

# Breve historia del Big Data

ALICIA HERNÁNDEZ MARTÍN

Hace unos meses se publicó un artículo de Bernard Marr titulado *"Una breve historia sobre Big Data que todo el mundo debería leer"*. Se trata de un repaso por la historia del análisis y almacenamiento de datos, desde la época antigua hasta nuestros días, que ayudará a comprender que los cimientos en que se basan los principales conceptos del Big Data surgieron mucho tiempo atrás.

Como veremos más adelante, mucho antes de que se extendiese el uso cotidiano de los ordenadores, en el mundo académico ya existía la idea de que se estaba creando un cuerpo cada vez mayor de conocimientos que debía ser analizado y almacenado para su posterior recuperación. Nuestra capacidad para capturar y procesar información se ha ido incrementando gradualmente, aunque ha sufrido

una gran aceleración desde finales del siglo pasado.

A continuación, basándome en el artículo mencionado, haré un repaso por la historia del conocimiento y la innovación en el almacenamiento de datos para conocer su desarrollo.

Comenzaremos mencionando la Historia Antigua de los datos. Una de las primeras evidencias conocidas del almacenamiento de datos es el Hueso de Ishango, descubierto

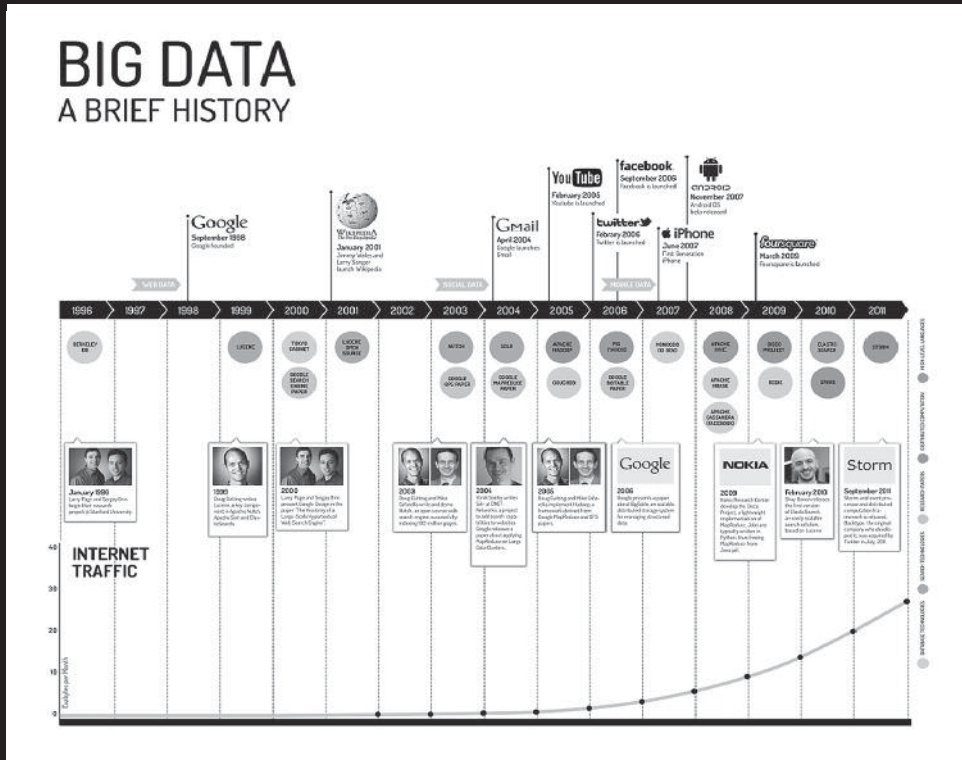
en 1960 y datado sobre el año 20000 A.C. En principio, se pensaba que era un simple palo de conteo, pero algunos científicos sugieren que las muescas incrustadas en el hueso indican un entendimiento

del que se tiene constancia. Se cree que fue desarrollado para fines astrológicos y para el seguimiento del ciclo de los Juegos Olímpicos.

La siguiente etapa en la historia del almacenamiento de

carga de trabajo del Censo de Estados Unidos, disminuyendo el tiempo empleado para la carga de datos de 10 años a 3 meses. Además funda IBM pasando a ser considerado el padre de la computación automatizada.

Los primeros días del almacenamiento de datos moderno vienen marcados por las predicciones de Nikola Tesla en



matemático que va más allá del conteo.

Sobre el año 18000 A.C., se tiene constancia de que los seres humanos empiezan a utilizar palillos para registrar datos que les permitía controlar la actividad comercial y el registro de inventarios.

En el 2400 A.C. se desarrolla el ábaco, instrumento construido específicamente para el cálculo. En esa época, además, surgieron las primeras bibliotecas en Babilonia.

Durante el período comprendido entre el 300 A.C. y 48 D.C. encontramos el mayor centro de almacenamiento de datos del mundo hasta ese momento, la Biblioteca de Alejandría.

Y entre los años 100 y 200 D.C se desarrolla en Grecia el mecanismo de Anticitera, el primer ordenador mecánico

datos vendría determinada por la aparición de la estadística. En el año 1963 John Graunt lleva a cabo los primeros experimentos estadísticos de análisis que se conocen, en un intento de frenar la peste bubónica en Europa.

En 1865 surge el término "inteligencia de negocios" cuando Richard Millar Devens lo utiliza en su *Enciclopedia de anécdotas comerciales y empresariales*, para describir cómo un banquero logra una gran ventaja sobre sus competidores mediante la recopilación y análisis de la información más relevante para sus actividades comerciales de manera estructurada.

En el año 1881, el ingeniero Herman Hollerith crea la máquina de tabulación Hollerith, que utiliza tarjetas perforadas para reducir la enorme



1926, al asegurar que en el futuro, cuando la tecnología inalámbrica esté instaurada en toda la tierra, un hombre será capaz de acceder y analizar grandes cantidades de datos utilizando un dispositivo tan pequeño que cabrá en un bolsillo.

Un par de años más tarde, el ingeniero Fritz Pfleumer crea un método de almacenamiento de datos magnético en cinta, que será la base de la tecnología digital moderna, que aún hoy se utiliza para el almacenamiento de datos en discos duros de ordenador.

Si decíamos que en 1865 apareció por primera vez el término "inteligencia de negocios", podemos afirmar que este concepto se desarrolló a partir de 1958 cuando el investigador de IBM Hans Peter Luhn lo define como "la capacidad de aprender las interrelaciones de los hechos presentados, de tal forma, que permiten orientar la acción hacia una meta deseada".

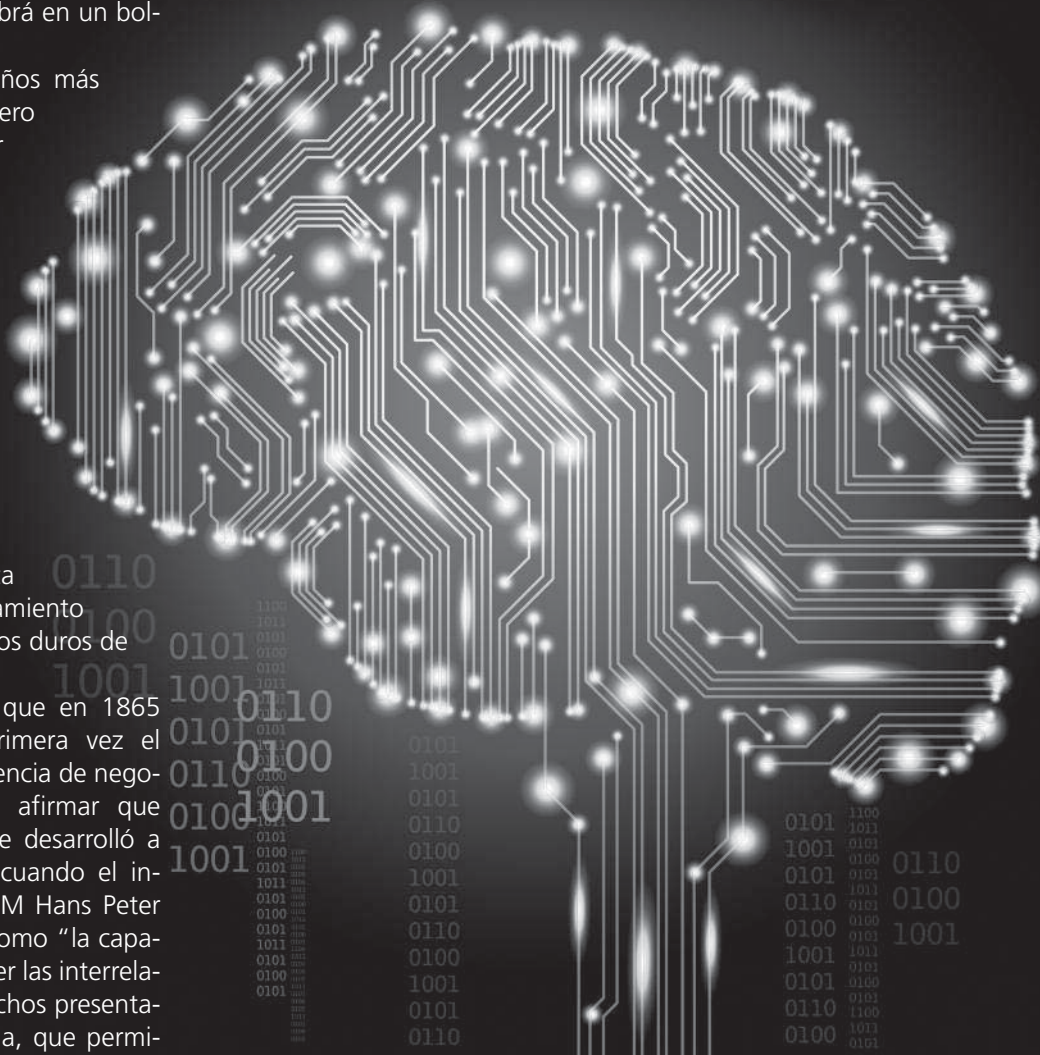
Por otra parte, comienza el desarrollo de los grandes centros de datos. En 1965 el Gobierno de Estados Unidos crea el primero de ellos para almacenar en cinta magnética 742 millones de declaraciones de impuestos y 175 millones de juegos de huellas digitales.

En 1970 el matemático Edgar F. Codd desarrolla para IBM el modelo de base de datos relacional que muchos servicios de información modernos utilizan para almacenar sus datos en un formato jerárquico y accesible para cual-

quier usuario que tenga claro lo que está buscando.

A partir de 1976 comienzan los primeros usos comerciales de los ordenadores,

En 1996 el precio del almacenamiento digital comienza a ser más rentable que en papel y en 1997 Google lanza su motor de búsqueda, el cual,



cuando los sistemas de planificación de requerimientos de materiales empiezan a ser de uso común en los negocios y la informática y el almacenamiento de datos se utilizan para tareas de rutina diaria.

Sin embargo, el gran cambio en cuanto al almacenamiento y recuperación de datos, se da en el año 1991 con el nacimiento de Internet. A partir de ese momento cualquier persona podrá navegar en línea, subir sus propios datos y analizar y recuperar los cargados por otros usuarios.

rápidamente se convertirá en el más popular del mundo. En esos momentos Mikel Lesk estima que el universo digital está aumentando diez veces más su tamaño cada año.

Las primeras ideas sobre Big Data surgen en 1999 cuando el término es utilizado por primera vez en un trabajo académico titulado "Visually Exploring Gigabyte Datasets in Realtime (ACM)".

En ese momento aparece también el concepto de "Internet de las cosas" en una presentación de negocios realizada

