

El big data

Analítica de negocios



La tendencia a la utilización de los sistemas de información en todos los ámbitos es cada vez más intensa y está generando cantidades enormes de datos, los cuales se cuentan por millones por hora y no pueden ser procesados en sistemas tradicionales por su volumen y complejidad; de ahí se deriva la creación del big data. Este término es para describir la captación de grandes volúmenes de datos con un crecimiento constante.

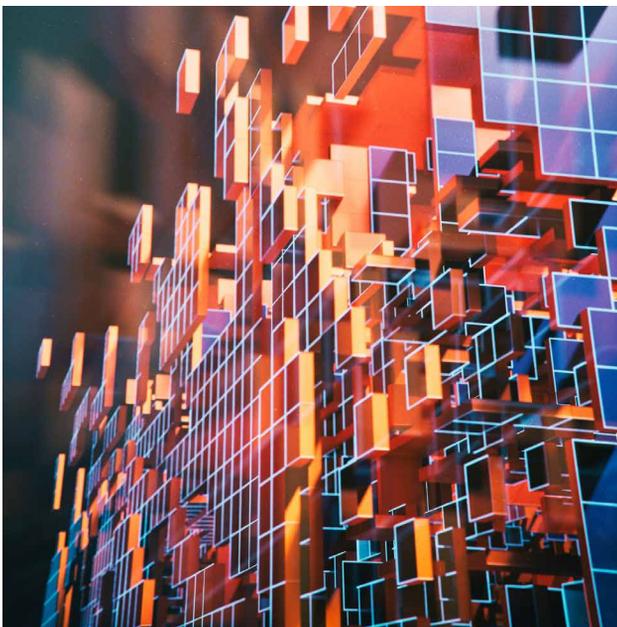


El big data se puede englobar en tres categorías:

- a. Estructurado
- b. No estructurado
- c. Semi estructurado

El big data tiene muchas características por su naturaleza; no obstante, las más reconocidas son las siguientes:

1. Volumen
2. Variedad
3. Velocidad
4. Variabilidad

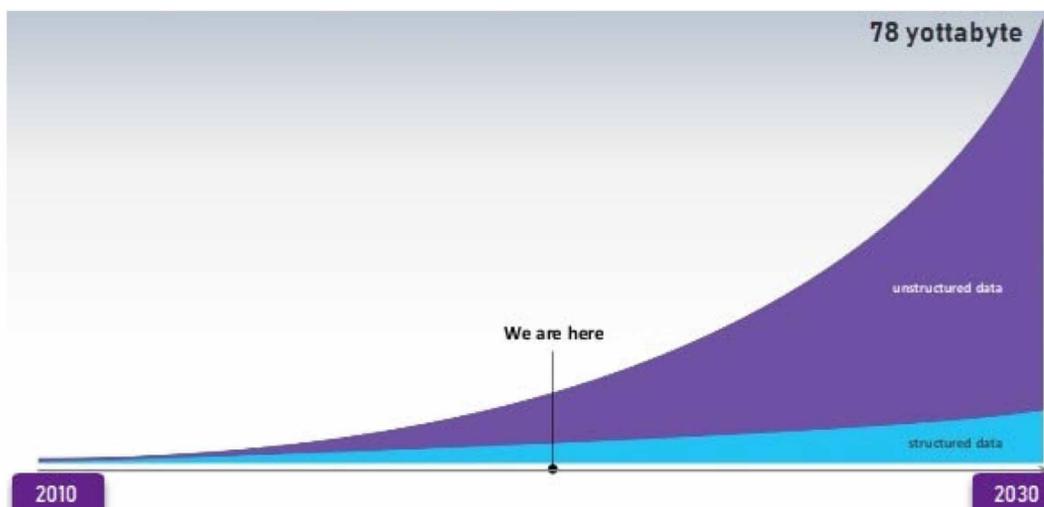


Estas cuatro características principales son las que le dan el nombre de big data. Como se mencionó, el volumen de los datos y su variedad hacen complejo el análisis de los mismos, además, la toma de decisiones demanda una pronta difusión de estos datos, por lo que el volumen de procesamiento debe ser muy intenso, ya que de esa prontitud depende el valor de los datos y la variabilidad dificulta el proceso de manejar y administrar, debido a la inconsistencia de los mismos, de forma efectiva.

Los diferentes tipos de datos se refieren a fuentes heterogéneas y a la naturaleza de los mismos tanto estructurados como no estructurados. En el pasado, las hojas de cálculo y las bases de datos eran las únicas fuentes de datos consideradas por la mayoría de las aplicaciones, pero hoy en día, los datos en forma de correos electrónicos, fotos, videos, dispositivos de monitoreo, PDF, audio, etc. también se están considerando en las aplicaciones de análisis. Esta variedad de datos no estructurados plantea ciertos problemas para el almacenamiento, la minería y el análisis de estos.

Las economías desarrolladas utilizan cada vez más tecnologías intensivas en datos. Hay 4,600 millones de suscripciones de teléfonos móviles en todo el mundo y entre 1,000 y 2,000 millones de personas acceden a Internet. Entre 1990 y 2005, más de mil millones de personas en todo el mundo ingresaron a la clase media, lo que significa que más personas se alfabetizaron, lo que a su vez condujo al crecimiento de la información. La capacidad efectiva mundial para intercambiar información a través de las redes de telecomunicaciones fue de 281 petabytes en 1986, 471 petabytes en 1993, 2.2 exabytes en 2000, 65 exabytes en 2007 y las predicciones indican que la cantidad de tráfico de Internet será de 667 exabytes anualmente para 2014. Según una estimación, un tercio de la información almacenada globalmente está en forma de texto alfanumérico y datos de imágenes fijas, que es el formato más útil para la mayoría de las aplicaciones de big data.

Crecimiento de datos a lo largo de los años.



Fuente: Hu, F. (Ed.). (2016). *Big data: storage, sharing, and security*. CRC Press.

En términos generales, el big data se define como algo de gran tamaño y es una colección de datos que continúa creciendo constantemente. Algunas de las fuentes más importantes de datos para big data son las redes sociales, *stock exchanges*, telefonía, etc.

La obra presentada es propiedad de ENSEÑANZA E INVESTIGACIÓN SUPERIOR A.C. (UNIVERSIDAD TECNILENIO), protegida por la Ley Federal de Derecho de Autor; la alteración o deformación de una obra, así como su reproducción, exhibición o ejecución pública sin el consentimiento de su autor y titular de los derechos correspondientes es constitutivo de un delito tipificado en la Ley Federal de Derechos de Autor, así como en las Leyes Internacionales de Derecho de Autor.

El uso de imágenes, fragmentos de videos, fragmentos de eventos culturales, programas y demás material que sea objeto de protección de los derechos de autor, es exclusivamente para fines educativos e informativos, y cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por UNIVERSIDAD TECNILENIO.

Queda prohibido copiar, reproducir, distribuir, publicar, transmitir, difundir, o en cualquier modo explotar cualquier parte de esta obra sin la autorización previa por escrito de UNIVERSIDAD TECNILENIO. Sin embargo, usted podrá bajar material a su computadora personal para uso exclusivamente personal o educacional y no comercial limitado a una copia por página. No se podrá remover o alterar de la copia ninguna leyenda de Derechos de Autor o la que manifieste la autoría del material.