

Historia del Big Data

La historia del Big Data es muy poco conocida. Como sucede frecuentemente con las tendencias, parece que cuando explotan es algo muy novedoso que acaba de aparecer; pero en muchos casos son la eclosión de algo que ha estado madurando durante largo tiempo.

Como hemos comentado en otras ocasiones, el Big Data es el análisis de un gran volumen de conjuntos de datos. Para recolectar, tratar y analizar esa ingente cantidad de información se necesitan fórmulas de procesamiento potentes y rápidas. Por ello, estas técnicas parecen haber surgido recientemente, gracias a los avances tecnológicos.

Esto, en gran medida es así: el término se comienza a utilizar de forma generalizada a finales de los años 90 y el boom llega con los avances experimentados en campos como internet, dispositivos móviles y conexión. Sin embargo, la historia del Big Data se remonta a mucho antes.

Hay incluso quienes lo sitúan en el paleolítico, con una lógica que relaciona el término con el primitivo interés de los seres humanos por lograr y procesar la información. A continuación ahondamos un poco más en esta cuestión compartiendo la historia del Big Data:

18000 AEC. En el Paleolítico Superior se empleaban rudimentarios métodos de almacenamiento de datos con el empleo de palos o muescas en huesos. Con este sistema, se podía llevar cuenta de provisiones, realizar cálculos básicos e incluso predecir necesidades de comida para el grupo. Quizá sea demasiado incluirlo en la historia del big data, pero es el primer momento documentado en el que la humanidad se interesa por los datos: el germen de todo lo que viene después. Si bien las cantidades no podían ser muy grandes, es la primera evidencia del interés por recopilar, contar y guardar datos.

2400 AC. En Babilonia se extiende el uso del ábaco, un sistema para realizar cálculos. En esta época surgen también las primeras bibliotecas como lugares para almacenar y consultar conocimiento.

48 AC. Los Romanos invaden Alejandría y accidentalmente destruyen su famosa biblioteca. Parte de los fondos se trasladaron a otros lugares, pero la mayoría de la colección fue quemada, perdida o robada. Hasta el momento, había logrado reunir medio millón de documentos con la intención de almacenar todo el conocimiento de la Humanidad.

Siglo II AC. Se desarrolla la primera computadora mecánica conocida en Grecia. El mecanismo de Anticitera era un aparato analógico de bronce diseñado para predecir posiciones astronómicas, probablemente una evolución de otros sistemas que no se han recuperado a día de hoy. Se empleó para el estudio astrológico y para marcar el calendario, fundamentalmente las fechas exactas de los antiguos Juegos griegos.

1663. John Graunt realiza el primer experimento de análisis de datos estadísticos conocido. Con los datos de defunciones, teoriza un sistema de alerta para la peste bubónica en toda Europa.

1792. Aunque hay constancia de análisis estadísticos desde las Guerras del Peloponeso y la palabra estadística se acuña en Alemania unos años antes; en 1792 se asocia el término a la "colección y clasificación de datos".

1865. Aparece por primera vez el término business intelligence, en la enciclopedia comercial de Richard Millar Devens. En ella describe cómo el banquero Henry Furnese logró una importante ventaja competitiva

Big Data / Análisis de Datos

Historia de Big Data 1/3

Agencia de
Aprendizaje
a lo largo
de la vida

recogiendo, estructurando y analizando datos clave de su actividad. La inteligencia de negocio es sin duda uno de los grandes motores de la analítica dentro de la historia del big data.

1880. Herman Hollerith, empleado del censo estadounidense, desarrolla su máquina tabuladora. Con ella consigue reducir un trabajo de 10 años a 3 meses. Este ingeniero funda una compañía que posteriormente se conocería como IBM.

1926. Nikola Tesla predice la tecnología inalámbrica. Según su visión, el planeta es un gran cerebro en el que todo está conectado, por lo que deberíamos ser capaces de simplificar el uso del teléfono. Predice que cada hombre llevará uno en su propio bolsillo.

1928. El ingeniero alemán Fritz Pfleumer patenta el primer sistema magnético para almacenar datos. Sus principios de funcionamiento se utilizan hoy en día.

1944. Primer intento de conocer la cantidad de información que se crea. Se trata de un estudio académico de Fremont Rider, que pronostica 200 millones de libros en la Universidad de Yale en 2040, almacenados en 6.000 millas de estanterías.

1958. El informático alemán Hans Peter Luhn, define la inteligencia de negocio: la habilidad de percibir las interrelaciones de los hechos presentados para guiar acciones hacia un objetivo deseado. En 1941 pasó a ser Gerente de Recuperación de Información en IBM.

1962. Se presenta IBM Shoebox en la Expo de 1962. Creada por William C. Dersch supone el primer paso en el reconocimiento de voz, capaz de registrar palabras en inglés en formato digital.

1965. Se proyecta el primer data center en Estados Unidos, para guardar documentación de impuestos y huellas dactilares en cintas magnéticas. Un año antes comienzan a surgir voces que alertan del problema de guardar la ingente cantidad de datos generada.

1970. IBM desarrolla el modelo relacional de base de datos, gracias al matemático Edgar F. Codd. Este científico inglés es también responsable de las doce leyes del procesamiento analítico informático y acuñó el término OLAP.

1976. Se populariza el uso de MRP (software de gestión de materiales), antecedentes de los ERP actuales, que mejoran la eficiencia de las operaciones en la empresa; además de generar, almacenar y distribuir datos en toda la organización.

1989. Erik Larson habla por primera vez de Big Data en el sentido que conocemos la expresión hoy en día. La revista Harpers Magazine recoge su artículo, en el que especula sobre el origen del correo basura que recibe. En torno a este año se empiezan a popularizar las herramientas de business intelligence para analizar la actividad comercial y el rendimiento de las operaciones.

1991. Nace internet, a la postre, la gran revolución de la recolección, almacenamiento y análisis de datos. Tim Berners-Lee establece las especificaciones de un sistema de red con interconexiones a nivel mundial accesible para todos en cualquier lugar.

1993. Se funda QlikTech, germen de la actual Qlik, que crea un sistema revolucionario de business intelligence (en 2012, Gartner comienza a hablar de business discovery para definir ese tipo de análisis).

1996. Los precios del almacenamiento de datos empiezan a ser accesibles con un coste eficiente en lo que es una de las grandes revoluciones en la historia del big data. El libro La evolución de los sistemas de almacenamiento, de 2003, establece esta fecha como el primer año en el que el almacenamiento digital es más barato que el papel.

Big Data / Análisis de Datos

Historia de Big Data 2/3



Agencia de
Aprendizaje
a lo largo
de la vida

1997. Google lanza su sistema de búsqueda en internet y en los siguientes años será de largo el primer lugar al que acudir en busca de datos en internet. Este mismo año, se publica el estudio ¿Cuánta información hay en el mundo?, de Michael Lesk. La conclusión es que hay tanta y crece a tal velocidad, que gran parte de ella no será vista por nadie jamás.

1999. El término Big Data es analizado por primera vez en un estudio académico. La Asociación de Sistemas Informáticos recomienda centrarse en el análisis de información ya que existe gran cantidad de datos y no todos son útiles. Recuerdan el propósito de la computación, que es el entendimiento, no los números.

2001. Doug Laney, de Gartner, define las 3 V's del Big Data. Este es un hito clave en la historia del big data. Se trata de tres conceptos que definen el término: volumen, velocidad y variedad. Al mismo tiempo, se populariza el concepto SaaS (software as a service).

2005. Nace la Web 2.0, una web donde predomina el contenido creado por los usuarios. Este mismo año se crea Hadoop, un entorno de trabajo Big Data de software libre.

2007. La revista Wired publica un artículo que lleva el concepto de Big Data a las masas.

2010. Los datos que se generan en dos días equivalen a la cantidad de datos generados desde el inicio de la civilización hasta 2003, según Eric Schmidt (Google).

2013. El archivo de mensajes públicos de Twitter en la Biblioteca del Congreso de Estados Unidos llega a los 170 billones de mensajes, creciendo a ritmo de 500 millones al día. Según la institución que alberga algunos de los documentos históricos más importantes del mundo, dicho archivo ofrece una imagen más amplia de las normas culturales, diálogos, tendencias y eventos de hoy en día. De este modo, contribuye a una mejora de la información en procesos legislativos, educación, definición de autoría de nuevos trabajos y otras cuestiones.

2014. Los móviles superan a los ordenadores en accesos a internet. La conexión casi continua contribuye a generar muchos más datos y mejora la conectividad con otros dispositivos.

2016. El Big Data se convierte en la palabra de moda. Se generaliza la contratación de expertos en Big Data, el Machine Learning llega a las fábricas y el Internet de las Cosas empieza a impregnarlo todo.

2017. Los datos llegan a las masas. La gente controla sus patrones de descanso con pulseras, sabe en qué se gasta el dinero con aplicaciones móviles y se informa sobre la posesión de balón de su equipo de fútbol. Los datos están en todas partes y la población está ya predispuesta a usarlos.

Futuro. ¿Qué nos deparará el futuro? Muy difícil de pronosticar, pero seguramente un aumento de datos y la consiguiente necesidad de tecnología para recogerlos, adaptarlos, almacenarlos y analizarlos. La computación cuántica está a la vuelta de la esquina y la historia del big data sigue avanzando.

Fuente: <https://ignsl.es/historia-del-big-data/>