

# **BIG DATA EN CONTEXTO**

# BIG DATA EN CONTEXTO

Las empresas se empezaron a dar cuenta de esto hace ya más de 25 años, cuando la minería de datos (data mining) se incorporó como disciplina dentro de las prácticas de la inteligencia de negocio (BI, Business Intelligence). Fallaban dos elementos importantes:

1. El primero de ellos era la **cultura empresarial**.
2. El segundo era el **estado de la tecnología**.

# BIG DATA EN CONTEXTO

Quizás lo más propio sería hablar de Big Data como un fenómeno que se da alrededor de 2010 cuando convergen cuatro factores:

- 1) La aparición y consolidación de nuevos modelos de negocio que se basan en la sobre los que se realimentan. El auge del internet de las cosas, con la explosión participación en las redes sociales o la preponderancia del comercio electrónico son algunos ejemplos.
- 2) Un cambio en la cultura empresarial, conducido por una nueva generación de formación interdisciplinar y acostumbrada a la tecnología.

# BIG DATA EN CONTEXTO

- 3) Un avance tecnológico que, girando entorno a la **computación en la nube** (cloud computing), proporciona nuevas funcionalidades y capacidades de unos modelos de consumo muy flexibles y con costes asequibles.
- 4) Una oportunidad de negocio en sí mismo, explotada por los departamentos de marketing de las grandes corporaciones tecnológicas.

# BIG DATA EN CONTEXTO

- 3) Un avance tecnológico que, girando entorno a la **computación en la nube** (cloud computing), proporciona nuevas funcionalidades y capacidades de unos modelos de consumo muy flexibles y con costes asequibles.
- 4) Una oportunidad de negocio en sí mismo, explotada por los departamentos de marketing de las grandes corporaciones tecnológicas.

# BIG DATA EN CONTEXTO

## 1. Modelo de las 5 V

Hay un cierto consenso a la hora de caracterizar el Big Data como la confluencia de **cinco propiedades** que presentan los datos que manejan las organizaciones.

# BIG DATA EN CONTEXTO

## 1. Modelo de las 5 V

- Volumen.
- Velocidad.
- Variedad.
- Veracidad.
- Valor.

# BIG DATA EN CONTEXTO

## 2. Computación en la nube.

Podemos ver el Big Data como una respuesta a una serie de retos tecnológicos que plantearon en su momento las características de los datos que hemos estado viendo.

A la hora de gestionar grandes y variados volúmenes de datos, que se generan y mueven a gran velocidad, la consolidación del modelo de **computación en la nube** (cloud computing) fue decisiva.



# BIG DATA EN CONTEXTO

## 2. Computación en la nube.

Podemos ver el Big Data como una respuesta a una serie de retos tecnológicos que plantearon en su momento las características de los datos que hemos estado viendo.

A la hora de gestionar grandes y variados volúmenes de datos, que se generan y mueven a gran velocidad, la consolidación del modelo de **computación en la nube** (cloud computing) fue decisiva.

# BIG DATA EN CONTEXTO

## 2. Computación en la nube.

La computación en la nube permite abstraer los distintos niveles que componen una aplicación, de forma que son gestionados por el proveedor de forma transparente y facturados por su uso. Esto ofrece mucha flexibilidad a la hora de mover a la nube cargas de trabajo existentes, agilizando al mismo tiempo la creación de aplicaciones nativas. Básicamente nos podemos encontrar **3 modelos** de consumo que se pueden combinar entre sí, siendo habitual que un mismo proveedor los ofrezca todos:

# BIG DATA EN CONTEXTO

## 2. Computación en la nube.

1. Un modelo de infraestructura como servicio (**IaaS**, *Infrastructure as a Service*), que permite el aprovisionamiento, configuración y consumo de servidores físicos y virtuales, almacenamiento y elementos de red. Proporciona también funcionalidades de balanceo de carga, autoescalado de alta disponibilidad o seguridad perimetral. Es el modelo que ofrece más flexibilidad, pero también un mayor esfuerzo de administración. Permite despegar rápidamente aplicaciones existentes sin prácticamente necesidad de modificarlas.

# BIG DATA EN CONTEXTO

## 2. Computación en la nube.

2. Un nivel por encima se sitúan los servicios de **plataforma** (**PaaS**, *Platform as a Service*). Aquí la abstracción cubre todos los elementos, exceptuando el propio código y los datos de las aplicaciones. Este nivel es muy amplio, ya que podemos encontrar desde componentes todavía muy próximos a la infraestructura, como plataformas de contenedores, hasta entornos “sin servidor” (serverless) orientados a la gestión de eventos. También encontramos aquí servicios de base de datos, integración de aplicaciones, aprendizaje automático, inteligencia artificial, componentes analíticos o herramientas de desarrollo. Es decir, en este nivel dispondremos de piezas que podremos ensamblar mediante código para conformar nuestra aplicación.

# BIG DATA EN CONTEXTO

## 2. Computación en la nube.

3. El modelo de **software** como servicio (**SaaS**, *Software as a Service*), donde consumimos, ya como usuarios finales, una aplicación (un cliente de correo electrónico, un procesador de textos o un gestor de campañas de marketing, por ejemplo). El proveedor se encarga de toda la gestión del ciclo de vida, incluyendo la liberación de nuevas versiones y la gestión de incidencias. La principal cualidad de este nivel sería la instantaneidad de la contratación y el acceso. Por el contrario, es el modelo menos flexible de todos, ya que las posibilidades de adaptación y programación suelen ser limitadas por motivos de estandarización.

# BIG DATA EN CONTEXTO

## 3. Gobierno y gestión del dato.

Cuando hablamos de gobierno del dato nos estamos refiriendo al **conjunto de procesos, roles, políticas, estándares y mediciones** que permiten el uso eficiente de los datos dentro de una organización. Va mucho más allá de un conjunto de buenas prácticas, convirtiéndose en un marco que orchestra y regula una empresa orientada por los datos, tal y como veíamos anteriormente. El gobierno del dato cubre un campo tremendamente deben cubrir todo el ciclo de vida del dato: captación, almacenamiento, transformación, consumo y eliminación.

# BIG DATA EN CONTEXTO

## 3. Gobierno y gestión del dato.

Un aspecto transversal en el gobierno del dato es la **gestión de los metadatos**. Los metadatos son propiedades de los datos, adicionales a su contenido, que nos sirven para documentarlos, tanto desde el punto de vista técnico como de negocio. Los datos parten de con una serie de metadatos descriptivos que nos informan de:

# BIG DATA EN CONTEXTO

## **3. Gobierno y gestión del dato.**

- Origen.
- Creador o formato,
- Metadatos que se han ido incorporando debido a las transformaciones que han ido sufriendo,
- Sus relaciones con otros datos,
- Su calidad.



# BIG DATA EN CONTEXTO

## **3. Gobierno y gestión del dato.**

- Origen.
- Creador o formato,
- Metadatos que se han ido incorporando debido a las transformaciones que han ido sufriendo,
- Sus relaciones con otros datos,
- Su calidad.