

Título de la ponencia: **Revisión crítica, política y pedagógica de la Minería de Datos y sistemas Tutoring en educación**

Nombre y Apellido del autor: Federico Ferrero

DNI: 32.462.564

Pertenencia institucional: Centro de Investigaciones María Saleme de Burnichon de la Facultad de Filosofía y Humanidades (UNC) (Conicet)

E-mail: [federicojferrero@gmail.com](mailto:federicojferrero@gmail.com)

Domicilio postal: Pueyrredón 256

Teléfono de contacto: 0351 152435917

Eje temático: Políticas sociales: nuevos y viejos derechos

## Resumen

Esta ponencia aborda el análisis de una tecnología que, actualmente, interroga dinámicas de producción, circulación, transmisión y aplicación del conocimiento en las instituciones educativas de nuestra región. El propósito, entonces, es revisar las bases conceptuales de la Minería de Datos (*Data Mining*) como tecnología y del sistema de tutorías inteligentes (*Tutoring*) como uno de los resultados de su aplicación al campo educativo. Se proponen tres dimensiones de análisis que emergen como principales conflictos: a) el problema de la pertenencia y del destino del uso de los datos, b) el problema del reemplazo laboral docente por los sistemas de tutorías inteligentes y c) el problema de la estandarización de los aprendizajes denunciado por las teorías de tradición socio-constructivista. Asimismo se analizan las alternativas que la Minería de Datos ofrece en educación como recurso metodológico para la administración de los sistemas y para la transmisión de conocimientos desde una perspectiva comprometida con el derecho a la educación y al acceso tecnológico.

La incorporación de nuevas tecnologías en el campo de la educación interroga dinámicas de producción, circulación, transmisión y aplicación del conocimiento en las instituciones. A saber: se constituye un nuevo escenario que integra sistemas de gestión institucional organizados con principios distintos a los de las clases tradicionales. Estos sistemas cubren dos formatos principales. Por un lado, el *E-learning* y *Learning Management System (LMS)* que “trasladan” gran parte de las prácticas educativas tradicionales a entornos virtuales (como por ejemplo las aulas virtuales de la plataforma *Moodle*). Por otro lado, aunque aún menos extendidos, aparecen en escena los Sistemas de Tutorías Inteligentes (*Intelligent Tutoring System - ITS*) con el objetivo de personalizar y adaptar la enseñanza a las necesidades de cada estudiante en particular. En ambos casos, el establecimiento virtual de estos sistemas genera una vasta cantidad de datos sobre cada una de las prácticas que llevan adelante los sujetos que participan en ellos.

El problema es que a partir de estas bases de datos, los análisis y las decisiones que se toman exigen, necesariamente, una reflexión lo más crítica posible que ordene estos fenómenos en contexto y que defienda una perspectiva comprometida con el derecho a la educación y el derecho al acceso tecnológico.

Ahora bien, más allá del caso concreto de aplicación al campo educativo, es importante señalar que primeramente la *big data* fue objeto de análisis en el mundo del comercio electrónico (*e-commerce*) y de los negocios electrónicos (*e-business*). Como es obvio suponer, la finalidad perseguida en ese campo era la mejora de los procesos mercantiles tanto en términos de eficacia como de eficiencia. De manera que considerando estos orígenes, no sorprende que la impronta comercial sea la que haya primado en el establecimiento de los principales criterios de investigación en la minería de datos aplicada a contextos educativos.

Con esta tradición a cuestas, el destino y utilidad de esta *big data* ha generado grandes discusiones sobre todo durante la última década en países europeos y en norteamérica. Como consecuencia de ello se ha producido un incremento del número de investigaciones, publicaciones y encuentros aglutinados

alrededor de dos comunidades principales distinguidas, en principio, por su procedencia profesional: docentes e investigadores con formación e interés pedagógico, y especialistas en computación y programadores que deciden trabajar sobre los datos producidos en contextos educativos. Los primeros han denominado al emergente campo como Analítica del Aprendizaje (*Learning Analytic - LA*), los segundos lo llamaron Minería de Datos en Educación (*Educational Data Mining - EDM*).

Aunque recientemente se han producido algunos acercamientos académicos (Siemens y Baker, 2012), estos dos campos se definen por sus diferencias y cruzan críticas que los deslindan en una versión más blanda y otra más dura. La Sociedad para la Investigación en *LA* lo ha definido como “la medida, colección, análisis y reporte de datos acerca de aprendices y sus contextos, para propósitos de comprensión y optimización del aprendizaje y de los entornos en los cuales ocurre” (Siemens y Gesavic, 2012: 1). Por su parte, la Sociedad Internacional de *EDM* establece que *Educational Data Mining* es una disciplina emergente, vinculada con el “desarrollo de métodos para explorar tipos de datos que provienen de contextos educativos, y para usar esos datos con el objeto de entender mejor a los estudiantes y a los contextos en los que ellos aprenden” (Romero y Ventura, 2010: 601).

Estas definiciones aparentemente similares pueden comenzar a distinguirse en cuanto se reconoce el *EDM* como un subtipo de *Data Mining* y se especifican las clases de métodos que comparten y desarrollan. Esta metodología común supone la “extracción automática de patrones implícitos e interesantes de una gran colección de datos” (Romero y Ventura, 2007: 135). Haremos hincapié aquí en el carácter automático de la obtención de los resultados.

En el caso del *LA*, el foco está dado en informar y “empoderar” a los actores involucrados (administradores, docentes, aprendices) de manera que, en definitiva, las decisiones las terminan tomando los mismos actores. Contrariamente, en el *EDM* el énfasis está puesto en la adaptación automática. Esto quiere decir que las decisiones sobre los sistemas establecidos en contextos educativos están basadas en resultados que provienen de hacer correr modelos cuyos *outputs* son ajenos a toda intervención humana.

Congruente con ello, el tipo de descubrimiento que es priorizado en cada caso también se diferencia. El *LA* se focaliza sobre criterios de evaluación realizada por los mismos actores involucrados, es decir que se aprovecha el juicio humano y el descubrimiento automático es sólo una herramienta más para conseguir el objetivo. Inversamente, en el *EDM* el descubrimiento automático a partir de la *big data* es la clave y el juicio humano es accesorio (Siemens y Baker, 2012).

Ahora bien, esta descripción breve debe ser profundizada a la luz de algunos interrogantes que pongan en evidencia compromisos compartidos por ambas líneas de desarrollo teórico. Con alguna distinción metodológica que arrastra mayor o menor simplificación del fenómeno educativo, ambos enfoques comparten el objetivo general del análisis de la *big data* en educación: conocer, predecir y mejorar la efectividad y la eficiencia del fenómeno educativo que ocurre en los entornos virtuales y en el que están involucrados estudiantes, profesores e instituciones.

Este objetivo surge en torno a algunos diagnósticos críticos del sistema tradicional de educación que son citados por los autores enrolados en estos enfoques (*LA* y *EDM*). Por ejemplo, Romero y Ventura sostienen que la educación tradicional ha sido criticada porque “alienta aprendizajes pasivos, ignora las diferencias individuales y las necesidades de los aprendices y no pone atención a la resolución de problemas, al pensamiento crítico o a otros tipos de habilidades de pensamiento superior” (2007: 137). Al mismo tiempo, trabajos como el de Jhonson, Arago, Shaik y Palma Rivas (2000) ponen en evidencia que el aprendizaje *on line* puede ser tan efectivo como el usual método cara a cara en muchos aspectos, a pesar de la creencia general que defiende la posición contraria. Aún así, no se puede desconocer que las más prestigiosas universidades del mundo se niegan tanto a reducir la cantidad de docentes por alumno como a “virtualizar” completamente sus cursos. De modo que la investigación realizada hasta el momento y los datos disponibles no permiten extraer conclusiones rápidas e irrefutables respecto de la aparente “inutilidad” de los métodos tradicionales en educación (que en algunos textos hasta se la llama “educación *off line*” ¡!).

Como se observará, esta importante discusión suele aparecer simplificada en mucha bibliografía alineada, sobre todo, a los principios del *EDM*. Se trata de posiciones que pueden ser problemáticas por adolecer de una cierta ingenuidad en el análisis. Al respecto y por ejemplo, dicen Greller y Drachler: “Es

esperado que el aprendizaje personalizado tiene el potencial de reducir costos de distribución mientras al mismo tiempo crea más eficiencia en las experiencias de aprendizaje, acelerando el desarrollo de competencias y aumentando la colaboración entre aprendices” (2012: 42). Creemos que esta y otras posiciones afines que ven en el análisis de la *big data* en educación una suerte de panacea para los problemas educativos actuales deben, necesariamente, ser revisadas seria y críticamente.

En este marco, proponemos tres tópicos centrales de discusión: 1) el problema de la pertenencia y del destino del uso de los datos, 2) el problema del reemplazo laboral docente por los sistemas de tutorías inteligentes y 3) el problema de la estandarización de los aprendizajes abordado desde el enfoque de las Teorías del Aprendizaje de línea socio-constructivista.

### **Datos y usos: ¿a quiénes les pertenecen los datos? y ¿para qué van a ser usados?**

El avance de los sistemas virtuales de gestión educativa relocaliza la importancia de la pregunta fundamental por la pertenencia de los datos. Pero esta necesidad de legislación invade no sólo a los sistemas educativos. El celebrado proyecto del “Internet de las cosas” (*Internet of Things – IoT*) es el programa que ha organizado el progreso en la construcción de bases de datos generadas por diferentes objetos del mundo y de diversa complejidad. A saber: desde el lavarropas con conexión *wi-fi* hasta, si se permite, los sistemas virtuales de gestión educativa que nos proponemos estudiar en esta comunicación. El mecanismo es, salvando las distancias, el mismo: a cambio de una suerte de comodidad y control remoto de las acciones, los usuarios de distintos objetos cotidianos conectados a internet van dejando su huella de prácticas en distintas bases de datos. Para seguir con el ejemplo previo: queda registrado desde la hora en la que un sujeto lava la ropa hasta el promedio general de la carrera universitaria elegida, etc. etc. etc. Tanto se ha sofisticado este sistema en las últimas décadas que se han diseñado tecnologías convergentes que son capaces de construir muy diversas bases de datos a partir del manejo de un mismo dispositivo cómodo y móvil: el caso típico es el del teléfono celular.

La pregunta, indefectiblemente, es por la política de control y vigilancia que subyace al uso que se hace de estas bases de datos. Nadie ya dudaría que actualmente, el análisis de estas bases de datos puede dar cuenta con fineza de historiales de vidas, perfiles, rutinas, etc. que, desde luego, al menos despiertan incertidumbre por el destino y uso de esos datos. Al respecto, es importante mencionar las acciones llevadas a cabo en torno al “derecho al olvido digital” especialmente en Europa durante los últimos años.

Es en este marco que debe incluirse la discusión en torno a la pertenencia de los datos y al uso de las bases de datos educativas. Así, la disponibilidad y la publicidad aparecen como grandes desafíos y exigen toma de decisiones particulares dado la vulnerabilidad de los sujetos a quienes les pertenecen los datos.

Pero esta situación que amenaza la autonomía de las instituciones y la libertad de sus actores - puesto que, en definitiva, los establece como objetos de vigilancia y control externo- no es de fácil resolución.

Por un lado, agencias y empresas privadas presionan por tener acceso a los datos y extraer análisis provenientes de entornos educativos. La aplicación de este conocimiento responde a intereses eminentemente financieros: por ejemplo la generación de *rankings* de universidades que ordenan sistemas de financiación, *rankings* de estudiantes de los que dependen las plazas vacantes de estudio, sistemas de predicción del comportamiento de los alumnos, etc.

Por otro lado, las instituciones que sostienen políticas de conocimiento abierto y, a la vez, deciden proteger algunos de los datos que ellas mismas generan para evitar entrar en el comercio académico al que referimos previamente. Este tironeo que se manifiesta entre la convicción de compartir la información que produce la institución y la necesidad de restringir ciertos usos de esos mismos datos, ha estimulado el diseño de propuestas que implican procesos de “anonimización” de la información (entre otras).

Al menos aquí lo importante es la claridad para distinguir quiénes demandan información, sobre quiénes se demanda y si los principales actores involucrados han concedido y aprobado el análisis de sus datos. Es entonces necesario puntualizar al menos tres comentarios. En primer lugar, sostener que la política de “anonimización” y protección de datos debe ser agenciada institucional y no individualmente por razones operativas y de capacidad (sobre este punto volveremos hacia el final). En segundo lugar, reconocer que las prácticas de evaluación en las instituciones no comienzan con los sistemas virtuales.

Esto significa que la evaluación dirigida a unidades académicas, a docentes, a alumnos, etc. preexiste a los nuevos sistemas virtuales de gestión de la educación y que, por ello, se ha valido de otros sistemas de información educativa tradicionales para lograr sus objetivos. En tercer lugar plantear que, desde nuestra perspectiva, los mayores problemas aparecen cuando las instituciones pierden el control sobre sus bases de datos de proveniencia virtual y terminan participando en *rankings* y en investigaciones que las evalúan con indicadores definidos externamente y que, además, desconocen sus dinámicas singularísimas de funcionamiento así como su historia y tradiciones. Este punto de vista se sostiene también en las investigaciones que muestran que las auto-evaluaciones obtienen mejores resultados que los controles externos.

Al respecto Greller y Drachsler (2012) realizan una distinción interesante. Estos autores proponen diferenciar dos clases de objetivos en el LA: reflejar prácticas y predecir comportamientos. A saber: el reflejo de los datos permite la auto-evaluación crítica a partir de la propia base de datos para obtener autoconocimiento, mientras que la predicción de las actividades del aprendiz supone, naturalmente, el pronóstico automático de sus comportamientos y la toma de decisiones en consecuencia. Sobre todo en este último caso -donde se ejerce control y vigilancia externa- se debate el problema ético vinculado con la posibilidad de limitar el potencial del alumno o de poner coto a toda forma de creatividad e innovación que divergiere de la media estadística predefinida. Desde un punto de vista pedagógico, tres fenómenos educativos ligados a las expectativas del profesor sobre el alumno pueden aquí utilizarse como argumentos para realizar una crítica al objetivo predictivo en LA: el “efecto Pigmalión o profecía de auto-cumplimiento”, la “educabilidad” y la homogeneización de los aprendizajes en tanto sólo se admite una forma correcta de aprender según lo establecen los “modelos evolutivos de la normalidad”.

#### **Sistemas Inteligentes de tutorías: ¿el problema laboral del reemplazo docente?**

El proyecto comeniano de escuela establecido en la modernidad ha definido un conjunto de principios que le han dado forma al dispositivo escolar tal como lo conocemos hoy: simultaneidad (áulica y sistémica) con la que ocurren los aprendizajes, universalidad del acceso a la educación, gradualidad en el avance de los contenidos a aprender y ritmo de aprendizaje parejo. Como es evidente, la “educación a distancia” ha modificado dos de los postulados centrales del proyecto moderno: el de la simultaneidad y el del ritmo común de avance. Es decir que los estudiantes no aprenden simultáneamente (al menos no en el mismo espacio físico) ni tampoco a la vez, sino que pueden hacerlo remotamente y según sus propias posibilidades de progreso. No obstante, las figuras del docente y del alumno se mantienen vigentes como roles centrales del trabajo pedagógico.

Lo interesante es advertir que los nuevos desarrollos tecnológicos aplicados a la gestión de lo educativo van más a fondo e interpelan la existencia de los roles clásicos que están involucrados en los procesos educacionales según el paradigma moderno. Analicemos a continuación los principales “formatos” educativos actuales.

<b>Descripción de formatos educativos actuales</b>				
<b>Principios del dispositivo escolar moderno comeniano y actores</b>		<b>Tipos de educación</b>		
		<b>Clases tradicionales</b>	<b><i>E-learning y Learning Managment System (LMS)</i></b>	<b><i>Intelligent tutoring system (ITS)</i></b>
Simultaneidad	Áulica (física)	Si	No	No
	Sistémica	Si	Puede ser	Sin evidencia
Universalidad		Si	Si	Sin evidencia
Gradualidad		Si	Si	Si
Ritmo		Común	Más o menos común	Personalizado
Actor/es principal/es		Docentes y alumnos	Docentes y alumnos	Alumnos

Los sistemas de *E-learning* y *Learning Management System (LMS)* son los que han tenido mejor recepción en las instituciones universitarias de nuestra región. La plataforma de gestión *Moodle* es un ejemplo típico de este tipo de sistemas en el que participan los clásicos roles de profesores (editores) y estudiantes. Supone la ausencia de la simultaneidad áulica (física) aunque se podría decir que, en la mayoría de los casos, se ha establecido una simultaneidad institucional aunque no sistémica. Por esta última razón, al menos en las universidades, se ha concretado la universalidad de su uso. Es importante observar que aunque se mantiene la gradualidad en la progresión de contenidos, lo que cambia es el ritmo de los aprendizajes en la medida que este puede ser gestionado con mayor libertad por parte de los estudiantes.

Veamos ahora brevemente los sistemas inteligentes de tutorías (*ITS*) cuya aplicación y extensión es aún incipiente en las universidades de la región. Al menos en la teoría, son abiertamente independientes de cualquier tipo de simultaneidad física puesto que se trata de sistemas de gestión personalísimos que van transformando su oferta y que recomiendan prácticas al estudiante según el programa educativo elegido por él. Respecto de la simultaneidad sistémica y de la universalidad de su alcance, aún no contamos con evidencia unívoca que muestre el modo con el que los sistemas educativos han legitimado su incorporación. Si bien la gradualidad en principio se mantiene, el ritmo es completamente personalizado y adaptado a las posibilidades de cada estudiante (según conocimientos previos, disponibilidad horaria, etc.). Lo distintivo de estos sistemas es la construcción y puesta a prueba de algoritmos que hacen posible la recomendación al estudiante cuando es oportuno. Por ejemplo, el sistema sugiere textos a leer, ejercicios a hacer, evaluaciones a completar, alarmas a tener en cuenta, etc. Obviamente que el sistema puede dar estos *feedback* porque -en base a modelos de estudiantes diseñados computacionalmente- realiza una evaluación constante de los *inputs* de cada alumno, detecta comportamientos no deseados y predice su *performance* próxima.

Si bien principalmente en los últimos cinco años hemos visto multiplicarse las investigaciones y desarrollos de este tipo de sistemas, no es irrisorio pensar que en el caso de que los *ITS* se perfeccionen y su uso se expanda, emergerá con fuerza el problema del “sentido del profesorado” como cuerpo de profesionales necesarios en la educación. Esto es así dado que mucha de las tareas típicas del docente, incluso el diseño de los mismos cursos, podrían ser reemplazadas por estos sistemas automáticos. Es natural entonces mencionar el evidente problema laboral que existiría en el caso de que este escenario hipotético se afiance en las instituciones.

### **Estandarización de los aprendizajes**

La instauración de un camino legítimo para aprender, es decir la construcción de modelos de aprendices en base a los cuales se examinarán las prácticas concretas de cada estudiante, hace ingresar desde luego, la discusión en torno a la estandarización y homogeneización de los aprendizajes. Desde la perspectiva de la teoría pedagógica sobre los aprendizajes podemos advertir que se trata de un problema de definición suficiente y pertinente de la unidad de análisis (unidad de análisis no en sentido metodológico sino en tanto recorte teórico que se hace de la realidad para abordar un fenómeno determinado). Aunque la *LA* se jacte de que su comprensión de los sistemas de aprendizaje abarca la complejidad con la que estos se manifiestan, creemos que los riesgos que toma no parecen evadir el reduccionismo en el que incurre abiertamente el *EDM*. Al dividir en “componentes” la realidad y analizar las relaciones que se establecen entre ellos, puede que se pierda la especificidad del fenómeno a estudiar. Al respecto, es oportuno mencionar que Vygotsky (1934/2010) ha mostrado este peligro analítico con su famosa analogía del H<sub>2</sub>O.

De modo que la simplificación y reducción del fenómeno del aprendizaje a unos pocos indicadores es un tema de seria discusión desde las Teorías del Aprendizaje. Desde nuestra perspectiva no existe problema en reconocer que un sistema de tutoría inteligente podría ser bien estudiado según los principios teóricos del programa conductista de investigación. Veamos la explicación que daba Skinner sobre las máquinas de enseñar a finales de los 50 del siglo pasado para encontrar una sugerente semejanza con los *ITS*.

“El dispositivo consiste de una caja de aproximadamente el tamaño de un pequeño toca-disco. En la superficie hay una ventanilla de vidrio, en la que puede verse una

pregunta o un problema impreso en una cinta de papel. El niño responde a la pregunta desplazando dos o más regletas deslizante, que llevan impresas las cifras 0 a 9. La respuesta aparece en orificios cuadrados, perforados en el mismo papel en que va impresa la pregunta. Cuando haya emitido su propia respuesta, el niño hará girar un botón. La operación es tan sencilla como el ajuste de un aparato de televisión. Si la respuesta es correcta, el botón girará fácilmente y puede hacerse que suene una campana o se proporcione algún otro reforzamiento condicionado. Si la respuesta es errónea, el botón no girará. Puede añadirse un contador para llevar la cuenta de las respuestas erróneas. A continuación, deberá hacerse girar ligeramente el botón a la inversa, efectuando un segundo intento de respuesta correcta. (A diferencia de las tarjetas impresas, el dispositivo indica que una respuesta es errónea sin dar la correcta). Cuando la respuesta sea correcta, otro giro del botón accionará un embrague que hará que pase a la mirilla el problema siguiente, pero ese movimiento no podrá completarse mientras las regletas no hayan regresado a cero” (Skinner, 1959, “*Cumulative Record*”)

Es evidente que podremos, al menos, establecer algunas comparaciones entre las famosas máquinas de enseñar y los sistemas de tutorías inteligentes. También es indudable que los métodos se han sofisticado pero los principios teóricos que ordenan su funcionamiento no son tan distintos. La pregunta allí es si aproximaciones constructivistas podrían tener algún acceso pertinente al estudio de estas máquinas. Hay que decir que muchos investigadores lo ponen en duda. Lo cierto es que la investigación debe concentrarse en esclarecer esta discusión que tiene vastas derivaciones prácticas y serias implicancias éticas.

### **Conclusiones y discusiones**

Esta breve ponencia ha tenido como propósito una revisión crítica y pedagógica de la minería de datos y especialmente de los sistemas de tutorías inteligentes en educación. Para ello, hemos abordado tres problemas principales: el de la pertenencia y destino del uso de los datos, el del hipotético reemplazo docente por los sistemas de tutorías inteligentes y el de la teoría pedagógica que subyace a la estandarización de los aprendizajes.

Finalmente, nuestra última reflexión apunta al rescate del nivel institucional como espacio de disputa privilegiado para lidiar con los problemas y riesgos que suscita la incorporación de estas nuevas tecnologías. Sostenemos que a nivel de los individuos las posibilidades de resistencia son cada vez menores dadas las estrategias de penetración tecnológica que han logrado incluir a muchos de los ciudadanos en el registro de cada vez más vastas bases de datos (vía *e-mails*, redes sociales, sitios de compra *on line*, etc.).

Esta decisión de preconizar el nivel institucional como instancia privilegiada que entra a tomar decisiones en los debates implica abandonar la tendencia que privilegia el estudio de la Tecnología Educativa como fenómeno típico del nivel de concreción áulico. La apuesta está dirigida a captar la complejidad del fenómeno tecnológico sin reduccionismos didácticos y a abrir el juego para conectar con problemáticas propias de niveles más amplios y para así robustecer la interdisciplinariedad que exige el trabajo con objetos de conocimientos de este calibre.

Al reconocer que estas tecnologías serán temas cada vez más ineludibles en los próximos años, en adelante, el desafío es sostener una mirada crítica y también identificar usos de la minería de datos que sean expresiones de una perspectiva comprometida con el derecho a la educación y al acceso tecnológico...

### **Bibliografía consultada**

Buckingham Shum, S. & Ferguson, R. (2012). Social Learning Analytics. *Educational Technology & Society*, 15 (3), 3-26.

Greller, W. y Drachsler, H. (2012). Translating Learning into numbers: a generic framework for Learning Analytics. *Educational Technology & Society*, 15 (3), 42-57.

Johnson, S., Arago, S., Shaik, N., y Palma-Rivas, N. (2000). Comparative analysis of learner satisfaction and learning outcomes in online and face-to-face learning environments. *Journal of Interactive Learning Research*, 11(1), 29–49.

Siemens, G. y Baker, R. S. (2012). Learning analytics and educational data mining: towards communication and collaboration. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 252-254). ACM.

Siemens, G. y Gesavic, D. (2012). Guest Editorial – Learning and Knowledge Analytics. *Educational Technology & Society*, 15 (3), 1-2.

Skinner, B. F. (1959). *Cumulative record*. New York: Appleton-Century-Crofts.

Romero, C. y Ventura, S. (2007). Educational data mining: A survey from 1995 to 2005. *Expert Systems with Applications*, 33, 135–146.

Romero, C. y Ventura, S. (2010). Educational Data Mining: A Review of the State of the Art. In *Proceedings in IEEE Transactions on Systems, Man, and Cybernetics: Applications and Reviews*, vol. 40, no. 6, 601-618.

Vygotsky, L. (1934/2010). *Pensamiento y lenguaje*. Paidós: Buenos Aires.