

Educación y Educadores

ISSN: 0123-1294

educacion.educadores@unisabana.edu.co

o

Universidad de La Sabana

Colombia

Rojas-Castro, Pablo

Learning Analytics: Una revisión de la literatura

Educación y Educadores, vol. 20, núm. 1, enero-abril, 2017, pp. 406-127

Universidad de La Sabana

Cundinamarca, Colombia

Disponible en: <http://www.redalyc.org/articulo.oa?id=83449754006>

- ▶ Cómo citar el artículo
- ▶ Número completo
- ▶ Más información del artículo
- ▶ Página de la revista en redalyc.org

Learning Analytics: una revisión de la literatura

Pablo Rojas-Castro

orcid.org/0000-0002-5732-1433

Universidad Metropolitana de Ciencias de la Educación
Pontificia Universidad Católica de Chile (Chile)
parojas14@uc.cl

Resumen

El propósito de este artículo es examinar la literatura reciente acerca del emergente campo de Learning Analytics; describir su evolución o crecimiento; analizar los elementos que lo diferencian de otros en cuanto campo, líneas, madurez, y establecer cómo se relaciona con el aprendizaje. Se revisaron 23 papers resultantes de la aplicación de una metodología que incorporó una serie de criterios de inclusión y exclusión, propuestos para cumplir el propósito. Los resultados señalaron que el principal elemento diferenciador de Learning Analytics en cuanto campo es que se centra en el mejoramiento u optimización del aprendizaje, y para alcanzar esta meta, utiliza distintas técnicas y métodos de variados campos, por lo que surge, además, un importante grupo de trabajos que se preocupa por delimitar el campo a través de modelos o frameworks de Learning Analytics.

Palabras clave

Entornos virtuales de aprendizaje; Learning Analytics; big data educacional; investigación educacional; informática educativa (Fuente: Tesauro de la Unesco).

Learning Analytics: A Literature Review

Abstract

The purpose of this article is to examine recent publications on the emerging field of learning analytics, to describe its evolution or growth, to analyze the elements that differentiate it from other disciplines in terms of field, lines of work and maturity, and to determine how it relates to learning. Twenty-three papers were reviewed, based on the application of a method that incorporated a series of inclusion and exclusion criteria proposed for that purpose. The results indicate the main element differentiating learning analytics in terms of its field is the focus on improving or optimizing learning. It uses different techniques and methods from different fields to accomplish that goal. As a result, an important body of work also has emerged that is concerned with delimiting the field through models or frameworks for learning analytics.

Keywords

Virtual learning environments; learning analytics; “big data” in education; educational research; computer uses in education (Source: Unesco Thesaurus).

Learning Analytics: uma revisão da literatura

Resumo

O propósito deste artigo é examinar a literatura recente sobre o emergente campo de Learning Analytics; descrever sua evolução ou crescimento; analisar os elementos que o diferenciam de outros no que se refere a campo, linhas, maturidade, e estabelecer como se relaciona com a aprendizagem. Revisaram-se 23 papers resultantes da aplicação de uma metodologia que incorporou uma série de critérios de inclusão e exclusão, propostos para cumprir com o propósito. Os resultados indicaram que o principal elemento diferenciador de Learning Analytics quanto a campo se centra no aperfeiçoamento ou na otimização da aprendizagem, e, para atingir essa meta, utiliza diferentes técnicas e métodos de variados campos, razão pela qual surge, além disso, um importante grupo de trabalhos que se preocupa por delimitar o campo por meio de modelos ou frameworks de Learning Analytics.

Palavras-chave

Ambientes virtuais de aprendizagem; big data educacional; informática educativa; Learning Analytics; pesquisa educacional (Fonte: Tesauro da Unesco).

Introducción

A comienzos del siglo XXI, la tecnología web da un giro que facilita la participación directa de los usuarios, con la habilitación de la lecto-escritura, que trajo consigo un amplio abanico de nuevas posibilidades, que incluyó la recolección de nueva información acerca de las actividades de los usuarios (Berners-Lee, Hendler y Lassila, 2001). En esta línea, la incorporación de entornos virtuales de aprendizaje (EVA), o sistemas de gestión de aprendizaje (LMS por su nombre en inglés), en educación superior inició un proceso de penetración importante; datos de Inglaterra señalan que en 1994 un 7% de instituciones de educación superior utilizaba EVA / LMS, un 40 % en el 2001 y sobre el 85 % en el 2003 (Britain y Liber, 2004).

El uso de entornos virtuales de aprendizaje genera una huella digital (registro de entrada, acciones, etc.), que aporta una gran cantidad de información acerca de las actividades de los usuarios (Romero, Ventura y García, 2008), a estos datos también se les conoce como *big data* educacionales o *big data* a secas, que son definidos como un conjunto de datos cuyo tamaño excede la capacidad del *software* tradicional para su captura, gestión y análisis (Mazza y Milani, 2004; Romero *et al.*, 2008).

El surgimiento de la huella digital trajo aparejado una serie de nuevas áreas de estudio, tales como el Analytics o Educational Data Mining (EDM) (Campbell, DeBlois y Oblinger, 2007; Romero *et al.*, 2008), acerca del comportamiento de los estudiantes en estos ambientes o sobre las interacciones que allí acaecen. Como señalan Campbell *et al.* (2007), en un contexto de crecientes presiones para rendición de cuentas en educación superior, especialmente para mejorar resultados de aprendizaje y éxito académico, la aparición del *analytics* viene a ayudar en la satisfacción de estas necesidades. El objeto de estudio de la presente revisión de literatura es el *learning analytics*, un tipo de *analytics*; es importante tener en consideración que en cuanto campo, su diferenciación aún no es visible en la discusión desarrollada en el periodo

anterior al revisado en el presente trabajo (véase la sección método); existe inicialmente una referencia general del campo en los trabajos como *analytics*, sin embargo, y más allá de la identificación genérica, en la práctica se está ejecutando una analítica del proceso de aprendizaje (Norris, 2008), o *learning analytics* (LA).

De esta manera, surgen una serie de experiencias en educación superior, en la implementación de *analytics*, como los casos de Baylor University, University of Alabama, Sinclair Community College, Northern Arizona University, o el célebre caso de la Purdue University y su programa llamado Signals, pionero en la aplicación exitosa de *learning analytics* (Campbell *et al.*, 2007; Norris *et al.*, 2008); estas iniciativas tempranas de *analytics* buscaban predecir qué estudiantes estaban en dificultades, permitiendo a los profesores o tutores personalizar la enseñanza o atender de forma precisa y acotada las necesidades de aprendizaje de los alumnos (Campbell *et al.*, 2007). Otros estudios comprenden las relaciones subyacentes entre las interacciones y el rendimiento académico de los estudiantes (Ramos y Yudko, 2008), o entre las interacciones, los niveles de participación y las tasas de deserción en los cursos en línea (Cocea y Weibelzahl, 2007).

Surgimiento de los paradigmas analíticos: EDM y analytics

El uso de EVA exige a los profesores que constantemente adapten sus cursos (en estructura y contenido), para asegurar comprensión, rendimiento y eficiencia en el aprendizaje de sus estudiantes (Gašević, Jovanović y Devedžić, 2007). Una retroalimentación comprensiva se basa en datos semánticamente relacionados entre sí acerca de la mayoría de los elementos principales de un proceso de aprendizaje, tales como: actividades de aprendizaje (por ejemplo, lectura y discusión), así como contenidos y resultados del aprendizaje, y estudiantes (Jovanović *et al.*, 2007). No obstante, en la mayoría de los EVA hacen falta herramientas de EDM y LA que sean utilizables por los profesores y que soporten la

investigación constante, y las que existen dejan mucho que desear (Hijon y Velázquez, 2006).

En este sentido, con la aparición progresiva de los *big data* educacionales, el campo del EDM emergió gradualmente. Si bien Romero y Ventura (2007) citan solamente dos artículos anteriores al año 2000, ellos sitúan su origen en 1995. En términos generales, el *data mining* es un campo de la computación que aplica una variedad de técnicas sobre las bases de datos (por ejemplo, árbol de decisión, inducción de reglas, análisis bayesiano, etc.), a fin de descubrir y mostrar patrones de datos desconocidos y potencialmente útiles. El EDM es un subconjunto de este campo que está centrado en desarrollar métodos para explorar tipos de datos únicos, provenientes de conjuntos de datos educacionales, utilizando estos métodos para entender mejor a los estudiantes y los ambientes en que ellos aprenden (Romero y Ventura, 2007).

El EDM surge del análisis de los archivos de registros de interacción estudiante-computador que predominó en los primeros años del siglo XX, los métodos de *relationship-mining*, seguidos por los de predicción (Romero y Ventura, 2007). No obstante su foco hacia los datos duros, este siempre tuvo un fuerte énfasis en el aprendizaje y la enseñanza. Zaiane y Luo (2001) señalan que la meta del EDM es hacer que los aprendices sean efectivos. Orientación que contrasta con el uso temprano del término *learning analytics*, para referirse a la inteligencia de negocios sobre el *e-learning* (Mitchell y Costello, 2000).

Allende el desarrollo de EDM, el *analytics* se constituye en un enfoque analítico centrado en el entendimiento y la optimización del aprendizaje (Dawson, 2006a; Dawson, McWilliam y Tan, 2008). Se destacan antecedentes de estudios con características de *analytics*, previos a surgimiento del *e-learning* o de los *big data*. En 1979, el Departamento de Investigación de Encuestas de la Open University del Reino Unido podía reflejar diez años de seguimiento de los progresos de sus miles de estudiantes a dis-

tancia, curso por curso, en varias etapas en el año académico (McIntosh, 1979). Otro caso es el de Tinto (1997), que en el contexto de un estudio acerca de persistencia estudiantil y deserción, recabó una base de datos de veinte años de diversas instituciones y tipos de estudiantes. En cuanto a los EVA, se destaca en los años noventa como pioneros a WebCT y TopClass.

El crecimiento del *analytics* está guiado por la necesidad de utilizar los conocimientos obtenidos del análisis de los datos para efectuar intervenciones con vistas a mejorar el aprendizaje (Campbell *et al.*, 2007), junto con otros factores, tales como abordar la tensión relacionada con la impersonalidad de los sistemas *online* (Mazza y Dimitrova, 2007).

Un hito en el desarrollo del *analytics* fue la integración del análisis de red social (Social Network Analysis, SNA), que puede usarse para investigar y promover conexiones colaborativas y cooperativas entre estudiantes, tutores y recursos, y ayudarles a extender y desarrollar sus capacidades (De Laat, Lally, Lipponen y Simons, 2007; Haythornthwaite, 2006).

Junto con lo anterior, y de manera incipiente, los constructos teóricos propios del campo educativo comienzan a surgir; por ejemplo, Aviv *et al.* (2003) situaron su trabajo dentro del paradigma del constructivismo, donde el conocimiento es una construcción mediada por una negociación social (De Laat *et al.*, 2007). De esta manera, el *analytics* se va configurando como un enfoque analítico centrado en el entendimiento y la optimización del aprendizaje (Dawson, 2006a; Dawson *et al.*, 2008), y sus aplicaciones van incorporando un sistema de análisis de la progresión del alumno con el fin de adaptar continuamente el ambiente (Azevedo *et al.*, 2005; Kalyuga, 2006). Ejemplos de lo anterior se pueden apreciar en el trabajo de Kennedy y Judd (2004), que utilizaron un análisis de conglomerado de las interacciones, cuyo resultado muestra que dentro de una única cohorte existen diferentes tipos de interacciones, señalando que estos grupos

tenían implicaciones tanto para el diseño del programa de aprendizaje como para la naturaleza de lo que los diferentes estudiantes pueden ser capaces de aprender de él. Existen otros estudios (Dawson, 2006b; Dawson *et al.*, 2008; Phillips, Baudains y Van Keulen, 2002), donde se realiza un seguimiento de cómo los estudiantes interactúan y participan con el contenido y con otros estudiantes en actividades como foros de discusión.

Problema y relevancia

No obstante la aparición de un cúmulo de *papers* que se ubican dentro del emergente campo del *analytics* (Campbell *et al.*, 2007; Dawson *et al.*, 2008; De Laat *et al.*, 2007; Norris *et al.*, 2008), la primera aproximación a este modelo se encuentra en el trabajo de Campbell y Oblinger (2007), quienes identifican cinco etapas de análisis: captura, reporte, predicción, acción y refinamiento. Esta aproximación es genérica y deja abierta la necesidad del establecimiento de un modelo o *framework* definitivo que incluya, entre otros elementos, cómo se está conceptualizando el aprendizaje (volviendo al eje inicial sobre el que emerge, *big data* educacional), pudiéndose identificar una brecha en la demarcación y definición del campo, que está a todas luces en pleno desarrollo.

Por otra parte, y en total concordancia con lo anterior, en el desarrollo temprano de los campos de EDM y *analytics* (que como se señaló surgen con el advenimiento de la huella digital), no existe una separación clara entre estos (Dawson *et al.*, 2008), teniendo el *analytics* referencias clave en EDM (Campbell *et al.*, 2007; Romero y Ventura, 2007), pudiéndose establecer solamente que las herramientas de EDM normalmente se diseñan más orientadas hacia el poder y la flexibilidad que hacia la simplicidad de uso, siendo demasiado complejas para los educadores (Romero y Ventura, 2007).

De acuerdo con los antecedentes hasta aquí expuestos, el primer factor que hace relevante la presente revisión de literatura es la necesidad de

aclarar el concepto de *learning analytics*, describir su evolución o crecimiento, líneas y madurez. Otro factor relevante son las consideraciones políticas, donde existe una demanda en diversos países por la medición, evidencia y mejora de rendimiento (Campbell *et al.*, 2007); por ejemplo, Norris *et al.* (2008) reportan que, en Estados Unidos, el Gobierno prepara una millonaria inversión en ese sentido. En esta secuencia, las preguntas que guían la revisión son: ¿cuáles son los elementos constitutivos y diferenciares del campo emergente? ¿Cómo se está entendiendo el aprendizaje en este paradigma analítico?

Propósito

El propósito de la revisión bibliográfica es examinar la literatura reciente acerca del emergente campo de *learning analytics*, describir su evolución o crecimiento, diferenciación de otros en cuanto campo, líneas, madurez, y cómo se relaciona con el aprendizaje. Para efectos del trabajo, la revisión se centra en educación superior, fundamentalmente porque el desarrollo inicial del campo se constriñe a este nivel, en concordancia con lo descrito por Campbell *et al.* (2007) y Norris *et al.* (2008).

Método

Para este objetivo, se consideró la revisión de artículos que describen o delimitan qué es *learning analytics*, y cuáles son sus métodos y herramientas (incluidos factores que afectan el aprendizaje), en el ámbito universitario; este proceso de búsqueda se describe a continuación.

La búsqueda se estructuró en dos fases: la primera de ellas centrada en la indagación de estudios de educación en tres bases de datos *online*: WoS, Scopus y Scielo. Las palabras clave utilizadas fueron *learning*, *learning analytics*, *learning analytics research*, *educational factors*, *Virtual Learning Environment*.

La segunda fase consistió en el escrutinio de las referencias de los artículos resultantes, lo que

permitió descubrir tópicos y nuevos autores específicos relacionados con el objetivo de búsqueda.

Criterios de inclusión exclusión

- Artículos que se vinculan con *learning analytics*.
- Artículos que se relacionen con aprendizaje.
- Artículos que relacionen los factores que inciden en el aprendizaje desde la visión de *learning analytics*.
- Artículos cuya fecha de publicación no superase los cinco años de antigüedad, esto es, desde enero del 2009 hasta el 2014, a causa de obtener una revisión de literatura totalmente actualizada, que dé cuenta de las discusiones, los resultados y las tendencias actuales del campo.
- Estudio cuyo campo fuera únicamente la educación terciaria, descartándose el estamento escolar, como, Ebner y Schön (2013) o Hershkovitz, De Baker, Gobert, Wixon y Pedro (2013).
- *Papers* en revistas con revisión de pares, en virtud de garantizar rigor científico y estándares de calidad en la literatura examinada. En consecuencia, se desacartaron fuentes como capítulos de libros, conferencias o ponencias.
- *Papers* escritos en inglés y español, no obstante, los resultados arrojaron únicamente referencias en inglés.
- No fueron incluidas investigaciones que no consideraran entornos virtuales de aprendizaje, debido a que el propósito de este estudio es examinar *learning analytics* en el contexto de los entornos virtuales de aprendizaje, como la aplicación de herramientas de LA en el uso de *software* educativo o juegos (Giannakos, Jaccheri y Leftheriotis, 2012; Rodriguez-Cerezo, Sarasa-Cabezuelo, Gomez-Albaran y Sierra, 2014).

Resultados

Como resultado, y bajo el tamiz de los criterios de inclusión y exclusión establecidos, el número de artículos que constituyen esta revisión de literatura asciende un total de 23 trabajos; quedaron descartados todos aquellos que no cumplieran lo señalado. El más antiguo data del año 2012 y el más reciente fue publicado en el año 2014. En este sentido, cabe destacar que existe literatura previa al 2012 que no cumple con los criterios de inclusión-exclusión, por ejemplo literatura técnica como el *Horizon report* del año 2011 (Johnson *et al.*, 2011), que aborda y define explícitamente LA. Debido a esto, en la parte final de los resultados se incorpora un acápite llamado “*learning analytics* en literatura técnica”, donde se presentan cinco documentos elaborados por organismos técnicos que abordan LA.

Los artículos fueron organizados en tres grandes categorías, con la finalidad de dotar de sentido la revisión, y apreciar de mejor manera las semejanzas y diferencias de la literatura revisada, a saber: i) Delimitadores del campo, ii) Centrados en el proceso de aprendizaje, iii) Centrados en herramientas LA. En la tabla 1, se presentan los artículos categorizados.¹

Tabla 1. Estudios reportados según organización categorial

Categoría	Papers
1. Delimitadores del campo	Becker (2013), Clow (2013), Chatti, Dyckhoff, Schroeder y Thüs (2012), Dietz-Uhler y Hurn (2013), Dringus (2012), Ferguson (2012), Greller y Drachsler (2012), Ifenthaler y Widanapathirana (2014), Martin y Sherin (2013), Mattingly, Rice y BergeZane (2012), Pardo y Siemens (2014), Picciano (2012), Siemens (2013), Slade y Prinsloo (2013).
2. Centrados en el proceso de aprendizaje	Gewerc, Montero y Lama (2014), Kotsiantis, Tselios, Filippidi y Komis (2013), Macfadyen y Dawson (2012).
3. Centrados en herramientas LA	Agudo-Peregrina, Iglesias-Pradas, Conde-Gonzalez y Hernandez-Garcia (2014), Ali, Hatala, Gasevic y Jovanovic (2012), Lockyer, Heathcote y Dawson (2013), Nistor <i>et al.</i> (2014).

¹ El acápite sobre *learning analytics* en literatura técnica no se considera parte de la categorización de artículos, y se incorpora como complemento en la parte final de los resultados, una vez revisadas las tres categorías.

Delimitadores del campo

Esta categoría está conformada por los *papers* que, de distintas maneras, precisan las fronteras o dan una definición del campo en su conjunto. De esta manera, el elemento que separa y distingue a los trabajos de esta categoría es que, por una parte, se enfocan no en un caso en particular, sino en el campo en su integridad; y por otra, buscan definirlo.

Al interior de la categoría, se pueden distinguir tres subconjuntos o dimensiones. Un grupo de trabajos que proporciona un modelo global de LA, que define sus alcances, métodos, impactos y beneficiarios (Greller y Drachsler, 2012; Martin y Sherin, 2013; Picciano, 2012; Siemens, 2013), otro grupo entrega un marco de referencia, que, entre otras cosas, establece los límites del campo (Becker, 2013; Clow, 2013; Dietz-Uhler y Hurn, 2013; Dringus, 2012; Ferguson, 2012; Ifenthaler y Widanapathirana, 2014; Macfadyen y Dawson, 2012; Mattingly *et al.*, 2012), y finalmente el tercer subconjunto realiza lineamientos éticos inéditos en el campo y marca un crecimiento de este (Pardo y Siemens, 2014; Slade y Prinsloo, 2013).

Modelos de LA

Greller y Drachsler (2012), utilizando una metodología de análisis morfológico general, analizan las discusiones en las comunidades de LA; a través de este método levantan un modelo, que se focaliza en su estructura fundamental. Este modelo posee seis dimensiones: actores, objetivos, datos, instrumentos, restricciones externas y limitaciones internas. De este modelo se puede destacar que los objetivos fundamentales de LA son la reflexión y la predicción. Otro aspecto que merece la pena relevar, y que tiene relación con su vinculación con otros campos afines, es que la dimensión instrumentos de LA, está orientada a la conversión de los *big data* desde su estado original (esto es, desestructurados, complejos, etc.), en información útil. Bajo esta premisa, LA se nutre de instrumentos analíticos desde diversos campos,

tales como EDM, *machine learning* o análisis estadístico clásico. Finalmente, es importante destacar que los autores identifican en la dimensión limitaciones internas, la necesidad de poseer ciertas competencias básicas para interpretar los resultados y determinar las acciones posteriores. Estas competencias de alto nivel, imprescindibles, tienen relación con habilidades evaluativas, interpretativas y críticas.

Un factor común, entre los modelos que presentan Martin y Sherin (2013), Picciano (2012), Siemens (2013), es que delimitan el ciclo vital del proceso de LA, desde un punto de partida común, como es la obtención de los datos, hasta una etapa final que varía de acuerdo con el modelo. La tabla 2 presenta las etapas de los respectivos modelos.

Tabla 2. Modelos de learning analytics

Modelo	Etapas del ciclo LA
Chatti <i>et al.</i> (2012)	<ol style="list-style-type: none"> 1. ¿Qué? (determinar qué datos se recolectarán y analizarán) 2. ¿Para quiénes? (Para cuáles actores se orienta el análisis; estudiantes, profesores o directivos) 3. ¿Para qué? (establecer propósito: monitoreo, predicción, adaptación, personalización, etc.) 4. ¿Cómo? (técnicas, métodos y herramientas de análisis)
Martin y Sherin (2013)	<ol style="list-style-type: none"> 1. Muestreo 2. Captura 3. Reducción 4. Hallazgo de patrones
Picciano (2012)	<ol style="list-style-type: none"> 1. Captura de "transacciones educacionales": conformación del big data de EVA 2. Tres estamentos distintos pueden aportar a la siguiente etapa: i) Sistema de información institucional, ii) <i>big data</i>, iii) experiencias, intuiciones o perspectivas de profesores o consejeros 3. Aplicación de software analítico 4. Cursos de acción (decisión basado en datos)
Siemens (2013)	<ol style="list-style-type: none"> 1. Recolección y adquisición de datos 2. Almacenamiento 3. Limpieza de los datos 4. Integración 5. Análisis 6. Representación y visualización 7. Acción (intervención, optimización, alertas, etc.) 8. Reinicio del proceso (<i>loop</i>)

En la tabla 2, se puede apreciar que el modelo de Martin y Sherin no incorpora de manera explícita la implementación de acciones luego de los análisis (implícitos en el hallazgo de patrones). El modelo propuesto por Picciano (2012), además de estar basado en la literatura, toma en cuenta la revisión de distintas experiencias exitosas de LA, tales como el proyecto PACE de Rio Salado Community College en Arizona, el sistema de advertencia temprana y retención llamado GPS (sistema de desempeño de grado), el proyecto Signals de la Universidad de Purdue, entre otros. Del modelo de Picciano es relevante acotar que la etapa 2 (tabla 2) incluye el aporte de información de tres estamentos (sistemas institucionales, EVA y profesores o consejeros) al proceso de análisis; y como destaca Picciano (2012), para aprovechar las ventajas de los *big data* y LA es casi un requisito que el procesamiento de transacciones sea electrónico en lugar de manual. La instrucción tradicional cara a cara puede apoyar los procesos de toma de decisiones basadas en datos tradicionales; sin embargo, para pasar a aplicaciones LA más extensas y especialmente sensibles al tiempo, es importante que las transacciones de instrucción se recojan a medida que ocurren. Esto sería posible en el caso de un EVA.

A la base del modelo de Chatti *et al.* (2012) se encuentra el ciclo iterativo de LA, que posee básicamente tres etapas: recolección y preprocesamiento, análisis-acción y posprocesamiento. Cabe destacar que la propuesta de Chatti *et al.* (2012), sostenida en cuatro dimensiones: datos y entornos, actores, objetivos y métodos, se basó en una revisión de la literatura, la que, entre otras cosas, indicó que los EVA representan la fuente de datos más utilizada para LA, que los objetivos aplicados con mayor frecuencia son adaptación, seguimiento y análisis, y que las técnicas más frecuentemente utilizadas son la clasificación y la predicción.

El modelo propuesto por Siemens (2013) posee un enfoque sistémico donde, además del proceso analítico en cuanto tal, los recursos de apoyo están justamente sistematizados. De este modelo destaca

que la recolección de información está sujeta al propósito del análisis, que puede incorporar motivaciones netamente investigativas, académicas, institucionales, de gestión, etc. En cuanto al análisis, las técnicas y herramientas son variadas e incluyen análisis de red social, programación neurolinguística, predicción, determinación de riesgo, búsqueda de apoyos o desarrollo de conceptos, entre otros.

Frameworks para LA

El conjunto de trabajos que constituye la sub-categoría *frameworks* para LA se caracteriza por proveer un marco conceptual que delimita el campo desde distintas dimensiones, ya sea desde sus componentes mínimos (Becker, 2013), ya sea desde una estructura conceptual holística y compleja (Ifenthaler y Widanapathirana, 2014). A continuación, en la tabla 3, se muestran los distintos *framework* LA hallados en la literatura.

Tabla 3. Frameworks según focalización

Autor	Foco del framework
Becker (2013)	Interacción
	Elementos interactivos constituyentes de LA: <i>timing, location y population</i> .
Clow (2013)	Ciclo iterativo (aprendices, datos, métrica e intervención)
	Elementos esenciales: modelamiento predictivo, análisis de red social, seguimiento, análisis de contenido y semántico
Dietz-Uhler y Hum (2013)	Objetivos de LA en tres niveles: nacional, institucional y local
	Elementos básicos: dos tipos de datos (EVA-profesor), posee capacidad descriptiva y predictiva, rol de la pedagogía en el análisis de los datos
Dringus (2012)	Requisitos esenciales para LA beneficioso: obtener datos correctos, transparencia, buenos algoritmos analíticos, evaluación y uso responsable, informar al proceso y la práctica
Ferguson (2012)	LA como campo separado
	EDM centrado en el desafío técnico; <i>academic analytics</i> * centrado en el reto económico / político; LA centrado en el reto educativo

* AA de ahora en adelante

Autor	Foco del framework
Ifenthaler y Widanapathirana (2014)	El marco LA integra los datos de: actores, su interacción con la web social y entorno de aprendizaje en línea, requisitos curriculares. Posteriormente viene la toma de decisiones
	El procesamiento y análisis de los datos combinados se lleva a cabo en un repositorio de datos de múltiples capas y retornan hacia los actores
Macfadyen y Dawson (2012)	LA se preocupa de optimizar el aprendizaje. Sin embargo, los números no son suficientes, la falta de influencia de LA a nivel institucional se debe a la ausencia de elementos culturales
Mattingly, Rice y Berge (2012)	Mejorar la enseñanza-aprendizaje
	Las acciones van desde la información inicial hasta la intervención

Como se aprecia en la tabla 3, en los distintos *framework* de LA se encuentra a la base la optimización o mejora del proceso de aprendizaje. En congruencia con lo anterior, la conceptualización de LA implica una revisión u observación del proceso de aprendizaje.

LA es un proceso que consta de distintas etapas de carácter cíclico (Becker, 2013; Clow, 2013; Dringus, 2012; Ifenthaler y Widanapathirana, 2014; Macfadyen y Dawson, 2012; Mattingly *et al.*, 2012). De acuerdo con Bernd Becker (2013), además de analizar el comportamiento de aprendizaje general de los estudiantes, LA permite una perspectiva de patrones de aprendizaje de subconjuntos. Dringus (2012), por su parte, se enfoca en que LA debe propender hacia la evaluación responsable y uso eficaz del aprendizaje en línea.

De los *framework* listados en la tabla 3, destaca por una parte Ferguson (2012), que realiza un esfuerzo por deslindar LA de campos afines; en este sentido, distingue y diferencia a cada enfoque analítico de acuerdo con una pregunta de fondo: EDM que está centrado en el desafío técnico, cuestiona ¿cómo se puede extraer valor de estos grandes conjuntos de datos relacionados con el aprendizaje? Por su parte AA, está focalizado en el reto económico o

político, teniendo como pregunta esencial ¿cómo se pueden mejorar sustancialmente las oportunidades de aprendizaje y los resultados educativos a nivel nacional o internacional? Finalmente, LA se encuentra centrado en el reto educativo ¿cómo se pueden optimizar las oportunidades para el aprendizaje en línea? Si bien reconoce que en el desarrollo de LA han existido solapamientos entre los tres grupos, es necesario reafirmar el foco y objeto final de cada campo, donde no existe superposición. En la misma línea, para Ferguson (2012), el desarrollo de *learning analytics* orientado al análisis de red social es una profundización del campo, y representa un movimiento más allá de la investigación basada en datos hacia la investigación sustentada fuertemente en las ciencias del aprendizaje y que lida, además, con la complejidad del aprendizaje permanente en una variedad de contextos.

Otro *framework* que se destaca es el de Ifenthaler y Widanapathirana (2014), fundamentalmente por lo ambicioso de sus pretensiones indagativas o, dicho de otra forma, por querer abarcar la mayor cantidad de información posible de los estudiantes. De esta manera, los datos de los actores incluyen una gran gama de fuentes e información, tales como características sociodemográficas, preferencias e intereses personales, habilidades demostradas o competencias (por ejemplo, alfabetización informática), rendimiento académico, etc. Por su parte, las interacciones relacionadas con la web social incluyen las preferencias de herramientas de medios sociales (por ejemplo, Twitter o Facebook) y las actividades de redes sociales (por ejemplo, recursos vinculados, amistades, grupos de pares, etc.). A lo anterior se agregan datos de fuera del sistema educativo que se recogen a través de diversas vías (por ejemplo a través de un sistema de biblioteca de la universidad o biblioteca pública). Otros datos físicos pueden incluir sensores y datos de localización de dispositivos móviles (por ejemplo, ubicación de estudio y tiempo), o estados afectivos recogidos a través de pruebas de reactivos (por ejemplo la

motivación, la emoción, la salud, el estrés, compromisos). Por su parte, el ambiente de aprendizaje en línea (EVA o blog de aprendizaje) proporciona datos ricos de los actores (por ejemplo, discusiones, resultados y calificaciones de las tareas, encuestas, etc.). Estos datos pueden ser agregados para producir rastros de datos, como los patrones de navegación y rutas de aprendizaje. En este sentido, está también al alcance rica información semántica en los foros de discusión, ensayos escritos, wikis, o entradas de blog, o las interacciones de los actores (entre pares o tutor-estudiante).

Lineamientos éticos

Esta subcategoría se diferencia de las demás en virtud del delineamiento ético que los trabajos buscan establecer en la aplicación y práctica de LA; en otras palabras, el propósito coincidente de los artículos fue exclusivamente dar cuenta y delinear los principios éticos y de privacidad del campo en su conjunto. A continuación, en la tabla 4, se presentan los principios resultantes de cada trabajo.

Tabla 4. Principios reguladores de learning analytics

Autores	Principios del LA
Pardo y Siemens (2014)	1. Transparencia
	2. Control del estudiante sobre los datos
	3. Derecho de acceso
	4. Rendición de cuentas y evaluación
Slade y Prinsloo (2013)	1. LA como práctica moral
	2. Estudiantes como agentes
	3. Identidad y desempeño del estudiante son constructos dinámicos y temporales
	4. El éxito del estudiante es un fenómeno complejo y multidimensional
	5. Transparencia
	6. Educación superior no puede permitirse la no utilización de los datos

Pardo y Siemens (2014), destacan que en el contexto de la recolección y el uso de información de los usuarios, los asuntos éticos y de privacidad no están debidamente abordados. Como metodología revisa-

san cómo estas dimensiones de ética y privacidad son tratadas en otros contextos, tales como el legislativo comercial y en el ámbito de la salud. Justamente, a la luz de las experiencias en otros campos y en consideración a diversas legislaciones, establecen cuatro principios que delimitan la implementación de LA: i) transparencia: en términos generales, los tres grupos de actores (estudiantes, profesores y directivos) deben tener acceso a la descripción de cómo se lleva a cabo el proceso de análisis y se debe informar del tipo de información que está siendo recopilada, incluida la forma en que se recopila, almacena y procesa; ii) control del estudiante sobre los datos: se relaciona con el principio de transparencia en el sentido de que, para que los estudiantes tengan el control sobre los datos que se recogen, ellos necesitan saber lo que se recogió, cuándo, cómo y cómo se manipula; iii) derecho de acceso: los datos recogidos deben estar bajo un conjunto claramente definido de derechos de acceso y la experiencia indica que, cuando no se observan adecuadamente estos derechos, las consecuencias en la confianza de los usuarios son graves; iv) rendición de cuentas y evaluación: es un principio que cruza toda la actividad de LA, y que se traduce en la robustez del proceso global. Por evaluación se entiende la responsabilidad de la institución para evaluar constantemente, revisar y perfeccionar la recogida de datos, seguridad, transparencia y rendición de cuentas.

El delineamiento propuesto por Slade y Prinsloo (2013) aborda la actividad de LA desde los cimientos éticos, esto es, desde un deber ser holístico que les permite, por ejemplo, establecer que LA es de forma primaria una actividad educativa y moral, que se encuentra al servicio (es decir, con una vocación) del mejor y más exitoso aprendizaje. Bajo esta premisa es que adquieran mayor relevancia y precisión principios tales como: i) LA como práctica moral: LA no solo debe centrarse en aquello que es efectivo, sino también en proporcionar indicadores pertinentes para decidir lo que es apropiado y moralmente necesario. La educación es ante todo una práctica

moral, no una causalidad. Por tanto, el análisis de aprendizaje debe funcionar principalmente como una práctica moral que resulta en la comprensión en lugar de medir; ii) estudiantes como agentes: junto con que los estudiantes den su consentimiento informado respecto a la recopilación, el uso y el almacenamiento de datos, deben colaborar voluntariamente para que las acciones de LA sirvan a su aprendizaje y desarrollo, y no solo para la eficacia de las intervenciones y los perfiles institucionales; iii) la identidad y el desempeño del estudiante son constructos dinámicos y temporales, la identidad es un constructo dinámico, y LA proporciona una fotografía del alumno en un momento y contexto particular; iv) el éxito del estudiante es un fenómeno complejo y multidimensional, el rendimiento académico es un fenómeno no lineal, que deviene la relación entre el estudiante, la institución y los factores sociales más amplios; v) transparencia: las instituciones de educación superior deben ser transparentes con respecto a los fines para los que se utilizarán los datos, las condiciones, quiénes tendrán acceso a los datos y las medidas de protección sobre la identidad de los estudiantes. La educación superior no puede permitirse la no utilización de los datos; ignorar información que podría ayudar activamente a perseguir los objetivos de la institución es una conducta miope al extremo.

Centrados en el proceso de aprendizaje

El elemento diferenciador de esta categoría lo constituyen *papers*, cuyo foco es el proceso de aprendizaje que se sostiene a través de algún EVA, su análisis y comprensión, además de establecer algún tipo de acción que genere mejoras en él (Gewerc *et al.*, 2014; Kotsiantis *et al.*, 2013; Macfadyen y Dawson, 2012).

Gewerc *et al.* (2014) se propusieron describir y comprender lo que sucede cuando los estudiantes usan una red social como contexto para llevar a cabo su aprendizaje, para esto ejecutan un estudio de caso de la asignatura de educación tecnológica del tercer año de Licenciatura en Educación de la

Universidad de Santiago de Compostela ($n=58$ año 2012-2013). El análisis se realizó bajo dos herramientas analíticas.

En primer lugar, se utilizó análisis de redes sociales, con lo que se revisaron las relaciones que se constituyeron a través del curso representadas en un gráfico (se usó Unicet y NetDraw). El procedimiento analítico generó dos tipos de métricas: i) densidad de la red (entendida como proporción de vínculos entre los nodos en el gráfico, todo esto en relación con el total de vínculos posibles), que permite apreciar la intensidad de la colaboración; ii) centralidad de un nodo, que muestra la importancia de este último en la red social como consecuencia de las relaciones que establece con otros nodos. La descripción de la composición de la red de interacciones permitió descubrir cuáles estudiantes tienen menor involucramiento, y cuáles están haciendo un trabajo que ayuda a incrementar la densidad de la red, lo que aporta en la implementación de actividades de apoyo a la enseñanza (andamiajes).

En segundo lugar, y con la finalidad de obtener un más completo y exhaustivo panorama del proceso, se implementó un análisis denominado extracción de información, que permite analizar el contenido de las entradas de los blogs generadas por los estudiantes y extraer de forma automatizada los términos o las palabras que caracterizan cada una de ellas. De esta manera, se efectuó un análisis cualitativo de los blogs y las páginas de los estudiantes; para ello, se preseleccionaron entradas relevantes aplicando la técnica de extracción (con una herramienta llamada ADEGA), que automáticamente entregó los términos que caracterizaron los posteos/entradas de cada estudiante. Se analizaron 474 entradas de blogs, de las que el 89,7% resultó ser una reflexión relevante (un post fue considerado relevante si contenía al menos diez palabras claves definidas por el profesor). Los resultados mostraron una interdependencia entre la enseñanza centrada en el estudiante (donde existe independencia, auto-regulación y colaboración) y los indicadores de LA

resultantes (como densidad de la colaboración, autonomía del trabajo, etc.).

Kotsiantis *et al.* (2013) realizan un estudio de caso ($n=337$ estudiantes de un curso *b-learning*, 8 hombres y 329 mujeres, se recogieron datos de 3 años, con grupos distintos cada año), con un objetivo triple: primero estudiar si, y en qué medida, el rendimiento de aprendizaje de los estudiantes se relaciona con su actividad en un EVA (*moodle*). En segundo lugar, para identificar qué tipos de datos recogidos son los mejores predictores. Con este fin, se utilizaron tanto los datos de interacción como las percepciones de los estudiantes acerca de la utilidad del enfoque adoptado y la facilidad del uso de Moodle. Y, en tercer término, si los conocimientos obtenidos (en el proceso) podrían informar el diseño de un sistema para proporcionar información en tiempo real tanto a estudiantes como a educadores.

Se aplicaron cuatro métodos para explorar los datos: i) visualización de cada distribución de las variables, los resultados señalaron que las variables más importantes fueron ingreso al curso, percepciones sobre la plataforma, acceso a los recursos, acceso a las tareas, visitas al foro y publicaciones en foro; ii) árbol de decisión, para identificar qué variables predicen la reprobación o aprobación de los estudiantes: los resultados muestran que la percepción de los estudiantes acerca de los materiales enviados vía *moodle* se diferencia significativamente de las variables que influyen en los resultados de los estudiantes, donde, por ejemplo, la inspección de tareas en más de 74 ocasiones condujo a pasar el grado, en cambio un bajo valor en la revisión de tareas es un predictor fuerte de fracaso; iii) reglas de asociación de clases: se generaron 26 reglas, la principal conclusión esbozada con este método fue que el fallo en el curso está asociado con la actitud y percepción negativa de los estudiantes hacia *moodle*; y por el contrario, las calificaciones excelentes están asociadas con el uso incrementado; iv) *clustering* (con *k-means* de *weka*): las percepciones más positivas hacia la usabilidad, capacidades y utilidad de *moodle* están

generalmente asociadas con las calificaciones más altas de los estudiantes. Además, los estudiantes cuyo total de acciones en el sistema están cerca de, o por sobre el valor promedio (693), tienden a recibir notas más altas. Esto también es aplicable para la vista de foro, vista de recursos y vista de actividades por parte del usuario. El bajo acceso al glosario diferencia fuertemente estudiantes que no lograron aprobar el curso. Por último, se deriva que los estudiantes con calificaciones excelentes tenían número significativamente más alto de registro de acciones relacionadas con la variable *course_view*. En síntesis, los principales resultados permitieron establecer que el fallo en el curso estuvo asociado con la actitud y percepción negativa de los estudiantes hacia *moodle*; y por el contrario, las calificaciones excelentes estuvieron asociadas con su mayor uso.

Macfadyen y Dawson (2012) presentan un estudio de caso de LA sobre la implementación y uso de un EVA en una universidad —que mantienen en el anonimato pero que al mismo tiempo destacan como muy prestigiosa—. Se recolectaron datos del EVA institucional y se llevó a cabo una observación participante. Esta última fue realizada en un comité consultor sobre tecnologías del aprendizaje, donde fueron observadas las actividades y discusiones colectivas de la comisión durante un periodo aproximado de 18 meses, tiempo durante el cual tuvo la tarea de evaluar el uso actual de las instituciones LMS y otras herramientas, y el desarrollo de una visión, hoja de ruta y un plan para la próxima generación del ambiente tecnológico de aprendizaje de la institución. Por otro lado, los datos de uso del EVA corresponden a un año académico (2009-2010); se utilizó una herramienta analítica (*software*), que permitió obtener reportes personalizados por usuarios, cursos, departamentos y facultades, además de capturar, categorizar y analizar datos de interacción en la red institucional. En este contexto, se obtuvieron datos como tiempo de uso del EVA de acuerdo con el rol (profesor, alumno, etc.), y se determinó aquello que los estudiantes realizaban en el EVA (correo

electrónico, discusión, tareas, agenda, revisar contenido, etc.). De la observación participante se destaca que en las primeras etapas de este proceso se presentó un informe de análisis detallado sobre el estado actual de la aplicación y uso de LMS existentes de la institución (posteriormente, fueron puestos a disposición de la comunidad universitaria actas e informes de la toma de decisiones de reuniones); a su vez, en la revisión de estos documentos, y de la participación en continuos debates de los comités, se observó que, aunque se finalizó el análisis del estado actual, en las etapas finales no hubo ninguna otra referencia o interpretación de los resultados. En síntesis, Macfadyen y Dawson (2012) destacan que una comprensión más detallada de lo que, exactamente, está ocupando el tiempo de los estudiantes en sitios de los cursos apoyados por el EVA proporciona una representación más significativa de cómo se utiliza un EVA, y, por tanto, el grado en que el uso de este último complementa las estrategias pedagógicas eficaces. No obstante, y a la luz de los resultados, el uso dominante del EVA se centra en la entrega de contenido.

Finalmente, es importante señalar que, de acuerdo con Gewerc *et al.* (2014), el proceso analítico debe ir más allá de la herramienta que se utiliza, siendo lo verdaderamente central la interdependencia de la herramienta y la propuesta pedagógica. Otra clave esencial para tener en consideración es que los entornos colaborativos no se generan espontáneamente por la mera existencia de un *software* específico; muy por el contrario, se requieren enfoques y propuestas que generen, sostengan y enriquezcan la colaboración (Gewerc *et al.*, 2014).

Centrados en herramientas LA

El elemento que distingue y diferencia a esta categoría lo constituye aquellos trabajos que están focalizados en la aplicación de alguna herramienta específica de LA; es decir, sin preocuparse por trazar los alcances del campo o enfocarse en todo el proceso de aprendizaje, sino enfocándose en alguna herramienta LA específica (Ali *et al.*, 2012; Dyckhoff, Zielke, Bultmann, Chatti y Schroeder, 2012), o como técnica complementaria de análisis (Lockyer *et al.*, 2013).

En primer lugar, es importante revelar que la tendencia en las herramientas presentadas apunta a la accesibilidad de LA para usuarios no expertos, tendencia que buscaría acercar LA a los docentes. Ali *et al.* (2012) se proponen examinar cómo las mejoras a la herramienta LOCO (emanadas en una primera evaluación) afectan la percepción de los usuarios acerca del valor de esta (que permite hacer LA con visualización de datos). LOCO es en general valorada positivamente, y se destaca que la representación visual de la retroalimentación permitiría a los educadores hacerse fácil y rápidamente de una visión general de los diferentes aspectos del proceso de aprendizaje. Algunos participantes atribuyen expresamente la eficacia de las retroalimentaciones a la forma en que fue representado en la nueva versión de la herramienta.

Dyckhoff *et al.* (2012) buscan presentar un marco, diseño, implementación y evaluación de eLAT (un kit de herramientas de LA), que permite a los profesores explorar y correlacionar el uso de objetos de aprendizaje con características y comportamiento del usuario, así como los resultados de evaluación basados en indicadores gráficos. Los resultados señalan que eLAT ha abordado requisitos de usabilidad, utilidad, interoperabilidad, extensibilidad, reutilización, operación en tiempo real y protección de datos (requisitos trazados en el marco elaborado). Si bien la evaluación mostró que eLAT es utilizable, aún faltan más pruebas para obtener datos más fiables sobre la utilidad pedagógica de este kit de herramientas. En este sentido, no se obtuvieron datos suficientes para responder sobre cuál de las variables puede ser suficientemente significativa a nivel pedagógico. Los indicadores implementados en eLAT son más bien generales y enfocados en la simplicidad y comprensibilidad.

Lockyer *et al.* (2013), en la exploración del diseño de aprendizaje, revisan cómo LA puede apoyar,

justamente, en la implementación de un diseño de aprendizaje. De esta manera, utilizando un marco que denominan “punto de control y análisis de proceso”, y una herramienta de análisis de red social, examinan las conductas y los comportamientos esperados del diseño de aprendizaje; sostienen que la información resultante permite un diseño de aprendizaje para ser evaluado en su contexto, en tiempo real, con datos basados en el comportamiento de cómo los estudiantes están actualmente interactuando dentro del ambiente de aprendizaje. Los resultados muestran que, principalmente, esta clase de analítica de punto de control (instrumental) y de proceso (interpretativa), entrega al profesor indicadores acerca del compromiso del estudiante, lo que puede ser usado tanto durante, como finalizado el curso. Bajo este contexto es que LA puede ayudar también a rediseñar el curso.

Learning analytics en literatura técnica

En esta sección de los resultados, se presentan documentos de literatura gris (esto es, documentos redactados por organismos técnicos), que aparecen citados en la revisión, y que abordan LA. La tabla 5 recoge de manera sintética los puntos principales sobre LA.

De los documentos señalados, el denominador común es que el foco de LA está puesto en mejorar el aprendizaje; el ciclo investigativo posee como finalidad última la optimización del proceso y del aprendizaje mismo. En este contexto, IBM (2011) da cuenta de una visión más global de *analytics*, siendo el logro educacional del estudiante una parte de la estrategia global en una organización educativa, que incluye: i) metas —de carácter institucional— de largo plazo ejecutables

Tabla 5. Síntesis de literatura gris

Documento	Puntos principales
Horizon Report 2011 (Johnson <i>et al.</i> , 2011)	LA: interpretación de un amplio rango de datos producidos y recolectados en favor de los estudiantes para evaluar el progreso académico, predecir rendimiento y detectar potenciales problemas
Analytics for Achievement (IBM, 2011)	Preguntas fundamentales del <i>analytics</i> : ¿cuáles son los principales indicadores y las razones detrás del bajo rendimiento? ¿Qué esfuerzos, inversiones y factores afectan al éxito estudiantil? ¿Cómo la asistencia, participación y disciplina se relacionan con el rendimiento? ¿Cuánto debemos invertir en el desarrollo del profesorado? Etapas del <i>analytics</i> : 1. Medición y seguimiento de los logros estudiantiles 2. Informes de resultado 3. Detectar valores atípicos para la intervención temprana 4. Potencialmente predecir 5. Prevención de deserción 6. Identificar y desarrollar atributos clave de los buenos maestros 7. Analizar las pruebas estandarizadas para igualar el rendimiento 8. Pruebas y planes de estudios en evolución
Horizon Report 2012 (Johnson, Adams y Cummins, 2012)	El propósito de LA es analizar la abundancia de datos disponibles sobre los estudiantes para permitir que los centros universitarios ajusten el aprendizaje de cada alumno a una nueva forma de observar los modelos que generan los datos complejos.
Horizon Report 2013 (Johnson <i>et al.</i> , 2013)	LA es un campo de investigación emergente que aspira a utilizar el análisis de datos para generar información que permita tomar mejores decisiones en cada nivel del sistema educativo. LA emplea datos relacionados con los alumnos para elaborar mejores pedagogías, atender a los alumnos con alto riesgo de abandono y evaluar si los programas diseñados para aumentar la retención han sido eficaces y deben mantenerse.
Horizon Report 2014 (Johnson, Adams Becker, Estrada y Freeman, 2014)	El modelo investigativo de LA utiliza análisis de datos para informar de las decisiones tomadas en cada sector del sistema educativo, generando datos sobre el rendimiento de los estudiantes para entregar un aprendizaje personalizado, pedagogía y prácticas adaptativas, e identificar los problemas de aprendizaje a tiempo para que se puedan resolver (además, tiene la expectativa de dar a legisladores y administradores indicadores de progreso educativo locales, regionales y nacionales, que puedan permitir que los programas e ideas sean medidos y mejorados). Los datos de aprendizaje adaptativo ya están proporcionando conocimientos acerca de la interacción de los estudiantes con cursos y textos <i>online</i> .

y medibles; ii) preocupación por el rendimiento académico, monitorear el progreso del estudiante, del currículo y otras actividades de aprendizaje; iii) objetivos financieros, por ejemplo, alineando los presupuestos y recursos operación; iv) requisitos operativos, reducir la incertidumbre de los datos y aumentar la transparencia. De acuerdo con *Horizon Report 2012* (Johnson *et al.*, 2012), el objetivo de LA es permitir a los profesores y a los centros adecuar las oportunidades educativas a las necesidades y capacidades de cada estudiante prácticamente en tiempo real. Por su parte, *Horizon Report 2013* (Johnson *et al.*, 2013), destaca que para profesores e investigadores LA ha sido crucial a la hora de averiguar cómo interactúan los alumnos con los textos y materiales disponibles por Internet, y a su vez, los alumnos también se están beneficiando con los resultados de LA, pues ha propendido a la creación de *software* para dispositivos móviles y plataformas por Internet que utilizan datos específicos de cada alumno, a fin de crear sistemas de apoyo que se ajusten a las necesidades de aprendizaje. En lo concerniente a *Horizon Report 2014* (Johnson *et al.*, 2014), se releva que desde que el tema apareció por primera vez, el 2011 (Johnson *et al.*, 2011), LA ha captado de forma continua el interés de diversos actores como responsables de políticas educacionales o profesionales del área, generando un interés que desborda muchas veces el ámbito netamente educativo. En síntesis, LA puede ayudar a cambiar la educación desde un estadio estándar (generalizada para todos), hacia una dimensión ágil y flexible, diseñada para satisfacer las necesidades particulares e intereses académicos de los estudiantes.

Por último, es importante señalar que en *Analytics for Achievement* (IBM, 2011) no hay explícita referencia a LA, no obstante lo describe. En este sentido, se puede contar con que la definición de LA suministrada por *Horizon Report 2011* (Johnson *et al.*, 2011), es una de las primeras en diferenciar a LA del *analytics* en general.

Discusión y consideraciones finales

En el periodo anterior a la revisión, esto es, antes del año 2009, iniciativas de LA son identificables bajo el rótulo de *analytics*, *academic analytics* o *action analytics* (Campbell *et al.*, 2007; Campbell y Oblinger, 2007; Norris *et al.*, 2008); recién el año 2011 surgen dos definiciones específicas de LA. La primera y más citada (Agudo-Peregrina *et al.*, 2014; Ali *et al.*, 2012; Clow, 2013; Dietz-Uhler y Hurn, 2013; Kennedy, Ioannou, Zhou, Bailey y O'Leary, 2013; Siemens, 2013), es la definición que surge en la “1st International Conference on Learning Analytics and Knowledge 2011”, donde se define LA como “medición, recopilación, análisis y presentación de datos sobre los estudiantes y sus contextos; para comprender y optimizar el aprendizaje y los entornos en que se produce” (Long, Siemens, Conole y Gašević, 2011). La segunda corresponde a la versión de LA que entrega *Horizon Report 2011*, que define LA como la interpretación de un amplio rango de datos producidos y recolectados en favor de los estudiantes en orden de evaluar el progreso académico, predecir rendimiento, y detectar potenciales problemas (Johnson *et al.*, 2011).

En este orden de ideas, y para despejar las posibles ambigüedades, si se sitúa en una organización temporal el desarrollo de LA, desde sus orígenes como *analytics*, se entiende como la búsqueda de predecir o anticipar las dificultades de los estudiantes, y personalizar los itinerarios de aprendizaje para abordar sus necesidades (Campbell *et al.*, 2007); todo lo anterior, desde la utilización o ligazón con grandes conjuntos de datos, técnicas estadísticas y modelos de predicción, o una especie de minería de datos institucionales (Campbell *et al.*, 2007). Desde este último punto surge, sin lugar a dudas, cierta líquidez de las fronteras entre LA y EDM; no obstante, es importante señalar que el origen, foco y objetivo de LA es propio e inequívoco: mejorar el aprendizaje, y para alcanzar esta meta, es que utiliza distintas técnicas y métodos de variados campos (Greller y Drachsler, 2012; Picciano, 2012; Siemens, 2013).

En la misma línea, se aprecia que, si bien LA incorpora o utiliza herramientas de EDM (Greller y Drachsler, 2012; Picciano, 2012; Siemens, 2013), no se restringe a estas, siendo su set de herramientas analíticas mucho más amplio, con la incorporación, por ejemplo, de análisis de redes sociales (Gewerc *et al.*, 2014), o *software* de análisis cuantitativo y cualitativo (Siemens, 2013), entre otros.

Junto con lo anterior, y como destacan Ifenthaler y Widanapathirana (2014), EDM se refiere al proceso de extracción de datos útiles de una gran colección de conjuntos de datos complejos educativos (Romero *et al.* 2011), buscando desarrollar métodos para el análisis de datos educativos (Clow, 2013), y tiende a centrarse más en los problemas técnicos que en cuestiones pedagógicas (Ferguson 2012). En cambio, LA utiliza información dinámica sobre los alumnos y sus entornos de aprendizaje, la analiza e integra, para el modelado en tiempo real del perfil de los estudiantes, predecir resultados y, finalmente, optimizar el proceso y ambiente de aprendizaje (Ifenthaler y Widanapathirana, 2014). En este sentido, vale la pena reiterar que, si bien tanto EDM como LA procesan grandes datos educativos, solamente LA se centra en la optimización del proceso y ambiente de enseñanza-aprendizaje (Clow, 2013; Ifenthaler y Widanapathirana, 2014), y constituye en definitiva, y antes de todo, una práctica moral y educativa (Slade y Prinsloo, 2013). Es también en esta línea cuando Slade y Prinsloo (2013) definen LA como la recolección, el análisis, el uso y la difusión adecuada de datos procesables, que son generados por los estudiantes con el propósito de crear un apoyo cognitivo adecuado, administrativo y efectivo para los estudiantes.

A su vez, Clow (2013) destaca que es problemático diferenciar un campo tan nuevo como es LA de desarrollos similares, como pueden ser AA; sin embargo, es necesario enfatizar que AA es el uso de Business Intelligence en educación, centrado en un nivel institucional y nacional, más que en los estudiantes y cursos individuales (Clow, 2013; Long y Sie-

mens, 2011), destacándose LA por poseer un ciclo analítico que incorpora una intervención (Clow, 2013).

Luego de revisar la evolución del concepto de LA, y establecer que su objeto esencial es la optimización del aprendizaje (Campbell *et al.*, 2007; Siemens, 2013), además de identificar que las principales iniciativas exitosas de LA apuntaron a la identificación de los estudiantes en riesgo y a facilitar algún tipo de soporte (Campbell *et al.*, 2007; Clow, 2013; Dietz-Uhler y Hurn, 2013; Mattingly *et al.*, 2012), queda abierta la brecha para abordar elementos específicos del aprendizaje como estilos de aprender (Becker, 2013). Ya está establecido que LA permite conocer y optimizar el proceso de aprendizaje; ahora bien, falta investigar con más especificidad sobre el tipo de aprendizaje, qué luces se pueden obtener de él, y cómo señalan las preguntas de salida; asimismo, hace falta investigar la idoneidad de los EVA hacia habilidades o aprendizajes específicos.

Al profundizar en la problemática y focalizar la mirada sobre el aprendizaje en sí, es factible hacer avanzar al campo incorporando variables desde la teoría pedagógica, que permitan formular preguntas de investigación y responder aspectos que se desconocen relacionados con habilidades cognitivas específicas; de tal forma es posible preguntar ¿qué habilidades cognitivas de orden superior se fortalecen o desarrollan con el uso de EVA? El desafío es afinar los análisis, profundizarlos y sacar a la superficie conocimientos acerca de la naturaleza del aprendizaje que es factible de desarrollar en los EVA, de manera tal que permitan seguir optimizando tanto el aprendizaje como el ambiente donde ocurre (Long *et al.*, 2011; Siemens, 2013).

Finalmente, es imprescindible precisar que las limitaciones de la revisión estuvieron determinadas fundamentalmente por la etapa temprana de desarrollo en que se encuentra el campo de LA. En este sentido, es importante señalar que la literatura gris incluida en la parte final de los resultados es una medida tendiente a mitigar los efectos de esta limitación.

Referencias

- Agudo-Peregrina, A. F., Iglesias-Pradas, S., Conde-Gonzalez, M. A. y Hernandez-Garcia, A. (2014). Can we predict success from log data in VLEs? Classification of interactions for learning analytics and their relation with performance in VLE-supported F2F and online learning. *Computers in Human Behavior*, 31, 542-550. doi: 10.1016/j.chb.2013.05.031
- Ali, L., Hatala, M., Gasevic, D. y Jovanovic, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education*, 58(1), 470-489. doi: 10.1016/j.compedu.2011.08.030
- Aviv, R., Erlich, Z., Ravid, G. y Geva, A. (2003). Network analysis of knowledge construction in asynchronous learning networks. *Journal of Asynchronous Learning Networks*, 7(3), 1-23.
- Azevedo, R., Cromley, J. G., Winters, F. I., Moos, D. C. y Greene, J. A. (2005). Adaptive human scaffolding facilitates adolescents' self-regulated learning with hypermedia. *Instructional science*, 33(5-6), 381-412.
- Becker, B. (2013). Learning Analytics: Insights Into the Natural Learning Behavior of Our Students. *Behavioral and Social Sciences Librarian*, 32(1), 63-67. doi: 10.1080/01639269.2013.751804
- Berners-Lee, T., Hendler, J. y Lassila, O. (2001). The semantic web. *Scientific american*, 284(5), 28-37.
- Britain, S. y Liber, O. (2004). A framework for pedagogical evaluation of virtual learning environments. DATOS.
- Campbell, J. P., DeBlois, P. B. y Oblinger, D. G. (2007). Academic analytics: A new tool for a new era. *EDUCAUSE Review*, 42(4), 40.
- Campbell, J. P. y Oblinger, D. (2007). Academic Analytics Educause. DATOS.
- Clow, D. (2013). An overview of learning analytics. *Teaching in Higher Education*, 18(6), 683-695. doi: 10.1080/13562517.2013.827653
- Cocea, M. y Weibelzahl, S. (2007). Cross-system validation of engagement prediction from log files Creating new learning experiences on a global scale. Springer, 14-25.
- 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), 318-331. doi: 10.1504/IJTEL.2012.051815
- Dawson, S. (2006a). The impact of institutional surveillance technologies on student behaviour. *Surveillance & Society*, 4(1/2).
- Dawson, S. (2006b). A study of the relationship between student communication interaction and sense of community. *The Internet and Higher Education*, 9(3), 153-162.
- Dawson, S. P., McWilliam, E. y Tan, J. P.-L. (2008). Teaching smarter: How mining ICT data can inform and improve learning and teaching practice. DATOS.

- De Laat, M., Lally, V., Lipponen, L. y Simons, R. J. (2007). Investigating patterns of interaction in networked learning and computer-supported collaborative learning: A role for Social Network Analysis. *International Journal of Computer-Supported Collaborative Learning*, 2(1), 87-103.
- Dietz-Uhler, B. y Hurn, J. E. (2013). Using learning analytics to predict (and improve) student success: A faculty perspective. *Journal of Interactive Online Learning*, 12(1), 17-26.
- Dringus, L. P. (2012). Learning analytics considered harmful. *Journal of Asynchronous Learning Network*, 16 (3), 87-100.
- Dyckhoff, A. L., Zielke, D., Bultmann, M., Chatti, M. A. y Schroeder, U. (2012). Design and Implementation of a Learning Analytics Toolkit for Teachers. *Educational Technology & Society*, 15(3), 58-76.
- Ebner, M. y Schön, M. (2013). Why learning analytics for primary education matters! *Bulletin of the Technical Committee on Learning Technology*, 15(2), 14-17.
- Ferguson, R. (2012). Learning analytics: Drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, 4(5-6), 304-317. doi: 10.1504/IJTEL.2012.051816
- Gašević, D., Jovanović, J. y Devedžić, V. (2007). Ontology-based annotation of learning object content. *Interactive Learning Environments*, 15(1), 1-26.
- Gewerc, A., Montero, L. y Lama, M. (2014). Collaboration and social networking in higher education. *Comunicar*, 21(42), 55-62.
- Giannakos, M. N., Jaccheri, L. y Leftheriotis, I. (2012). Learning and creativity through tabletops: A learning analytics approach. *Bulletin of the Technical Committee on Learning Technology*, 14(4), 11-13. doi: 10.1007/s00779-012-0525-2
- Greller, W. y Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, 15(3), 42-57.
- Haythornthwaite, C. (2006). Facilitating collaboration in online learning. *Journal of Asynchronous Learning Networks*, 10 (1), 7-24.
- Hershkovitz, A., de Baker, R. S. J., Gobert, J., Wixon, M. y Pedro, M. S. (2013). Discovery With Models: A Case Study on Carelessness in Computer-Based Science Inquiry. *American Behavioral Scientist*, 57(10), 1480-1499. doi: 10.1177/0002764213479365
- Hijon, R. y Velázquez, Á. (2006). *E-learning platforms analysis and development of students tracking functionality*. Paper presented at the World Conference on Educational Multimedia, Hypermedia and Telecommunications.
- IBM, S. G. (2011). *Analytics for Achievement*. Ottawa: IBM.

Ifenthaler, D. y Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning*, 19(1-2), 221-240. doi: 10.1007/s10758-014-9226-4

Johnson, L., Adams Becker, S., Estrada, V. y Freeman, A. (2014). *NMC Horizon Report: 2014 Higher Education Edition*. Austin: The New Media Consortium.

Johnson, L., Adams, S. y Cummins, M. (2012). *Informe Horizon del NMC: Edición para la enseñanza universitaria 2012*. Austin: The New Media Consortium.

Johnson, L., Adams, S., Cummins, M., Estrada, V., Freeman, A. y Ludgate, H. (2013). *The NMC horizon report: 2013 higher education edition*. Austin: The New Media Consortium.

Johnson, L., Smith, R., Willis, H., Levine, A. y Haywood, K. (2011). *The 2011 Horizon Report*. Austin: The New Media Consortium.

Jovanović, J., Gašević, D., Brooks, C., Devedžić, V., Hatala, M., Eap, T. y Richards, G. (2007). Using semantic web technologies to analyze learning content. *IEEE Internet Computing*, 45-53.

Kalyuga, S. (2006). Assessment of learners' organised knowledge structures in adaptive learning environments. *Applied cognitive psychology*, 20(3), 333-342.

Kennedy, G., Ioannou, I., Zhou, Y., Bailey, J. y O'Leary, S. (2013). Mining interactions in immersive learning environments for real-time student feedback. *Australasian Journal of Educational Technology*, 29(2), 172-183.

Kennedy, G. E. y Judd, T. S. (2004). Making sense of audit trail data. *Australian Journal of Educational Technology*, 20(1), 18-32.

Kotsiantis, S., Tselios, N., Filippidi, A. y Komis, V. (2013). Using learning analytics to identify successful learners in a blended learning course. *International Journal of Technology Enhanced Learning*, 5(2), 133-150. doi:10.1504/IJTEL.2013.059088

Lockyer, L., Heathcote, E. y Dawson, S. (2013). Informing Pedagogical Action: Aligning Learning Analytics With Learning Design. *American Behavioral Scientist*, 57(10), 1439-1459. doi:10.1177/0002764213479367

Long, P. y Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *EDUCAUSE Review*.

Long, P., Siemens, G., Conole, G. y Gašević, D. (2011). *Proceedings of the 1st International Conference on Learning Analytics and Knowledge*, Banff, Alberta, Canada.

Macfadyen, L. P. y Dawson, S. (2012). Numbers Are Not Enough. Why e-Learning Analytics Failed to Inform an Institutional Strategic Plan. *Educational Technology & Society*, 15(3), 149-163.

Martin, T. y Sherin, B. (2013). Learning Analytics and Computational Techniques for Detecting and Evaluating Patterns in Learning: An Introduction to the Special Issue. *Journal of the Learning Sciences*, 22(4), 511-520. doi:10.1080/10508406.2013.840466

Mattingly, K. D., Rice, M. C. y Berge, Z. L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management and E-Learning*, 4(3), 236-247.

Mattingly, K. D., Rice, M. C. y Berge Zane, L. (2012). Learning analytics as a tool for closing the assessment loop in higher education. *Knowledge Management & E-Learning: An International Journal (KM&EL)*, 4(3).

Mazza, R. y Dimitrova, V. (2007). CourseVis: A graphical student monitoring tool for supporting instructors in web-based distance courses. *International Journal of Human-Computer Studies*, 65(2), 125-139.

Mazza, R. y Milani, C. (2004). *Gismo: a graphical interactive student monitoring tool for course management systems*. Paper presented at the TEL'04 Technology Enhanced Learning'04 International Conference.

McIntosh, N. E. (1979). Barriers to implementing research in higher education. *Studies in Higher Education*, 4 (1), 77-86.

Mitchell, J. y Costello, S. (2000). *International e-VET market research report: A report on international market research for Australian VET online products and services*. DATOS.

Nistor, N., Baltes, B., Dascalu, M., Mihaila, D., Smeaton, G. y Trausan-Matu, S. (2014). Participation in virtual academic communities of practice under the influence of technology acceptance and community factors. A learning analytics application. *Computers in Human Behavior*, 34, 339-344. doi:10.1016/j.chb.2013.10.051

Norris, D., Baer, L., Leonard, J., Pugliese, L. y Lefrere, P. (2008). Action Analytics: Measuring and Improving Performance that Matters in Higher Education. *EDUCAUSE Review*, 43(1), 42.

Pardo, A. y Siemens, G. (2014). Ethical and privacy principles for learning analytics. *British Journal of Educational Technology*, 45(3), 438-450.

Phillips, R., Baudains, C. y van Keulen, M. (2002). An evaluation of student learning in a web-supported unit on plant diversity. DATOS.

Piccciano, A. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education. *Journal of Asynchronous Learning Networks*, 16(3), 9-20.

Ramos, C. y Yudko, E. (2008). "Hits" (not "Discussion Posts") predict student success in online courses: A double cross-validation study. *Computers & Education*, 50(4), 1174-1182. doi: <http://dx.doi.org/10.1016/j.comedu.2006.11.003>

Rodriguez-Cerezo, D., Sarasa-Cabezuelo, A., Gomez-Albaran, M. y Sierra, J. L. (2014). Serious games in tertiary education: A case study concerning the comprehension of basic concepts in computer language implementation courses. *Computers in Human Behavior*, 31, 558-570. doi:10.1016/j.chb.2013.06.009

Romero, C. y Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33(1), 135-146. doi:10.1016/j.eswa.2006.04.005

Romero, C., Ventura, S. y Garcia, E. (2008). Data mining in course management systems: Moodle case study and tutorial. *Computers & Education*, 51(1), 368-384. doi: 10.1016/j.compedu.2007.05.016

Siemens, G. (2013). Learning Analytics: The Emergence of a Discipline. *American Behavioral Scientist*, 57(10), 1380-1400. doi: 10.1177/0002764213498851

Slade, S. y Prinsloo, P. (2013). Learning Analytics: Ethical Issues and Dilemmas. *American Behavioral Scientist*, 57(10), 1510-1529. doi: 10.1177/0002764213479366

Tinto, V. (1997). Colleges as communities: Taking research on student persistence seriously. *The review of higher education*, 21(2), 167-177.

Zaiane, O. R. y Luo, J. (2001). *Web usage mining for a better web-based learning environment*. Paper presented at the Proceedings of conference on advanced technology for education.

