



1 de enero de 2016 | Vol. 17 | Núm. 1 | ISSN 1607 - 6079

# ARTÍCULO

## **BIG DATA EN LA EDUCACIÓN**

<http://www.revista.unam.mx/vol.17/num1/art06/>

*Javier Salazar Argonza (responsable de tecnologías de información y comunicación en el área de Acervos Digitales, DGTIC, UNAM)*

## **BIG DATA EN LA EDUCACIÓN**

### **Resumen**

Se realiza un acercamiento a la tecnología de *big data* y su influencia en la educación explicando los principales métodos educativos y las nuevas opciones profesionales derivadas de la analítica que éste ofrece. Se habla sobre los aspectos técnicos, legales y éticos para la adopción y uso de esta tecnología en las instituciones educativas.

**Palabras clave:** *big data*, métodos educativos, TIC, aprendizaje adaptativo, EBC, aula invertida, aprendizaje combinado, gamificación, aprendizaje Móvil, analítica del aprendizaje, Hadoop, SGBD.

### **BIG DATA IN THE EDUCATION**

#### *Abstract*

*In this paper it is given an approach to big data technology and its influence on education explaining the main educational methods and new professional options derived from big data analytics. It discusses the technical, legal and ethical aspects for the adoption and use of big data technology in educational institutions.*

**Keywords:** *big data, educational methods, adaptive learning, EBC, flipped classroom, blended learning, gamification, mobile learning, learning analytics, Hadoop, DBMS.*

## BIG DATA EN LA EDUCACIÓN

### Introducción

El uso de la tecnología de *big data* en la educación ha comenzado ya a brindar grandes beneficios en lo referente a la mejora de la gestión educativa, el desarrollo de nuevos métodos para la enseñanza y el aprendizaje, la creación de nuevas carreras y opciones profesionales para los estudiantes, así como en la explotación y aprovechamiento de acervos digitales generados por las instituciones educativas.



Sección de logo tomado de:  
<http://www.ididactic.com/edblog/big-data-y-educacion-ii/>, Nivard Arboix, enero, 2015.

Es la analítica del *big data* el recurso clave actual para analizar, visualizar, entender y mejorar la educación. Por tal motivo, en este artículo, dirigido a los responsables de guiar el rumbo de ésta, académicos, investigadores, docentes y público en general, se intenta dar una visión del actual grado de avance de esta tecnología en la educación, así como las facilidades que hay para su instrumentación en las instituciones educativas.

La tecnología de *big data* permite recolectar, almacenar y preparar grandes volúmenes de datos para analizar o visualizar la relación entre ellos. Esto inclusive a partir de datos que se estén generando en tiempo real y que provienen de redes sociales, sensores, dispositivos de diversa índole o de fuentes de audio y video. De esta manera, es posible identificar y predecir con detalle las causas o efectos de eventos, sucesos o procesos complejos.

En el ámbito educativo, esta tecnología ciertamente ha comenzado a brindar grandes beneficios. Por ejemplo, en la mejora de la gestión educativa, al desarrollo de nuevos métodos para la enseñanza y el aprendizaje, la creación de nuevas carreras y opciones profesionales, así como en la generación y almacenamiento de acervos digitales que constituyen el producto de años de actividad académica, docente y de investigación.

Actualmente, se ha trascendiendo hacia un estado en el cual se necesitan de nuevos métodos sustentados en la tecnología para poder hacer el seguimiento de los alumnos, mejorar sus tutorías, obtener datos objetivos de sus evaluaciones, predecir los riesgos académicos o simplemente comprender el comportamiento de los grupos escolares. En este continuo cambio los profesores están más conscientes de la necesidad de actualizarse tecnológicamente para poder ofrecer una educación más efectiva y adecuada a las necesidades de la población escolar actual.

El método de observación en el aula ha dejado de ser la forma más efectiva para mejorar el proceso educativo. Por su parte, la analítica del *big data* es el recurso clave actual para entenderlo y mejorarlo. Asimismo, los acervos digitales, generados de manera regular por las instituciones educativas, deben ser integrados en el lago de datos y así mejorar su difusión, explotación y aprovechamiento no sólo en los procesos de analítica de *big data*, sino por toda la sociedad.

## ¿Qué es el *big data*?

Es un concepto que hace referencia al tratamiento y análisis de enormes repositorios de datos tan desproporcionadamente grandes que resulta imposible manejarlos con las herramientas de bases de datos y analíticas convencionales.<sup>1</sup>

De hecho, esta tecnología se ocupa de todas las actividades relacionadas con los sistemas que manipulan grandes conjuntos de datos. Las dificultades más habituales vinculadas a la gestión de estos últimos se centran en la recolección y el almacenamiento, búsqueda, compartición, análisis y visualización de la información. La tendencia a manipular enormes cantidades de datos se debe a la necesidad en muchos casos de emplear dicha información en la creación de informes estadísticos y modelos predictivos que pueden ser utilizados en muchas áreas del quehacer humano.<sup>2</sup>

## Métodos educativos derivados del *big data*

Entre los principales métodos educativos derivados a partir de la analítica del *big data* (cuyos orígenes se remontan al año 2000<sup>3</sup>), y su integración con los nuevos dispositivos inteligentes y tecnología web que hoy en día ya se están aplicando en el ámbito educativo, se tienen:

- Aprendizaje adaptativo
- Educación basada en competencias
- Aula invertida y aprendizaje combinado (flipped classroom y blended learning)
- Gamificación
- Aprendizaje móvil (mobile learning)

Estos métodos, que permiten hacer llegar la educación a una mayor población estudiantil y satisfacer las continuas demandas de los alumnos (cada vez exigen una formación más flexible y personalizada), están brindando resultados muy prometedores y han cambiando la forma en la cual los estudiantes se relacionan entre sí, con sus profesores e instituciones educativas. A continuación se explica de forma breve en que consiste cada uno de ellos.

### • El aprendizaje adaptativo

El *aprendizaje adaptativo* es un método educativo basado en la modificación de los contenidos y formas de enseñanza de acuerdo con las necesidades particulares de cada estudiante. Para ello, el lago de datos de conocimientos, que facilitará la analítica para crear el plan de estudios a la medida, se conformará recolectando información sobre los hábitos de aprendizaje, conocimientos, debilidades y fortalezas de cada estudiante.<sup>4</sup>

Este método hace énfasis en las áreas en las que los estudiantes tienen más dificultades, con el fin de adecuarse a su forma y ritmo de aprendizaje y crear un camino de enseñanza personalizado, diferenciado y adaptado para cada estudiante.



[1] Véase: <http://www.enri-quedans.com/2011/10/big-data-una-pequena-introduccion.html>

[2] Véase [https://es.wikipedia.org/wiki/Big\\_data](https://es.wikipedia.org/wiki/Big_data)

[3] Véase la historia de la analítica del aprendizaje (Learning Analytics): [http://etec.ctlt.ubc.ca/510/wiki/Learning\\_Analytics:\\_An\\_Introduction\\_and\\_Critical\\_Analysis#History](http://etec.ctlt.ubc.ca/510/wiki/Learning_Analytics:_An_Introduction_and_Critical_Analysis#History)

[4] Véase: <http://centrodeinnovaciontecnologia.blogspot.mx/2014/01/que-es-aprendizaje-adaptativo.html>

Hoy en día existen diversas plataformas tecnológicas innovadoras que buscan adaptar el ritmo de estudio de ciertas materias a las necesidades particulares de cada estudiante. Estas plataformas proporcionan un aprendizaje único y totalmente acoplado a las necesidades de los estudiantes.

Un ejemplo es Knewton, cuya tecnología es capaz de identificar las fortalezas y debilidades de cada alumno y, en función de éstas, ofrecerle los materiales más apropiados.<sup>5</sup>



Figura 1. Logotipo de la plataforma de aprendizaje adaptativo, KNEWTON.

### • Educación basada en competencias (EBC)

La educación basada en competencias surge como una concepción formativa para el trabajo. Su propósito está relacionado con la productividad, esto en el sentido de lograr que los egresados de los programas de capacitación se incorporen rápidamente y de forma provechosa a su actividad laboral.<sup>6</sup>

Este método educativo consiste en adaptar el proceso de aprendizaje al ritmo y las necesidades de cada estudiante. De esta manera, el alumno sólo superará una determinada materia cuando consiga dominarla con destreza, independientemente del tiempo que necesite para conseguirlo. En este sentido, el estudiante será el único responsable de fijar su ritmo de aprendizaje, lo cual disminuirá notablemente sus niveles de estrés y elevará su motivación hacia el estudio.

Aunque este procedimiento no es una corriente nueva, las actuales plataformas educativas basadas en competencias, derivadas del poder de la analítica del *big data* bajo la métrica del aprendizaje, sí lo son. De hecho, éstas potencian la independencia de los estudiantes y, al mismo tiempo, les permiten a los profesores evaluar avances. Como ejemplo se encuentra *Mastering*.

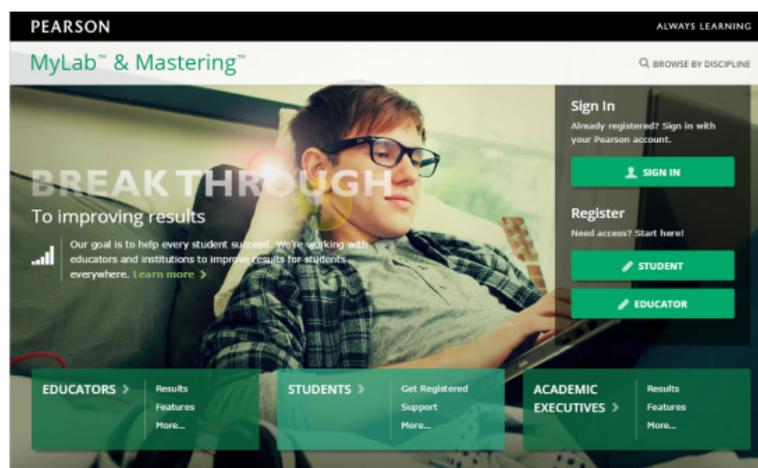


Figura 2. Pantalla de acceso a la plataforma de aprendizaje basada en competencias, Mastering.



[5] Para mayor información véase: <https://edshelf.com/tool/knewton/>

[6] Véase: <http://blogs.monografias.com/para->

- Aula invertida y aprendizaje combinado (Flipped Classroom y Blended Learning)**

En los últimos años, los avances en el desarrollo de las páginas web y la facilidad para realizar contenidos digitales en audio y vídeo contribuyeron al crecimiento de metodologías de enseñanza basadas en el modelo de aula invertida y aprendizaje combinado.

El aula invertida (*flipped classroom*) se basa en el estudio en casa y la práctica en clase. En este sentido, las nuevas tecnologías de información permiten al profesor impartir sus contenidos en línea reservando el tiempo de clase para una formación más práctica.<sup>7</sup>

El aprendizaje combinado o *blended learning* también se basa en la combinación de la formación en línea y presencial, pero de un modo distinto: algunos contenidos se imparten en clases presenciales, mientras que otros se desarrollan exclusivamente desde Internet.<sup>8</sup>

Una plataforma típica que utiliza estos métodos educativos es Moodle. Ésta es utilizada en las principales universidades del mundo y permite a los estudiantes adquirir conocimientos a distancia, mismos que luego compartirán con el profesor y sus compañeros a través de un aula virtual.<sup>9</sup>



Figura 3. Pantalla del portal de distribución de la plataforma de aprendizaje basada en el aula invertida y aprendizaje combinado, Moodle.



[7] Véase: <http://profesor3punto0.blogspot.mx/2013/07/que-es-eso-de-la-flipped-classroom-para.html>

[8] Véase: <http://cantelimitioblog.blogspot.mx/2010/02/la-aparicion-de-las-nuevas-tecnologias.html>

[9] Véase: [https://moodle.org/?lang=es\\_mx](https://moodle.org/?lang=es_mx)

[10] Véase: <http://www.cae.net/gamificacion-cae/>

Cabe destacar que las bases de conocimientos que generan este tipo de plataformas sirven para complementar los lagos de datos de una institución educativa y, por ende, facilitar la analítica del aprendizaje.

- Gamificación**

Se puede definir a la gamificación como el empleo de mecánicas de juego en entornos de aprendizaje con el fin de potenciar la motivación, la concentración y el esfuerzo de los estudiantes. Asimismo, constituye una poderosa estrategia para influir y animar a grupos de alumnos. Su intención es estimular el proceso de enseñanza–aprendizaje entre los miembros de una comunidad estudiantil.<sup>10</sup>

Una correcta implementación de la gamificación permite obtener un compromiso, consiguiendo que los estudiantes de una escuela participen de manera dinámica y proactiva en acciones que generalmente requieren un esfuerzo de voluntad.

La integración de dinámicas de juego en entornos de aprendizaje no es algo nuevo. De hecho, el crecimiento exponencial del uso de los videojuegos despertó el interés de los expertos educativos para descifrar las claves que hacen de éstos un medio tan eficaz para enseñar.

Un ejemplo de plataforma implementada con este método de aprendizaje es **ClassDojo**. Esta herramienta ayuda a los profesores a mejorar el aprendizaje en el aula de forma eficiente y divertida. En particular, enriquece conductas específicas e impulsa la participación activa de los estudiantes mediante la emisión de premios en tiempo real.

Finalmente, esta plataforma entrega informes de seguimiento para informar el progreso de los alumnos tanto a los padres y profesores.

Figura 4. Pantalla del portal de distribución de la plataforma de aprendizaje basada en gamificación, ClassDojo.



## • Aprendizaje móvil (Mobile Learning)

Los dispositivos móviles han transformado la vida cotidiana y la educación no ha sido una excepción. Hoy en día, muchas universidades, como la UNAM, ya están comenzando a utilizar esta tecnología para ofrecer una educación más amigable y eficaz con sus alumnos.

Actualmente, estos dispositivos promueven el aprendizaje por medio de contenidos educativos, como libros de texto electrónicos o cursos interactivos que coadyuvan a la mejora del rendimiento escolar. Asimismo, permiten la aplicación de exámenes de evaluación o de colocación que agilizan y simplifican el proceso de evaluación de los estudiantes.<sup>11</sup>

Entre las principales ventajas que los alumnos tienen al usar este método de aprendizaje se tienen:

- Disponer de información en tiempo real que permite a los alumnos descargar y revisar sus archivos, así como compartirlos y hacer aportes desde cualquier lugar



[11] Véase: <http://ideasqueinspiran.com/2015/06/23/mobile-learning-nuevas-posibilidades-nuevos-retos/>  
<http://www.uv.mx/cienhombre/revistae/vol20num2/articulos/aprendizaje/>  
[https://es.wikipedia.org/wiki/Aprendizaje\\_](https://es.wikipedia.org/wiki/Aprendizaje_)

en el que se encuentren.

- Atender a sus obligaciones académicas.
- Retroalimentar su aprendizaje entre pares o con grupos de estudiantes con intereses similares.
- Interactuar con sus profesores u otros estudiantes de manera simultánea o no.
- Participar en cursos a distancia.

Finalmente, estos dispositivos constituyen uno de los pilares más importantes para recabar información de estudiantes y de los procesos educativos a los que atienden, incrementando así las posibilidades de realizar analítica con la tecnología de *big data* e identificar nuevas áreas de oportunidad para mejorar la educación.

## ¿En qué consiste la analítica del aprendizaje ?

Se trata de la medición, recopilación, análisis y presentación de datos sobre los estudiantes, sus contextos y las interacciones que allí se generan, con el fin de comprender el proceso de aprendizaje que se está desarrollando y optimizar los entornos en los que se produce.<sup>12</sup>

En este caso, el lago de datos para el análisis se conforma a partir de acciones explícitas de los estudiantes, por ejemplo, su actividad académica (específicamente a través de la realización y envío de exámenes o tareas por medios electrónicos). Sin embargo, también se retoman otras actividades que no son parte del trabajo del alumno y que pueden englobar los sitios web que visitan, su ubicación, las obras que consultan, su actividad social en línea, etcétera.

Una vez extraídos los datos, éstos se analizan, se relacionan y se evalúan con el objeto de poder enfocar la enseñanza a las capacidades específicas de cada estudiante.

El concepto de analítica del aprendizaje buscará que exista una acción posterior a los resultados obtenidos, por ejemplo, que se elabore un nuevo plan de estudios, se busquen las técnicas para personalizar y adaptar la educación, así como predecir e intervenir sobre las capacidades del estudiante con el objetivo de mejorarlas.

La inmediatez y la eficacia de la analítica del aprendizaje servirán para que tanto el alumno conozca cuáles son sus propias dificultades, como para que sus maestros comprendan más a fondo las dificultades personales y grupales para enseñar.

Hoy en día, la analítica del aprendizaje es usada para:

- Detectar las debilidades y fortalezas de los sistemas educativos.
- Desarrollar nuevos planes de estudios y métodos educativos.
- Mejorar los cursos que imparten las instituciones educativas.
- Reflexionar sobre los logros y los patrones de comportamiento de cada alumno respecto a sus compañeros o al currículum educativo utilizado.
- Predecir en qué ámbitos específicos los estudiantes necesitarán apoyo y atención personalizada.
- Ayudar a los maestros y personal de apoyo en sus intervenciones.
- Ayudar a las instituciones de enseñanza en la gestión y la toma de decisiones.
- Adecuar los procesos de evaluación y mostrar a las necesidades cada vez más con-



[12] Véase: [https://es.wikipedia.org/wiki/Learning\\_analytics](https://es.wikipedia.org/wiki/Learning_analytics)

- cretas y específicas de los estudiantes.
- Diseñar currículos formativos más ajustados a las realidades de las instituciones educativas.

## Nuevas carreras y opciones profesionales derivadas a partir de la analítica del *big data*

El *big data* está propiciando también el surgimiento de nuevas carreras y opciones profesionales, como las que se comentan a continuación.

- **Científico de datos o consultor de *big data*.** Esta carrera está relacionada con la creciente necesidad de contar con profesionales aptos para el manejo, el análisis y la gestión de la información. Básicamente se trata de una persona formada con amplios conocimientos en matemáticas y estadística, que domina diferentes lenguajes de programación, las ciencias de la computación y la analítica. El científico de datos analiza, interpreta y comunica las nuevas tendencias encontradas en su área de investigación y las traduce a su empresa o institución para que ésta haga uso de ellas y de esta manera, puedan adaptar sus productos o servicios y crear nuevas oportunidades de negocio o de innovación que les resulten ventajosas.
- **Experto en analítica del aprendizaje.** Esta profesión está relacionada con los últimos avances en el ámbito de la enseñanza, como el desarrollo de nuevos entornos personalizados de aprendizaje y los MOOC (Massive Online Open Courses), cursos en línea masivos y abiertos. Las funciones de estos profesionales pueden ser: orientador especializado en la mejora del rendimiento escolar, administrador de las plataformas en línea, especialista en la atención de los alumnos, investigador educativo o asesor y orientador de directivos y profesores, etcétera.
- **Maestros digitales.** La mayoría de los nuevos empleos que van a generar la economía digital en los próximos años estarán ligados a la educación, sobre todo los maestros digitales. La formación de posgrado también estará en sintonía con la demanda del mercado laboral. Hoy en día, hace mucha falta profesionales que sean capaces de trasladar los contenidos educativos a la web. Este tipo de profesionistas deben tener una formación en tecnología más amplia y dominar otras disciplinas como la psicología, dado que requieren manejar las necesidades de los estudiantes, quienes se convierten en usuarios en línea.
- **Administrador de servidores y sistemas de almacenamiento.** Será responsable de administrar toda la información disponible en los sistemas de información. Quienes se dediquen a ello deberán garantizar que toda la información disponible en los servidores y sistemas de almacenamiento masivo esté debidamente actualizada, protegida, respaldada y disponible para su acceso en línea. Asimismo, serán responsables de la seguridad en línea.
- **Ciber-abogado.** Es el experto en delitos informáticos tales como las estafas a través de la web, los delitos en las redes sociales, *ciber-bullying*, la protección de datos y derechos de autor, la pornografía, el espionaje, robo de identidad, *phishing*, etcétera.
- **Desarrollador de Internet de las Cosas.** El Internet de las cosas o Internet of Things

(IoT), exige una nueva serie de profesionales capaces de crear software que conecte sensores con dispositivos móviles y sistemas de cómputo, o que éstos se enlacen directamente con los centros de datos. No hay que olvidar que mucha de la información que requiere el *big data* actualmente proviene de los diferentes tipos de dispositivos móviles y sensores alojados en las instalaciones de empresas e instituciones.

## Aspectos técnicos para la adopción y uso de la tecnología de *big data*

Una de las principales limitaciones para la adopción de la tecnología *big data* en cualquier institución son los recursos económicos, de infraestructura y de capital humano experimentado que se requieren para su instrumentación. No obstante, hoy en día, comienzan a aparecer alternativas relativamente económicas para poder realizar analítica con *big data*, entre ellas destaca BigQuery y algunas herramientas especializadas que están diseñadas para facilitar diferentes aspectos de esta tecnología.

BigQuery es una potente plataforma de análisis *big data* que puede ser utilizada por todo tipo de empresas u organizaciones a través de una tarifa accesible. En ella es posible almacenar y analizar terabytes de datos en segundos aprovechando la infraestructura tecnológica de Google.

Figura 5. Logotipo comercial del producto BigQuery de Google.



La plataforma de BigQuery es relativamente sencilla de utilizar y permite a los desarrolladores y analistas estudiar las bases de datos casi en tiempo real. Para esto provee consultas del tipo SQL sobre conjuntos de datos de terabytes de información, las cuales arrojan resultados en pocos segundos. Dichos resultados se pueden almacenar en tablas o exportar para su análisis o visualización posterior.

Técnicamente, BigQuery puede ser utilizado en tres modalidades: desde un navegador web, por medio de alguna herramienta de línea de comandos, o mediante la realización de llamadas a su API<sup>13</sup> REST<sup>14</sup> utilizando una gran variedad de librerías de Java, .NET o Python. También soporta una gran variedad de herramientas de terceros, tanto para la visualización como para la carga de datos.

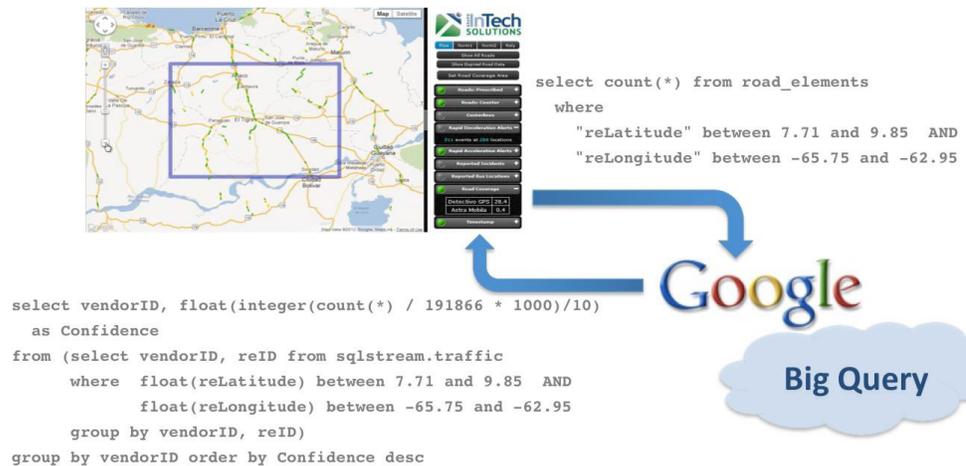
En la figura que se presenta a continuación se muestra un ejemplo de una consulta SQL de *big data* implementada con la plataforma de BigQuery de Google que aprovecha el Conector SQLstream (tecnología para la extracción y procesamiento de datos de un área geográfica en tiempo real).



[13] API (Application Programming Interface). Interfaz de programación de aplicaciones.

[14] REST (Representational State Transfer). Transferencia de Estado Representacional. Es un estilo de arquitectura software para sistemas hipermedia distribuidos como la *World Wide Web* basados en el HTTP.

Figura 6. Ejemplo de consulta SQL con Big Query de Google. Fuente: <http://www.sqlstream.com/solutions/google-bigquery/>



Amazon Inc., por su parte, también ofrece un conjunto integral de servicios de informática en la nube denominado AWS (Amazon Web Services), el cual permite realizar analítica con *big data*. Este servicio de paga, ofrece soluciones tecnológicas para cada fase del ciclo de vida de un proyecto de *big data* (recopilación, transmisión, almacenaje, uso de manejador de base de datos relacional o NoSQL, análisis y archivado).

En cuanto a la parte de analítica, que es la importante, Amazon presenta la potente estructura de Apache Hadoop a través de su producto Elastic MapReduce (EMR) para brindar un servicio fácil de utilizar. EMR es básicamente es un servicio web que facilita la gestión y el procesamiento rápido y rentable de grandes cantidades de datos.

Cabe mencionar que EMR administra de manera segura y confiable los casos de uso de *big data*, incluidos el análisis de registros, la indexación web, el almacenamiento de datos, el aprendizaje automático, el análisis financiero, la simulación científica y la bioinformática.

Figura 7. Logotipo comercial del producto EMR de Amazon. Fuente: <https://aws.amazon.com/es/elasticmapreduce/>



¿Pero qué es el Apache Hadoop?

El nuevo mundo de oportunidades que supone el *big data* no sería nada sin el software necesario para distribuir y procesar rápidamente "conjuntos de datos" con algoritmos especializados dentro de los sistemas computacionales. En este sentido, Apache Hadoop permite resolver varios aspectos importantes como:

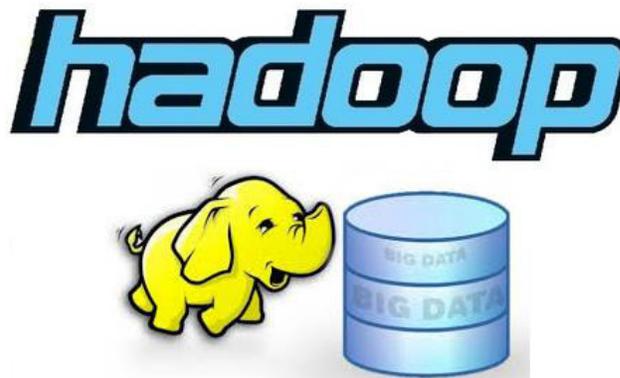
- Establece el punto de contacto con la ciencia de los datos (*data science*) que provee

las técnicas necesarias para manipular y tratar la información desde un punto de vista estadístico/matemático.

- Hace posible el procesamiento distribuido de grandes conjuntos de datos en arreglos de computadoras (*clusters*) utilizando modelos sencillos de programación.
- Permite escalar la infraestructura de cómputo, es decir, utilizar desde unos pocos servidores (nodos) y sistemas de almacenamiento de información, hasta miles de equipos, todos ellos ofreciendo una calidad idéntica de servicio.
- Permite a las aplicaciones desarrolladas trabajar con miles de nodos y petabytes de datos, tantos como necesidades de análisis de información se requieran.
- Dispone de facilidades para realizar consultas avanzadas sobre los conjuntos de datos.
- Permite el uso de extensiones que facilitan el trabajo, manipulación y seguimiento de toda la información que maneja.
- Permite ejecutar procesos en paralelo en todo momento.

Se trata, en definitiva, de un *framework*<sup>15</sup> de software de código abierto que se utiliza para almacenar, procesar y analizar grandes volúmenes de datos. Éste se orienta hacia la computación distribuida, en la cual la escalabilidad y la fiabilidad son los dos atributos más importantes. En otras palabras, Hadoop es el complemento perfecto de *big data*.

Figura 8. Logotipo comercial del framework Hadoop.  
 Fuente: <https://hadoop.apache.org/>



## ¿Qué hay sobre los sistemas de gestión de bases de datos (SGDB) utilizados para el *big data*?

Aquí se emplean básicamente de dos tecnologías, la SQL y la NoSQL. La primera contempla todos aquellos manejadores de bases de datos de alto nivel que incluyen un lenguaje de consulta estructurado o structured query language (SQL). Estos emplean bases de datos relacionales y se caracterizan porque soportan el manejo del álgebra y el cálculo relacional que permite efectuar consultas con el fin de recuperar información de la base de datos, así como hacer cambios en ella. Esta tecnología es ampliamente utilizada para tratar información estructurada.



[15] En el desarrollo de software, un *framework* o infraestructura digital, es una estructura conceptual y tecnológica de soporte definido, normalmente con artefactos o módulos concretos de software, que puede servir de base para la organización y desarrollo de éste.

En cambio, Los sistemas NoSQL difieren de las bases de datos relacionales en varios aspectos importantes, el más destacado es que no usan SQL como el principal lenguaje para las consultas. Asimismo, los datos que almacenan no emplean estructuras fijas como tablas, tampoco soportan operaciones para combinar registros de dos o más tablas en una base de datos relacional, ni garantizan completamente la atomicidad, consistencia, aislamiento y durabilidad de la información. Los sistemas NoSQL se denominan a veces “no sólo SQL” para destacar el hecho de que también pueden soportar lenguajes de consulta de tipo SQL extendido.

Ante esto, ¿por qué se usan ambas tecnologías?, la respuesta es sencilla: se complementan. Las bases de datos relacionales modernas han mostrado poca eficiencia en determinadas aplicaciones, por ejemplo, para el indexado de un gran número de documentos. Por ello se utilizan más en aplicaciones en las cuales se requiere manejar modelos transaccionales de datos. Por otro lado, la tecnología NoSQL permite manejar bases de datos orientadas a documentos. Entre las aplicaciones más populares implementadas esta tecnología se encuentran los manejadores de MongoDB, MemcacheDB, CouchDB, Apache Cassandra y HBase, etcétera.

## Los aspectos legales y éticos sobre el uso del *big data* en la educación

Si bien es cierto que los resultados obtenidos con la tecnología de *big data* pueden ser sorprendentes, su uso también plantea serios problemas legales y éticos.

Hay que recordar que detrás de cada número, de cada dato, hay una persona que no se tiene que ver como una sumatoria de patrones de conducta o de hábitos de estudio, sino como un ser humano que siente, que cuenta con derechos y valora su privacidad.

No se puede, ni se debe comenzar a recopilar datos sobre los alumnos y los profesores o generar informes analíticos sobre ellos sin su consentimiento. En este sentido, el asesoramiento legal es vital para el éxito de nuestro proyecto.

Algunos de los aspectos que se deberán considerar son:

- La propiedad de los datos que son recolectados. Éstos regularmente provienen de:
  - La actividad registrada por los estudiantes.
  - Sistemas en los que interactúan.
  - Mecanismos de retroalimentación de los estudiantes, como su participación en encuestas, portales de sugerencias y de atención o asesoría de servicios complementarios.
  - Redes sociales.
- Los datos que requieren ser compartidos entre sistemas, organizaciones y actores involucrados.
- Los principios de colección, conservación y explotación de los datos por parte de las instituciones educativas involucradas.

- La misión educativa y los temas que permanecen ocultos y no pueden verse fácilmente en la gestión del aprendizaje.
- Las expectativas de los centros educativos que decidan utilizar la analítica del *big data*.
- La motivación para el desarrollo del análisis.

Queda claro que el trabajo de las instituciones educativas en materia de *big data* debe apegarse al esquema legal para generar un entorno seguro y de confianza en el cual su comunidad se sienta cómoda. Los datos son la voz de los estudiantes y de la propia institución educativa a la que asisten y se deben analizar para poder extraer conocimiento con la tecnología, pero siempre en un marco ético-legal que permita reportar múltiples beneficios.

Hoy en día, se cuenta con ciertos mecanismos legales que permiten proteger con las leyes existentes a muchos de los avances tecnológicos que han surgido, como patentes, marcas, diseños industriales, derechos de autor; o bien mediante contratos. No obstante, existen muchas otras situaciones en materia de informática que se escapan de la ley o que los creadores, desarrolladores o productores de los mismos (incluso la propia legislación) no contemplan.

Es claro que a la legislación le queda un largo camino por recorrer para cubrir todos los aspectos legales que se puedan derivar de la analítica del *big data* y, en cierto modo, se debería estar consciente de que en la práctica hay muchas otras cuestiones legales que simplemente no se van a poder trasladar rápidamente a la realidad. Para ello, cada legislador deberá asesorarse de los especialistas en informática necesarios para encontrar soluciones reales, válidas y eficaces a las situaciones que presenta la analítica del *big data*.

## Conclusiones

- Se debe dar mayor impulso a la introducción de la analítica del *big data* en las instituciones de educación superior.
- Hace falta formar capital humano para realizar la analítica con *big data*.
- Hoy en día, existen mayores facilidades tecnológicas para adoptar y utilizar la tecnología de *big data* en las instituciones de educación superior.
- Los acervos digitales, que constituyen el producto de años de actividad académica, docente y de investigación, y que son generados de manera regular por las instituciones educativas, deben ser incorporados en el lago de datos para su aprovechamiento en los procesos de analítica de *big data*.
- El uso de la tecnología de *big data* plantea serios problemas legales y éticos que deben ser tomados en cuenta para su introducción en las instituciones educativas. 🌟

## Referencias bibliográficas

- [1] ÁLVAREZ, Paz, "15 profesiones confuturo", *Cinco Días. Diario Económico Español*, 2015, [en línea]: <[http://cincodias.com/cincodias/2015/01/30/sentidos/1422612761\\_107465.html](http://cincodias.com/cincodias/2015/01/30/sentidos/1422612761_107465.html)>, [Consultado: noviembre de 2015].
- [2] ALM, "Tendencias y desafíos para el e-learning en 2015", *America Learning Media*, 2014, [en línea]: <<http://www.americlearningmedia.com/edicion-035/395-tendencias/6201-tendencias-y-desafios-para-el-e-learning-en-2015>>, [Consultado: noviembre de 2015].
- [3] AMAZON, "Big Data", *Amazon Web Services*, 2015, [en línea]: <<https://aws.amazon.com/es/big-data/>>, [Consultado: noviembre de 2015].
- [4] BENÍTEZ MORENO, Rafael, "5 ejemplos de gamificación en el aprendizaje", *El Divan Digital*, 2013, [en línea]: <<http://www.benitezrafa.es/5-ejemplos-de-gamificacion-en-el-aprendizaje/>>, [Consultado: noviembre de 2015].
- [5] BOLLATTI, Rodolfo, "Bigdata en la Educación", *Workshop de Investigadores en Ciencias de la Computación*, 2013, [en línea]: <<http://imgbiblio.vaneduc.edu.ar/fulltext/files/TC112745.pdf>>, [Consultado: octubre de 2015].
- [6] EQUIPO MULTIDISCIPLINARIO, "Qué es 'Aprendizaje Adaptativo'", *CEDDIE Guadalajara*, 2015, [en línea]: <<http://centrodeinnovacionytecnologia.blogspot.mx/2014/01/que-es-aprendizaje-adaptativo.html>>, [Consultado: noviembre de 2015].
- [7] EQUIPO DE INVESTIGACIÓN UP, "Los Big Data en la educación", *Universo UP. Revista Digital de la Universidad de Padres*, 2014, Núm. 5, [en línea]: <<http://universoup.es/5-explorandoelhorizonte/los-big-data-en-la-educacion/>>, [Consultado: octubre de 2015].
- [8] EVAÏN, Agathe, "La Revolución del Big Data en la Educación", *EnderMetrics*, 2015, [en línea]: <<http://endermetrics.com/es/2015/06/la-revolucion-del-big-data-en-la-educacion/>>, [Consultado: noviembre de 2015].
- [9] GOOGLE, "¿What is BigQuery?", *Google Cloud Plataform, Alphabet Inc.*, 2015, [en línea]: <<https://cloud.google.com/bigquery/what-is-bigquery>>, [Consultado: noviembre de 2015].
- [10] LASKOWSKI, Nicole, "No hay un único camino para construir una plataforma de analítica de big data", *Searchdatacenter en español*, 2015, [en línea]: <[http://searchdatacenter.techtarget.com/es/cronica/No-hay-un-unico-camino-para-construir-una-plataforma-de-analitica-de-big-data?utm\\_medium=EM&asrc=EM\\_](http://searchdatacenter.techtarget.com/es/cronica/No-hay-un-unico-camino-para-construir-una-plataforma-de-analitica-de-big-data?utm_medium=EM&asrc=EM_)>

EDA\_48279080&utm\_campaign=20151005\_Consejos%20para%20construir%20una%20plataforma%20de%20an%e1lisis%20de%20Big%20Data\_&utm\_source=EDA>, [Consultado: octubre de 2015].

- [11] MORE, Mireia, "Los beneficios de usar Big Data en Educación", *IEBSchool*, 2015, [en línea]: <<http://comunidad.iebschool.com/iebs/formacion/usar-big-data-en-educacion/>>, [Consultado: octubre de 2015].
- [12] NIVARD ARBOIX , "Big data y educación II", *Blog de iDidactic*, 2015, [en línea]: <<http://www.ididactic.com/edblog/big-data-y-educacion-ii/>>, [Consultado: octubre de 2015].
- [13] -----, "Big data y educación III", *Blog de iDidactic*, 2015, [en línea]: <<http://www.ididactic.com/edblog/big-data-y-educacion-iii/>>, [Consultado: octubre de 2015].
- [14] PEARSON, "Las 5 principales tendencias de la educación 3.0", *Pearson Educación*, 2015, [en línea]: <<http://ideasqueinspiran.com/2015/05/07/principales-tendencias-de-la-educacion-3-0/>>, [Consultado: noviembre de 2015].
- [15] POWERDATA, "¿Cómo se relacionan Big Data y Hadoop?", *Blog de Especialistas en Gestión de Datos*, PowerData Solutions, 2013, [en línea]: <<http://blog.powerdata.es/el-valor-de-la-gestion-de-datos/bid/328879/C-mo-se-relacionan-Big-Data-y-Hadoop>>, [Consultado: noviembre de 2015].
- [16] RECUERO, Paloma, "Ética y Big Data. ¿Hasta dónde se debe llegar?", *TIC, TAC, TEP: Aprender en el siglo XXI*, 2015, [en línea]: <<https://palomarecuero.wordpress.com/2015/08/14/etica-y-big-data-hasta-donde-se-debe-llegar/>>, [Consultado: octubre de 2015].
- [17] -----, "Big Data en Educación: El problema de la privacidad", *TIC, TAC, TEP: Aprender en el siglo XXI*, 2015, [en línea]: <<https://palomarecuero.wordpress.com/2015/06/01/big-data-en-educacion-el-problema-de-la-privacidad/>>, [Consultado: octubre de 2015].
- [18] SÁNCHEZ, Escarlata, "Big Data al servicio de la educación", *Euronews*, Sección Learning world, 2015, [en línea]: <<http://es.euronews.com/2015/05/22/big-data-al-servicio-de-la-educacion/>>, [Consultado: octubre de 2015].
- [19] WIKIPEDIA, "Learning analytics", 2015, [en línea]: <[https://es.wikipedia.org/wiki/Learning\\_analytics](https://es.wikipedia.org/wiki/Learning_analytics)>, [Consultado: noviembre de 2015].