afectan. La forma de presentarlo es similar a como se podrían llevar a cabo otros proyectos relacionados con ciencia de datos, pero en este caso específicamente adaptado al contexto educativo.

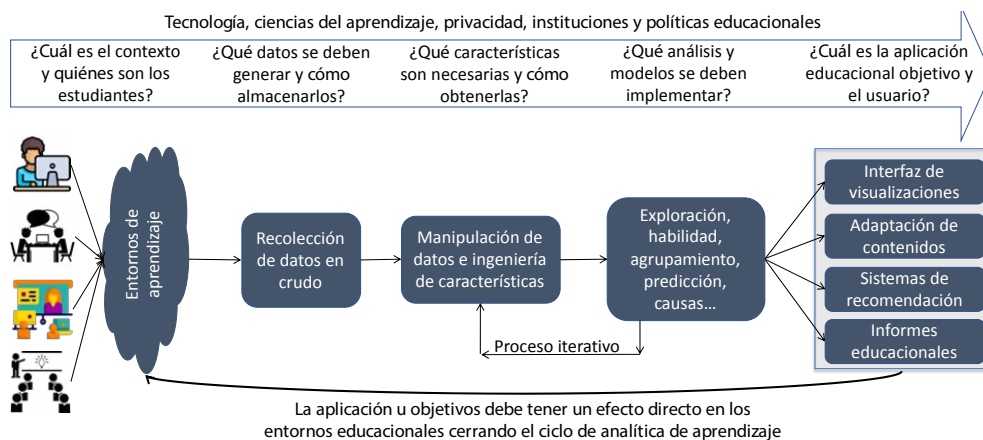
## **EL PROCESO DE IMPLEMENTACIÓN DE ANALÍTICAS DE APRENDIZAJE**

Este proceso está basado en la experiencia de desarrollar distintos proyectos de analítica de aprendizaje, y guarda cierta similitud con los presentados en los trabajos relacionados, pero se encuentra más centrado en las etapas de implementación que sigue un proyecto de analítica de aprendizaje, así como los elementos y preguntas que se deben tener en cuenta. El proceso tiene las siguientes cinco etapas:

- 1. Entornos de aprendizaje:** ¿Cuál es el contexto y cuáles son los estudiantes?
- 2. Recolección de datos en crudo:** ¿Qué datos se deben generar y cómo almacenarlos?
- 3. Manipulación de datos e ingeniería de características:** ¿Qué características son necesarias y cómo obtenerlas?
- 4. Análisis y modelos:** ¿Qué análisis y modelos se deben implementar?
- 5. Aplicación educacional:** ¿Cuál es la aplicación educacional objetivo y el usuario?

Además, de forma transversal hay otros elementos que se deben tener en cuenta, como las tecnologías a utilizar, las teorías y ciencias del aprendizaje, la privacidad de los usuarios, así como las instituciones y las políticas educacionales. Por último, que dependiendo del proyecto no es necesario que se pasen por todas estas etapas, la mayoría de investigación en el área de analítica de aprendizaje se centra sólo en los pasos 3 y 4. Ahondamos ahora en cada una de estas etapas.

Figura 2. El proceso de implementación de analíticas de aprendizaje



## Entornos de Aprendizaje

El primer paso del proceso sucede en el entorno y contexto de aprendizaje y con los usuarios que intervienen en el mismo. Con la llegada de la educación a través de entornos digitales, la posibilidad de recolectar datos fácilmente se fortaleció. Tradicionalmente, los entornos usados en educación a distancia o digital han sido los Learning Management Systems (LMSs), como Moodle, Sakai, dotLRN entre muchos otros (Romero, Ventura y García, 2008). Con la llegada de los Cursos Masivos Abiertos en Línea (MOOCs), hubo una explosión en términos de ser capaces de recolectar grandes cantidades de datos de estudiantes de todo el mundo (Breslow et al., 2013), lo que favoreció la aparición de numerosos estudios en el área e incluso generando sus propios marcos de referencia de analítica de aprendizaje y MOOCs (Drachler y Kalz, 2016). Los proyectos de analítica de aprendizaje también se han llevado a cabo en entornos menos habituales, como tutores inteligentes (Jaques, Conati, Harley y Azevedo, 2014), juegos educativos (Freire et al., 2016) o incluso entornos basados en servicios externos (Ros et al., 2013). Cada uno de estos entornos tienen sus especificidades que hacen que la implementación de estas analíticas pueda ser muy diferente.

Por otra parte, estas herramientas se pueden usar en educación 100% online, o con metodologías híbridas o invertidas (Redondo et al., 2015). Este tipo de herramientas se usan normalmente en educación superior (Viberg et al., 2018), pero también se encuentran casos de estudio en entornos de educación primaria y media (Monroy, Snodgrass Rangel y Whitaker, 2014). Además, también se ha desarrollado investigación en proyectos de analítica de aprendizaje en otro tipo de entornos de aprendizaje, como los no formales (Rezaei y Yaraghtalaie, 2019) o en entornos

profesionales de trabajo (de Laat y Schreurs, 2013). Finalmente, mencionar también que las últimas tendencias están llevando a combinar las analíticas de aprendizaje a través tanto de espacios físicos como digitales (Hsiao, Huang y Murphy, 2017), y combinar datos de múltiples plataformas al mismo tiempo (Mangaroska, Vesin y Giannakos, 2019).

## Recolección de Datos en Crudo

Estos entornos de aprendizaje deben generar datos y guardarlos, para que posteriormente podamos generar la analítica. La forma más común que han seguido la mayoría de entornos virtuales de aprendizaje, es la de guardar las trazas de todos los clics que los usuarios realizan en el entorno como eventos, en lo que se suele conocer en la literatura como *clickstream data*. Esta aproximación no es solo seguida en entornos educativos, sino que es universal en multitud de dominios digitales para modelar el comportamiento humano (Bollen et al., 2009). Esto ha causado que como cada sistema mantiene normalmente su propio esquema de datos, una de las mayores barreras del área ha sido la interoperabilidad entre modelos y soluciones de analítica de aprendizaje (Hoel y Chen, 2014). Hasta la fecha, los dos formatos más extendidos que se han propuesto de cara a alcanzar interoperabilidad en entornos e-learning han sido IMS Caliper (Serrano-Laguna et al., 2017) y xAPI (Berg, Scheffel, Drachsler, Ternier, y Specht, 2016). Hay muchos otros métodos para la generación de datos educacionales más allá de los entornos digitales, como por ejemplo mediante el uso de encuestas (Suchithra, Vaidhehi y Easwaran Iyer, 2015).

En los últimos años también se ha vuelto cada vez más común el uso de sensores en educación, tanto para capturar señales audiovisuales (por ejemplo, audio de conversaciones o vídeo de la clase), como de señales biométricas de los estudiantes (por ejemplo, el latido del corazón o la actividad electrodérmica). Esta área que combina distintos tipos de fuentes de datos se conoce como analítica de aprendizaje multimodal (Blikstein y Worsley, 2016) y es especialmente prometedora para la evaluación de tareas complejas. En algunos de estos casos y dependiendo del contexto, se capturan grandes cantidades de datos y para ejecutar estas analíticas son necesarias arquitecturas de *Big Data* que permitan un almacenaje, acceso y procesado eficiente (Demchenko, Laat, Membreel, y Al., 2014), lo que supone a las universidades un nuevo desafío para la transición tecnológica a nuevos entornos de *Big Data* educacionales (Prinsloo, Archer, Barnes, Chetty y Zyl, 2019).

## Manipulación de Datos e Ingeniería de Características

Una vez se han recolectado los datos, se inicia el proceso de su manipulación. Debido a que estos eventos de bajo nivel representan acciones aisladas de los estudiantes, no son muy informativos en crudo, y por lo tanto es necesario pasar por un proceso que se conoce como ingeniería de características en la que estos

datos se procesan para transformarlos en información educacional útil y sobre la que se pueda actuar. Por ejemplo, estos entornos virtuales guardarán las entradas y salidas al sistema de cada estudiante, pero esta información no será muy útil hasta que algorítmicamente no calculemos el tiempo total que ha pasado el usuario interactuando activamente con la plataforma de aprendizaje. Este proceso requiere de alto nivel de conocimientos técnicos para la manipulación de los datos, experiencia en el contexto, de manera que se conozcan qué características son útiles y creatividad (Ramasubramanian y Singh, 2017). Además, suele ser una de las etapas que mayor esfuerzo requiere dentro de los proyectos de análisis de datos.

Pocos trabajos han tratado este tema específico en el caso de proyectos de analítica de aprendizaje, pero Veeramachaneni y sus colegas han analizado el caso específico de ingeniería de características en entornos de MOOCs (Veeramachaneni, Reilly y Taylor, 2014), así como Paquette y sus colegas lo han hecho en entornos de tutores inteligentes (Paquette, Carvalho, Baker y Ocumpaugh, 2014). Se han reportado algunas herramientas abiertas y gratuitas en la literatura que realizan este proceso de manipulación de datos e ingeniería de características en entornos de MOOCs, como el entorno de trabajo de *edx2bigquery* (Lopez, Seaton, Ang, Tingley y Chuang, 2017) o la herramienta de visualizaciones para Open edX ANALYSE (Ruipérez-Valiente, Muñoz-Merino, Gascón-Pinedo y Delgado Kloos, 2016). Existen diversas iniciativas en otros campos de uso intensivo de datos como la visión por computador o bioinformática para la generación automática de características (Kaul, Maheshwary y Pudi, 2017), pero no parece una dirección muy prometedora en el campo de analítica de aprendizaje debido a la alta complejidad, bajas posibilidades de generalización y los numerosos factores influyentes en el área.

## Análisis y Modelado

El análisis y los modelos que se implementan sobre las características finales que se tienen es clave para comprender nuestros datos y sacar un beneficio educacional de ellos. Muchos de los estudios en torno a analítica buscan entender las interacciones de los estudiantes con el entorno de aprendizaje de forma retrospectiva. Tal como dijo Steve Jobs, “No puedes conectar los puntos mirando hacia adelante; solo puedes conectarlos mirando hacia atrás” (Jobs, 2005); en estas ocasiones el investigador suele buscar la historia que dicen los datos mediante distintos análisis y sin una idea clara de inicio, y cuando la encuentra la puede contar una vez entiende todo lo que ha sucedido.

En otras ocasiones los proyectos tienen objetivos y aplicaciones futuras claras, siendo por lo tanto los métodos a aplicar más obvios y directos. Describimos ahora una serie de tipos de algoritmos que son los más comúnmente utilizados en este tipo de estudios con una aplicación clara: Los casos que los investigadores buscan generar predicciones de resultados de aprendizaje futuros, se aplicarán algoritmos de aprendizaje supervisado que sean capaz de modelar el futuro en función de



conjuntos de datos históricos, como por ejemplo predecir si van a conseguir un certificado en MOOCs (Ruipérez-Valiente, Cobos, Muñoz-Merino, Andújar y Delgado Kloos, 2017). La adaptación de contenidos o recomendaciones personalizadas al usuario, por ejemplo en función de la dificultad y nivel de habilidad actual de usuario mediante la teoría de respuesta al ítem (Chen, Lee, y Chen, 2005). Algoritmos para modelar el comportamiento humano en base a heurísticas y modelos de reglas, por ejemplo, para analizar el interés por los elementos de ludificación ( Ruipérez-Valiente, Muñoz-Merino y Delgado Kloos, 2017). Las conexiones entre estudiantes, o de estudiantes con contenidos, para ver cómo han conectado con los elementos del curso, como por ejemplo mediante el análisis del discurso en foros aplicando técnicas de análisis de redes conectadas y grafos (Milligan, Littlejohn y Margaryan, 2013). Como último ejemplo, el análisis de causas, donde frecuentemente se realiza experimentación y comparan los grupos de control y tratamiento (Kizilcec y Cohen, 2017), y en algunas ocasiones se aplican algoritmos específicos para realizar un análisis causal (Koedinger, Kim, Jia, McLaughlin y Bier, 2015). Estos son algunos de los ejemplos de los métodos de análisis y modelado más comunes que se suelen usar en el área de analítica de aprendizaje, pero otros se pueden aplicar sin problema igualmente. Por último, destacar que el proceso de manipulación de datos e ingeniería de características, en conjunto con esta fase de análisis y modelado, es un proceso iterativo que se puede repetir hasta que se alcancen los resultados deseados.

## **Aplicación Educativa**

Los análisis de la fase anterior normalmente suelen tener asociada una aplicación educativa, pero la realidad es que en numerosas ocasiones no se suele llevar dicha aplicación a la práctica y, la investigación suele terminar en la fase de análisis y modelado. Esto hace que la mayoría de la investigación no se transfiera a la práctica y por lo tanto no se pueda evaluar si realmente tiene un impacto positivo educativo o no; esta es una de las principales problemáticas del área de analítica de aprendizaje en la actualidad (Viberg et al., 2018). Existe una tensión entre la investigación que se hace y cómo transferirla a casos de estudio educacionales reales.

Las aplicaciones más típicas que se ven en proyectos de analítica de aprendizaje son: Interfaces de visualización, que normalmente son usadas por instructores para monitorizar como están avanzando sus alumnos pero también pueden ser accedidas por estudiantes para reflexionar sobre su propio proceso de aprendizaje, como por ejemplo ALAS-KA para Khan Academy (Ruipérez-Valiente, Muñoz-Merino, Leony y Delgado Kloos, 2015), ANALYSE para Open edX (Ruipérez-Valiente, Muñoz-Merino, Gascon-Pinedo y Kloos, 2016) o LAPA para el LMS Moodle (Park y Jo, 2015). También son comunes las aplicaciones que introducen módulos o sistemas completamente adaptativos en la que los contenidos que se muestran van cambiando en función de cómo el usuario se comporta con el sistema y aprende nuevos contenidos (Shute y Zapata-Rivera, 2012). Los sistemas de recomendación también han sido comunes,

envían recomendaciones personalizadas al estudiante sobre qué podría interesarle, por ejemplo sobre contenidos o hilos de discusión en un foro (Wang y Wu, 2011). Por último, también es común que la aplicación final del análisis sea simplemente generar un informe a alto nivel sobre un contexto educativo de una institución, curso o materia, lo que puede ayudar a entender mejor el contexto y adaptar políticas educacionales (Reich y Ruipérez-Valiente, 2019).

Estas aplicaciones o análisis final deberían generar una retroalimentación en los contextos educativos donde se generaron los datos que permita mejorar el proceso de aprendizaje, y con ello se cerraría el ciclo de la implementación de analíticas de aprendizaje. Además, el efecto de estos cambios debería ser evaluado, lo que podría realizarse de nuevo mediante una metodología de analítica de aprendizaje. Esta evaluación es imprescindible para ser capaces de medir el impacto de los cambios introducidos en el contexto educativo.

## **Tecnología, ciencias del aprendizaje, privacidad, instituciones y políticas educacionales**

El trabajo presentado en este artículo se centra en el proceso para la implementación de analíticas de aprendizaje, pero hay otros numerosos factores transversales a tener en cuenta durante este proceso que son de alta importancia para alcanzar el éxito. En primer lugar, resaltar que el auge de la investigación en analítica de aprendizaje ha venido dada por la introducción de la tecnología en educación y que conforme esta se siga introduciendo más en el tejido educacional (Thomas, 2016), podríamos esperar una mayor demanda de estas analíticas y a su vez mayor facilidad en ciertas partes de la implementación de este proceso.

En segundo lugar, resaltar la necesidad de asentar los proyectos de analítica de aprendizaje en aplicaciones educacionales reales que puedan mejorar el aprendizaje (Gašević et al., 2015) y en la teoría sobre las ciencias de aprendizaje y pedagogía desarrollada con anterioridad (Friend Wise y Williamson Schaffer, 2015), sino se corre el riesgo de implementar tecnología y analíticas totalmente desconectadas de las mejores prácticas pedagógicas y teorías educativas desarrolladas en las últimas décadas.

En tercer lugar, la privacidad de los estudiantes y los profesores es fundamental para un desarrollo ético de esta tecnología que no genere rechazo, problemática que al mismo tiempo ya está siendo modelada por numerosas políticas (Drachler, Hendrik y Greller, 2016). Una pregunta crítica es si la educación quiere moverse en la misma dirección que las grandes empresas de Internet, las cuales realizan una continua vigilancia de sus usuarios (Slade y Prinsloo, 2013). Para ello hay que tener en cuenta que los estudiantes deben consentir el que sus datos se usen para estos fines, y tener la libertad de elegir qué datos se pueden usar o si desean retirar ese consentimiento en cualquier momento del proceso; todo esto requiere una orquestación importante

entre los actores de este proceso, no solo estudiantes, sino también profesores, diseñadores y administradores de sistemas (Pardo y Siemens, 2014).

Por último, destacar la necesidad de que para la verdadera implementación y sistematización de analíticas de aprendizaje en educación es crítico alcanzar la involucración de las instituciones educativas y de la generación de políticas educacionales nacionales, y esto genera problemáticas que inicialmente podrían no ser obvias (Macarini et al., 2019). Diversos trabajos han analizado las políticas que instituciones están promoviendo para desarrollar capacidades y competencias para implementar analíticas de aprendizaje de forma sistemática (Arnold et al., 2014), así como para resaltar cuáles son los principales desafíos que las instituciones afrontan (Tsai y Gasevic, 2017), como por ejemplo los programas de reciclaje profesional e incentivos hacia educadores para conseguir que empiecen a adquirir las habilidades necesarias para ser capaces de entender y usar en sus clases analíticas de aprendizaje (Wolff, Moore, Zdrahal, Hlosta y Kuzilek, 2016).

## CONCLUSIONES

En este trabajo se ha presentado una visión del proceso para la implementación de analíticas de aprendizaje que esperamos pueda resultar útil para investigadores del área, analistas de datos educacionales, así como educadores que busquen introducirse con una lectura orientada a la práctica de forma sencilla. A pesar del despegue en popularidad a nivel de investigación, así como de opinión pública del área de analítica de aprendizaje, la transferencia que está teniendo a la práctica está siendo muy baja. Se espera que este texto pueda ayudar a mejorar la transferencia proveyendo de un marco práctico y pragmático para estos proyectos que puede facilitar su organización y desarrollo.

En los próximos años, el campo y las instituciones educacionales afrontan el desafío de empezar a implementar analíticas de aprendizaje a gran escala y de forma sistemática para finalmente poder alcanzar el elevado potencial que todos los investigadores confluyen que tiene. Para ello, hay numerosas barreras en la que toda la comunidad de actores que intervienen en la educación, deberá contribuir a derrumbar. Se deberán desarrollar políticas que permitan una correcta implementación con impacto educacional, sin perder de vista la importancia de la ética y seguridad en el uso de estos datos, preservando la privacidad a la cual tienen derecho los estudiantes. Estos proyectos, deberán estar centrados en los usuarios, empoderándolos y poniéndolos en el centro del desarrollo para lograr implementar aplicaciones que se les pueda dar un uso sostenible a lo largo del tiempo. Otro objetivo será la re-educación de los usuarios, para que en el caso de los profesores aprendan a incorporarlas en sus dinámicas educacionales y en el caso de los alumnos como parte de las herramientas para soportar la auto-regulación de su aprendizaje.

Para ello, se recomienda que las instituciones en los próximos años se centren en implementar algunas de las aplicaciones que han sido desarrolladas de forma

más extensiva y que son sencillas de implementar de forma sistemática, de las cuales se han desarrollado guías para su implementación adecuada, como interfaces de visualización o evaluación de contenidos mediante analíticas. Estas aplicaciones de analítica de aprendizaje pueden sentar las bases de otros de los factores transversales que hemos discutido para empezar a implementarlas a nivel institucional y nacional. En un futuro más lejano, al área le esperan otros desafíos a nivel de investigación, como la transferencia de modelos de aprendizaje máquina cuando se cambia de contextos e incluso plataformas o cómo facilitar la interpretabilidad de todas estas analíticas para que sean más accesibles a aquellos usuarios que tengan competencias limitadas en análisis de datos. El área también pasa por un periodo crítico para empezar a desarrollar ciencia educativa más abierta (Zee y Reich, 2018), en la que las hipótesis se pre-registren, los datos y análisis sean abiertos y los artículos publicados gratuitos para cualquier investigador y educador; sin una ciencia abierta, el avance y el impacto se ven resquebrajados. Futuras extensiones de este trabajo pueden poner a prueba este proceso trabajando con investigadores o instituciones desarrollando este tipo de proyectos siguiendo estas directrices.

La mayoría de los sectores empresariales llevan una gran ventaja a las instituciones educativas en cómo se han re-inventado utilizando los datos que provienen de sus usuarios para mejorar sus negocios. Habiendo sido las instituciones educativas tan conservadoras de forma histórica, alcanzar el verdadero potencial de las analíticas de aprendizaje requerirá de estrecha colaboración y conversación entre todos los actores involucrados en su desarrollo: administradores de instituciones, políticos, educadores, estudiantes, ingenieros e investigadores deberán sentar unas bases que permitan la implementación de analítica de aprendizaje de forma sistemática y productiva.

## NOTAS

1. [https://scholar.google.ca/citations?view\\_op=top\\_venues&hl=en&vq=soc\\_educationaltechnology](https://scholar.google.ca/citations?view_op=top_venues&hl=en&vq=soc_educationaltechnology)
2. <https://tekri.athabascau.ca/analytics/>

## REFERENCIAS

- Aldowah, H., Al-Samarraie, H., y Fauzy, W. M. (2019). Educational data mining and learning analytics for 21st century higher education: A review and synthesis. *Telematics and Informatics*, 37(January), 13-49. <https://doi.org/10.1016/j.tele.2019.01.007>
- Arnold, K. E., Lynch, G., Huston, D., Wong, L., Jorn, L., y Olsen, C. W. (2014). Building institutional capacities and competencies for systemic learning analytics initiatives. In *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge* (pp. 257-260).
- Berg, A., Scheffel, M., Drachsler, H., Ternier, S., y Specht, M. (2016). The dutch xAPI experience. In *Proceedings of the Sixth*

- International Conference on Learning Analytics & Knowledge* (pp. 544-545).
- Blikstein, P., y Worsley, M. (2016). Multimodal learning analytics and education data mining : using computational technologies to measure complex learning tasks. *Journal of Learning Analytics*, 3(2), 220-238. <https://doi.org/http://dx.doi.org/10.18608/jla.2016.32.11>
- Bollen, J., Van de Sompel, H., Hagberg, A., Bettencourt, L., Chute, R., Rodriguez, M. A., y Balakireva, L. (2009). Clickstream data yields high-resolution Maps of science. *PLoS ONE*, 4(3). <https://doi.org/10.1371/journal.pone.0004803>
- Breslow, L., Pritchard, D. E., DeBoer, J., Stump, G. S., Ho, A. D., y Seaton, D. T. (2013). Studying Learning in the Worldwide Classroom: Research into edX's First MOOC. *Research and Practice in Assessment*, 8, 13-25.
- Campbell, J. P., y Oblinger, D. G. (2007). Academic Analytics. *Educause Review*, (October), 1-20. Recuperado de <https://er.educause.edu/articles/2007/7/academic-analytics-a-new-tool-for-a-new-era>
- Chatti, M. A., Dyckhoff, A. L., Schroeder, U., y Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning*, 4(5/6), 318. <https://doi.org/10.1504/IJTEL.2012.051815>
- Chen, C. M., Lee, H. M., y Chen, Y. H. (2005). Personalized e-learning system using Item Response Theory. *Computers and Education*, 44(3), 237-255. <https://doi.org/10.1016/j.compedu.2004.01.006>
- Clow, D. (2012). The learning analytics cycle: closing the loop effectively. In *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge* (pp. 134-138). <https://doi.org/10.1145/2330601.2330636>
- Daniel, J. (2012). Making Sense of MOOCs: Musings in a Maze of Myth, Paradox and Possibility. *Journal of Interactive Media in Education, Perspectiv.* <https://doi.org/10.1145/2316936.2316953>
- de Laat, M., y Schreurs, B. (2013). Visualizing Informal Professional Development Networks: Building a Case for Learning Analytics in the Workplace. *American Behavioral Scientist*, 57(10), 1421-1438. <https://doi.org/10.1177/0002764213479364>
- Demchenko, Y., Laat, C. De, Membreele A., y Al, E. (2014). Defining architecture components of the Big Data Ecosystem. In Collaboration Technologies and Systems (CTS). In *International Conference on IEE* (p. 104:112). Retrieved from <http://www.uazone.org/demch/papers/bddac2014-bd-ecosystem-archi-vo5.pdf>
- Drachsler, H., y Kalz, M. (2016). The MOOC and learning analytics innovation cycle (MOLAC): A reflective summary of ongoing research and its challenges. *Journal of Computer Assisted Learning*, 32(3), 281-290. <https://doi.org/10.1111/jcal.12135>
- Drachsler, Hendrik, y 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). ACM. <https://doi.org/10.1145/2883851.2883893>
- Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5-6), 304-317. <https://doi.org/10.1504/IJTEL.2012.051816>
- Freire, M., Serrano-laguna, Á., Iglesias, B. M., Martínez-ortiz, I., Moreno-ger, P., y Fernández-manjón, B. (2016). Game Learning Analytics: Learning Analytics for Serious Games. In *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy* (pp. 1-29). Springer.

- <https://doi.org/10.1007/978-3-319-17727-4>
- Friend Wise, A., y Williamson Schaffer, D. (2015). Why theory matters more than ever in the age of big data. *Journal of Learning Analytics*, 2(2), 5-13. <https://doi.org/10.18608/jla.2015.22.2>
- Gasevic, D., Dawson, S., Mirriahi, N., y Long, P. D. (2015). Learning Analytics – A Growing Field and Community Engagement. *Journal of Learning Analytics*, 2(1), 1-6. <https://doi.org/10.18608/jla.2015.21.1>
- Gašević, D., Dawson, S., y Siemens, G. (2015). Let's not forget : Learning analytics are about learning. *TechTrends*, 59(1), 64-71. <https://doi.org/10.1007/s11528-014-0822-x>
- Greller, W., y Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology and Society*, 15(3), 42-57.
- Hoel, T., y Chen, W. (2014). Learning Analytics Interoperability - Looking for low-hanging fruits. In *Workshop Proceedings of the 22nd International Conference on Computers in Education, ICCE 2014* (pp. 253-263).
- Hsiao, I.-H., Huang, P.-K., y Murphy, H. (2017). Integrating programming learning analytics across physical and digital space. *IEEE Transactions on Emerging Topics in Computing*.
- Jaques, N., Conati, C., Harley, J., y Azevedo, R. (2014). Predicting Affect from Gaze Data during Interaction with an Intelligent Tutoring System. In *Proceedings 12th International Conference, Intelligent Tutoring Systems 2014, Honolulu, HI, USA*. (pp. 29-38). Springer International Publishing. [https://doi.org/10.1007/978-3-319-07221-0\\_4](https://doi.org/10.1007/978-3-319-07221-0_4)
- Jobs, S. (2005). You've got to find what you love. *Stanford News*. Recuperado de <http://news.stanford.edu/news/2005/june15/jobs-061505.html>
- Kaul, A., Maheshwary, S., y Pudi, V. (2017). Autolearn—Automated feature generation and selection. In *2017 IEEE International Conference on Data Mining (ICDM)* (pp. 217-226).
- Kizilcec, R. F., y Cohen, G. L. (2017). Eight-minute self-regulation intervention raises educational attainment at scale in individualist but not collectivist cultures. *Proceedings of the National Academy of Sciences*, 114(17), 4348-4353. <https://doi.org/10.1073/pnas.1611898114>
- Koedinger, K. R., Kim, J., Jia, J. Z., McLaughlin, E. A., y Bier, N. L. (2015). Learning is not a spectator sport: doing is better than watching for learning from a MOOC. In *Proceedings of the Second ACM Conference on Learning@Scale* (pp. 111-120). <https://doi.org/10.1145/2724660.2724681>
- Leitner, P., Khalil, M., y Ebner, M. (2017). Learning Analytics: Fundamentals, Applications, and Trends. *Learning Analytics: Fundamentals, Applications, and Trends, Studies in Systems, Decision and Control*, 94(February), 1-23. <https://doi.org/10.1007/978-3-319-52977-6>
- Lopez, G., Seaton, D. T., Ang, A., Tingley, D., y Chuang, I. (2017). Google BigQuery for education: Framework for parsing and analyzing edX MOOC data. In *L@S 2017 - Proceedings of the 4th (2017) ACM Conference on Learning at Scale* (pp. 181-184). <https://doi.org/10.1145/3051457.3053980>
- Macarini, L. A., Ochoa, X., Cechinel, C., Rodés, V., Dos Santos, H. L., Alonso, G. E., ... Díaz, P. (2019). Challenges on implementing learning analytics over countrywide K-12 data. *ACM International Conference Proceeding Series*, 441-445. <https://doi.org/10.1145/3303772.3303819>
- Mangaroska, K., Vesin, B., y Giannakos, M. (2019). Cross-platform analytics: A step towards personalization and adaptation in education. In *9th International Learning Analytics & Knowledge*

- Conference (LAK19) (pp. 71-75). <https://doi.org/10.1145/3303772.3303825>
- Milligan, C., Littlejohn, A., y Margaryan, A. (2013). Patterns of engagement in connectivist MOOCs. *MERLOT Journal of Online Learning and Teaching*, 9(2), 149-159.
- Monroy, C., Snodgrass Rangel, V., y Whitaker, R. (2014). A Strategy for Incorporating Learning Analytics into the Design and Evaluation of a K-12 Science Curriculum. *Journal of Learning Analytics*, 1(2), 94-125. <https://doi.org/10.18608/jla.2014.12.6>
- Paquette, L., Carvalho, A. M. J. de, Baker, R., y Ocuppaugh, J. (2014). Reengineering the Feature Distillation Process : A Case Study in the Detection of Gaming the System. *Proceedings of the 7th International Conference on Educational Data Mining (EDM)*, 284-287.
- Pardo, A., Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438-450. <https://doi.org/10.1111/bjet.12152>
- Park, Y., y Jo, I. (2015). Development of the Learning Analytics Dashboard to Support Students' Learning Performance Learning Analytics Dashboards ( LADs ). *Journal of Universal Computer Science*, 21(1), 110-133.
- Prinsloo, P., Archer, E., Barnes, G., Chetty, Y., y Zyl, D. Van. (2019). Big(ger) Data as Better Data in Open Distance Learning Paul. *International Review of Research in Open and Distributed Learning*, 16(1), 284-306.
- Suchithra, R., Vaidhehi, V., y Easwaran Iyer, N. (2015). Survey of Learning Analytics based on Purpose and Techniques for Improving Student Performance. *International Journal of Computer Applications*, 111(1), 22-26. <https://doi.org/10.5120/19502-1097>
- Ramasubramanian, K., y Singh, A. (2017). Feature Engineering. In *Machine Learning Using R* (pp. 181-217). Springer.
- Redondo, D., Muñoz-Merino, P. J., Ruipérez-Valiente, J. A., Delgado Kloos, C., Pijera Díaz, H. J., y Santofimia Ruiz, J. (2015). Combining Learning Analytics and the Flipped Classroom in a MOOC of maths. In *International Workshop on Applied and Practical Learning Analytics* (pp. 71-79). Recuperado de [http://ceur-ws.org/Vol-1599/9WAPLA\\_2015.pdf](http://ceur-ws.org/Vol-1599/9WAPLA_2015.pdf)
- Reich, J., y Ruipérez-Valiente, J. A. (2019). The MOOC pivot. *Science*, 363(6423), 130-131. <https://doi.org/10.1126/science.aav7958>
- Rezaei, M. S., y Yaraghtalaie, M. (2019). Next learning topic prediction for learner's guidance in informal learning environment. *International Journal of Technology Enhanced Learning*, 11(1), 62-70. <https://doi.org/10.1504/IJTEL.2019.096738>
- Romero, C., Ventura, S., y García, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384. <https://doi.org/10.1016/j.compedu.2007.05.016>
- Ros, S., Hernández, R., Robles-Gomez, A., Caminero, A. C., Tobarra, L., y Ruíz, E. S. (2013). Open service-oriented platforms for personal learning environments. *IEEE Internet Computing*, 17(4), 26-31.
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Gascon-Pinedo, J. A., y Kloos, C. D. (2016). Scaling to Massiveness With ANALYSE: A Learning Analytics Tool for Open edX. *IEEE Transactions on Human-Machine Systems*. <https://doi.org/10.1109/THMS.2016.2630420>
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Leony, D., y Delgado Kloos, C. (2015). ALAS-KA: A learning analytics extension for better understanding the learning process in the Khan Academy platform. *Computers in Human*

- Behavior*, 47. <https://doi.org/10.1016/j.chb.2014.07.002>
- Ruipérez-Valiente, J. A., Cobos, R., Muñoz-Merino, P. J., Andújar, Á., y Delgado Kloos, C. (2017). Early Prediction and Variable Importance of Certificate Accomplishment in a MOOC. In *Fifth European MOOCs Stakeholders Summit*. Madrid, Spain.
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., y Delgado Kloos, C. (2017). Detecting and Clustering Students by their Gamification Behavior with Badges: A Case Study in Engineering Education. *International Journal of Engineering Education*, 33(2-B), 816-830.
- Ruipérez-Valiente, J. A., Muñoz-Merino, P. J., Gascón-Pinedo, J. A., y Delgado Kloos, C. (2016). Scaling to Massiveness with ANALYSE: A Learning Analytics Tool for Open edX. *IEEE Transactions on Human-Machine Systems*, 47(6), 909-914. <https://doi.org/10.1109/THMS.2016.2630420>
- Serrano-Laguna, Á., Martínez-Ortiz, I., Haag, J., Regan, D., Johnson, A., y Fernández-Manjón, B. (2017). Applying standards to systematize learning analytics in serious games. *Computer Standards & Interfaces*, 50, 116-123.
- Shute, V., y Zapata-Rivera, D. (2012). Adaptive educational systems. In *Adaptive Technologies for Training and Education* (pp. 7-27). Recuperado de [http://books.google.com/books?hl=en&lr=&id=mOlPSl6iJa1-C&oi=fnd&pg=PA7&dq=Adaptive+Educational+Systems&ots=igwDCR\\_zkc&sig=Sb-l4gFHsItZupoZpSuUYZRCBKc](http://books.google.com/books?hl=en&lr=&id=mOlPSl6iJa1-C&oi=fnd&pg=PA7&dq=Adaptive+Educational+Systems&ots=igwDCR_zkc&sig=Sb-l4gFHsItZupoZpSuUYZRCBKc)
- Siemens, G., y Baker, R. S. J. d. (2012). Learning Analytics and Educational Data Mining: Towards Communication and Collaboration. *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge - LAK '12*, 252-254. <https://doi.org/10.1145/2330601.2330661>
- Slade, S., y Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, 57(10), 1510-1529. <https://doi.org/10.1177/0002764213479366>
- Thomas, S. (2016). Future Ready Learning. Reimagining the Role of Technology in Education. *Office of Educational Technology, US Department of Education*. ERIC.
- Tsai, Y. S., y Gasevic, D. (2017). Learning analytics in higher education - Challenges and policies: A review of eight learning analytics policies. In *Seventh international learning analytics & knowledge conference* (pp. 233-242). <https://doi.org/10.1145/3027385.3027400>
- Zee, T., y Reich, J. (2018). Open Education Science. *AERA Open*, 4(3), 2332858418787466.
- Veeramachaneni, K., Reilly, U. O., y Taylor, C. (2014). Towards feature engineering at scale for data from massive open online courses. *ArXiv Preprint ArXiv:1407.5238*.
- Viberg, O., Hatakka, M., Bälter, O., y Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89(October 2017), 98-110. <https://doi.org/10.1016/j.chb.2018.07.027>
- Wang, S., y Wu, C. (2011). Application of context-aware and personalized recommendation to implement an adaptive ubiquitous learning system. *Expert Systems with Applications*, 38(9), 10831-10838. <https://doi.org/10.1016/j.eswa.2011.02.083>
- Wolff, A., Moore, J., Zdrahal, Z., Hlosta, M., y Kuzilek, J. (2016). Data literacy for learning analytics. In *Proceedings of the Sixth International Conference on Learning Analytics & Knowledge* (pp. 500-501). ACM. <https://doi.org/10.1145/2883851.2883864>



## PERFIL ACADÉMICO Y PROFESIONAL DEL AUTOR

**José A. Ruipérez-Valiente.** Investigador Juan de la Cierva en el Departamento de la Ingeniería de la Información y las Comunicaciones, de la Facultad de Informática en la Universidad de Murcia (España). Investigador afiliado al MIT Playful Journey Lab. Sus líneas de investigación se centran en el aprendizaje mejorado por tecnología, con un alto grado de foco en la analítica de aprendizaje y minería de datos educacionales. ID: <https://orcid.org/0000-0002-2304-6365>. Web: <http://joseruiperez.me/>  
E-mail: [jruiperez@um.es](mailto:jruiperez@um.es)

Dirección:  
Departamento de la Ingeniería de la Información y las Comunicaciones  
Facultad de Informática – Universidad de Murcia  
Calle Campus Universitario, S/N  
30100 Murcia, España

**Fecha de recepción del artículo:** 03/01/2020  
**Fecha de aceptación del artículo:** 17/02/2020  
**Fecha de aprobación para maquetación:** 07/03/2020