



Data 101



Motivación



- El big data está teniendo un crecimiento anual de entre el 8 y el 10%
- Hay un problema serio: la falta de personal cualificado
- Todo gira cada día más en torno al Big Data:
 - Lo “conocido”: Inteligencia de Negocio como la conocemos
 - La inteligencia artificial
 - Traducción de documentos
 - Reconocimiento facial
 - Coches autónomos
 - Detección de contenidos inadecuados
 - Asistentes personales
- Hemos pasado de ser unos frikis raros a unos frikis sexys



■ Nuestra fuerza



"Without data
you're just another person
with an opinion."

W. Edwards Deming



Buzzwords

Confusión

LARGER
CONTINUES
TENS
COMPLEX
ANALYTICS
USED
CREATED
NOW
CAPACITY
BIOLOGICAL
RECORDS
NETWORKS
HUNDREDS
DATABASES
SEARCH

SET
USE
RELATIONAL
SOCIAL
INDEXING
CITATION
UBQUITOUS
SOLID
TYPES
CURRENTLY
GARTNER
DIFFICULTY
OPPORTUNITIES
WORKING
AMOUNT
ELAPSED
FORWARD
SYSTEMS
PARALLEL
MASSIVELY
GROW

BIG DATA

TOOLS
DISK
TARGET
EXAMPLES
SHARED
SENSOR
RECONSIDER
ZETTABYTES
PRACTITIONERS
INTERNET
MANAGEMENT
DESCRIBING
RADIOS
FREQUENCY
DISTRIBUTED
SETS
TERABYTES
CASE
INCLUDE
TOLERABLE
ABILITY
SIZE
RESEARCH
LOGS
DEFINING
RELATED
COMPUTING
QUALITIES
SAN
STORAGE
SYSTEMS
PARALLEL
MASSIVELY
GROW

ALSO
EVERY
MOVING
WITHIN
THROUGH
CURRENT
MAY
WIRELESS
ARCHIVES
MANAGE
BIOGEOCHEMICAL
TECHNOLOGIES
GENOMICS
NEEDED
TIME
SOFTWARE
PERFORMANCE
PROCESS
INFORMATION
PRESENTATIONS
STORE
COMBAT

Solape



Buzzwords

Término paraguas

Big
Data

Analytics



■ Luego...

Big Data Analytics



¿Qué ocurre cuando uno se decide a estudiar “Big Data”?



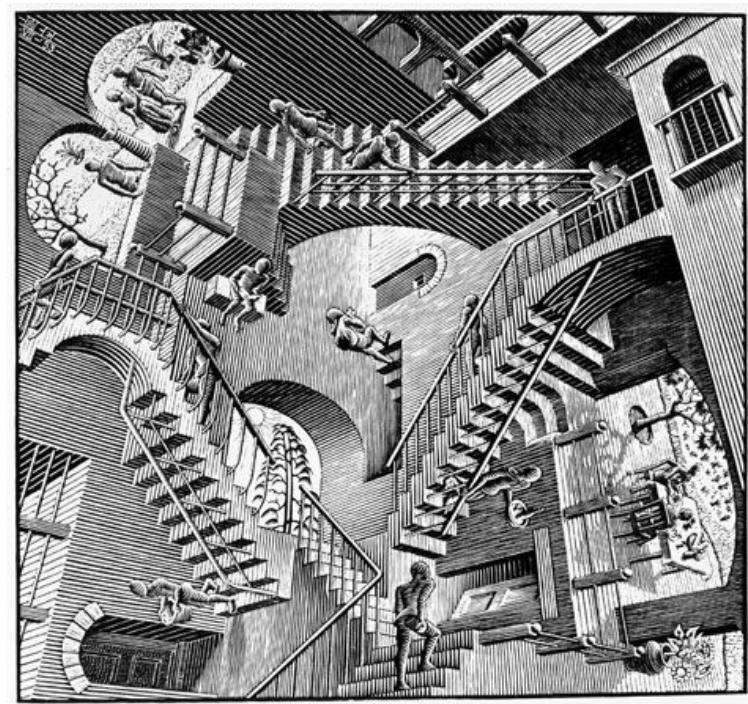
Final 2018 version, updated 03/15/2018

© Matt Turck (@mattturck), Remi Obavomi (@demi_obavomi), & FirstMark (@firstmarkcap)

mattturck.com/bigdata2018



■ Me perdí...

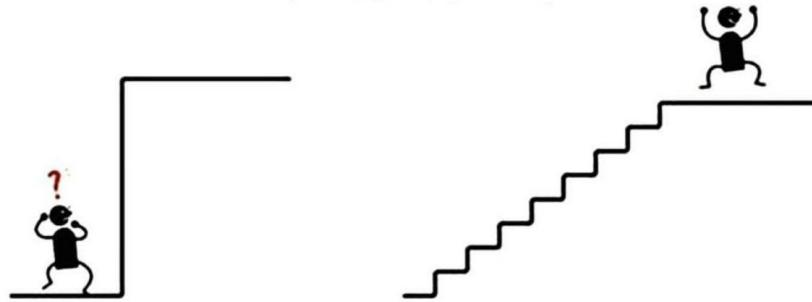


Estás más perdido que un muelle cayendo por las escaleras de Hogwarts



■ ¿Qué pretendemos conseguir?

**NINGUNA META ES DIFÍCIL
SI LA DIVIDIMOS EN
OBJETIVOS PEQUENOS
Y ALCANZABLES.**





■ ¿Utilizas Amazon, Spotify o Netflix?



...Entonces te estás beneficiando del Big Data

- Nuestra **huella digital** es cada vez más importante.
 - Separación entre la huella digital profesional y personal
 - Perfil digital coherente - Borrado de trazas antiguas
- Es importante entender que el big data plantea una **oportunidad** para tu vida digital, y por otro, vale la pena conocer qué **medidas inteligentes** podemos tomar para lograr que el big data trabaje en nuestro beneficio a la par que **protejamos nuestra identidad**.
- Precio de nuestros datos: ¿Deberíamos cobrarlos?



■ Cuando uno piensa que conoce la necesidad de sus usuarios...





■ Big Data en el mundo multimedia...



- Disney lanza los trailers de Star Wars analizando la expectación en redes sociales
- Se usa en la investigación para averiguar las preferencias de los clientes potenciales sobre argumentos, castings, etc. de series de televisión y películas
- Netflix tiene etiquetadores (taggers) para la clasificación de sus series
- En el caso de “La Gran Muralla” se eligió a Matt Damon como personaje principal ya que su nombre superaba doblemente al de cualquier otro actor en las búsquedas por internet en China.
- Big Data: El “asesino” de los guionistas



■ ¿Conocéis a una persona,
dispositivo u organización que no
genere datos?



■ ¿Conocéis a una persona u organización que no utilice los datos?



■ ¿Conocéis a una persona u organización que no aproveche los datos que tiene a pesar de usarlos?





■ Contexto



- “Tengo un montón de datos y necesito guardarlos para que después aporte valor en mi empresa”
- Hace años, los datos eran para Business Operation
- Creación del departamento de Inteligencia de Negocio para ayudar a la toma de decisiones
- “El cliente más valioso es aquel que nunca pierdes”

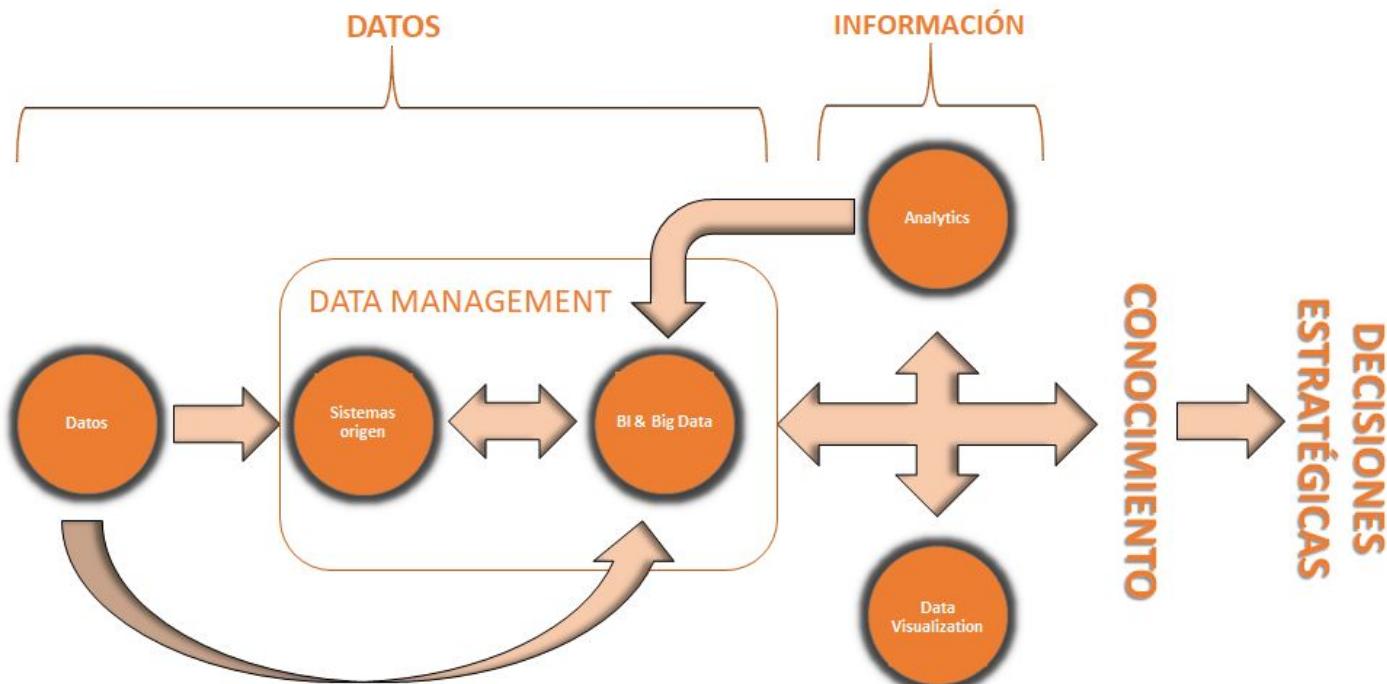




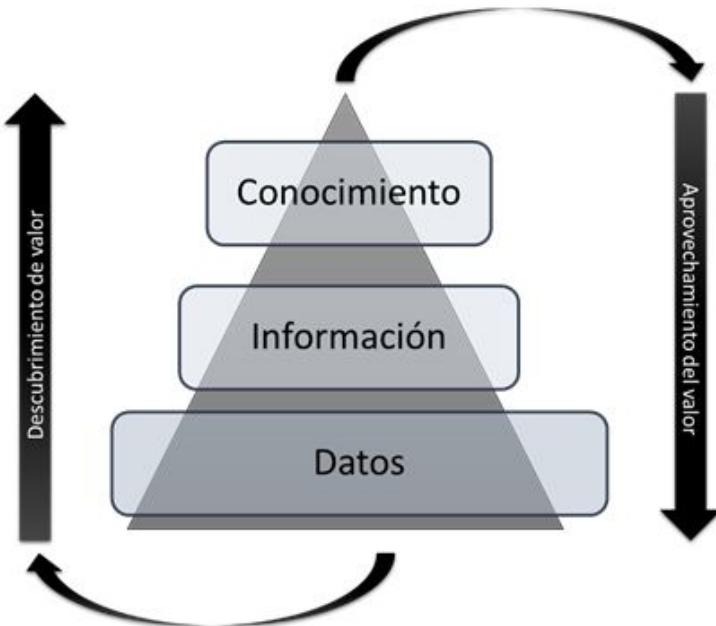
■ De los datos a las decisiones estratégicas



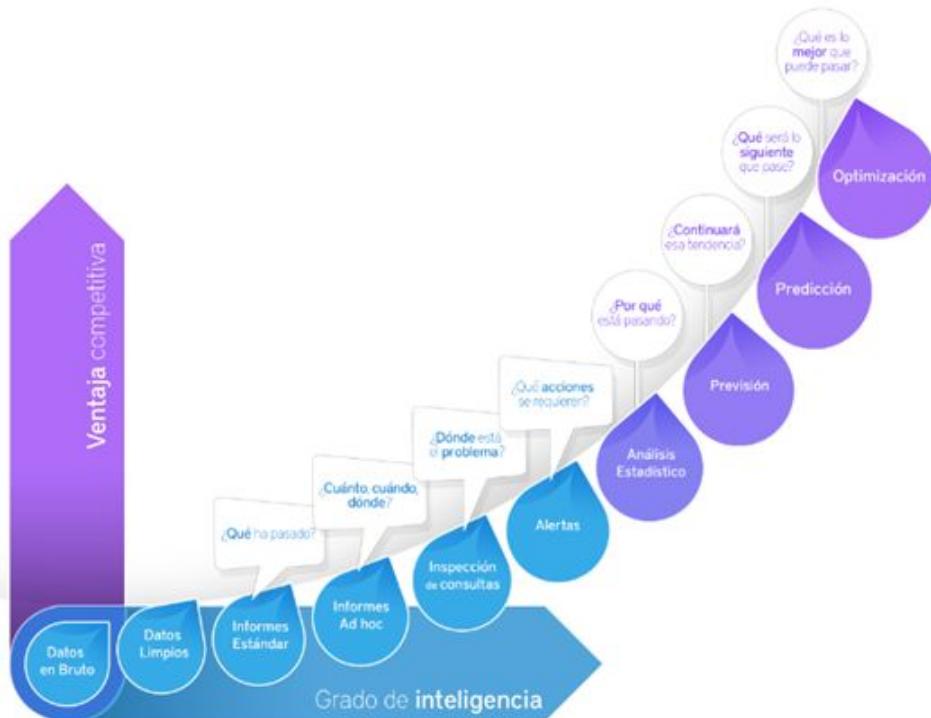
Ciclo de vida del dato



■ La pirámide del conocimiento



■ ¿Cuál es el DIKW real que necesitamos?



Data Information Knowledge Wisdom



Corporate Performance Management



■ La pirámide de la información



■ Organizaciones basadas en la estrategia (I)

PRINCIPIOS

- No se puede aplicar una estrategia que no se pueda describir.
- La actividad de la organización es algo más que la suma de las partes, necesita estar alineada con la estrategia.
- Hay que hacer que la estrategia sea el trabajo diario de todos los empleados. No se trata de dirigir desde arriba hacia abajo, sino de comunicar desde arriba hacia abajo.
- Hay que hacer de la estrategia un proceso continuo.
- Movilizar el cambio mediante el liderazgo de los directivos.



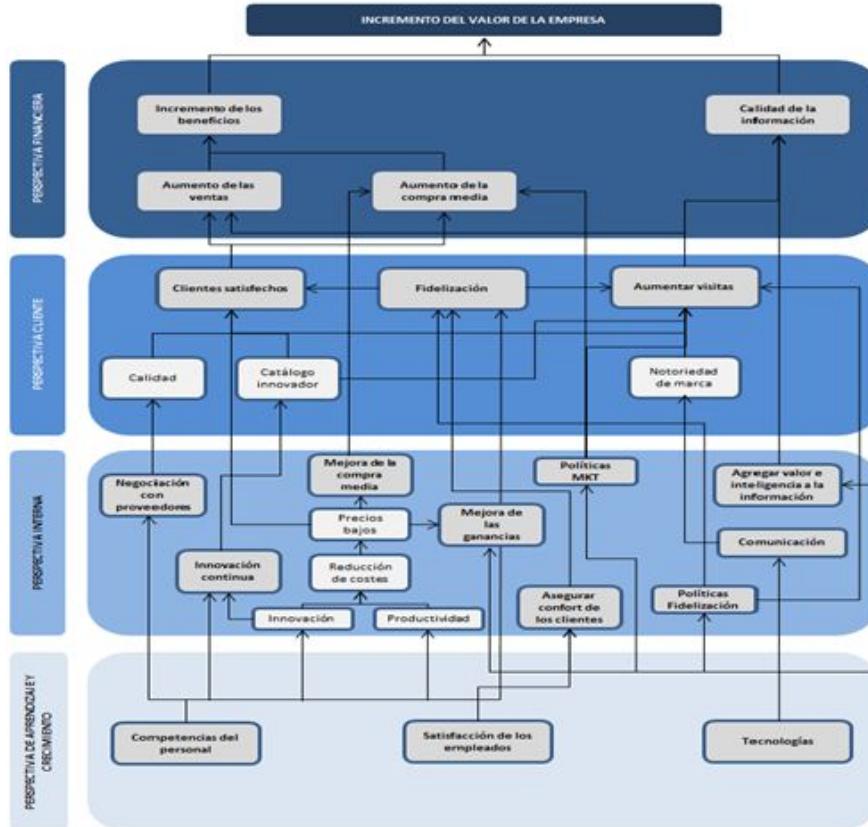
■ Organizaciones basadas en la estrategia (II)

CADENA DE CREACIÓN DE VALOR



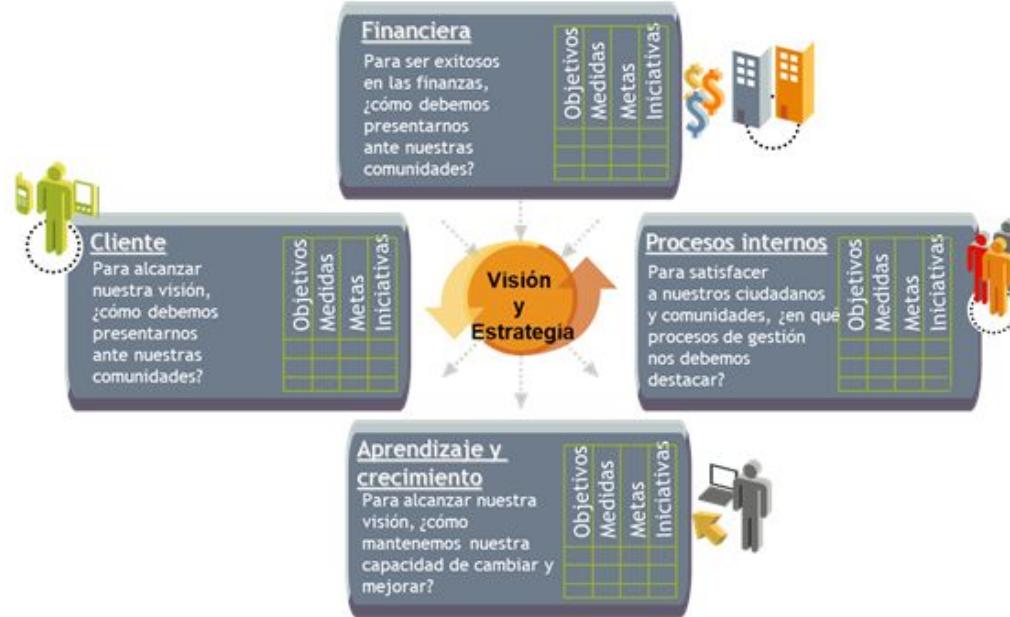
■ Organizaciones basadas en la estrategia (III)

MAPA ESTRATÉGICO



■ Organizaciones basadas en la estrategia (IV)

CUADRO DE MANDO INTEGRAL



■ Organizaciones basadas en la estrategia (V)

OMNICANALIDAD

- Conseguir una **experiencia de usuario de calidad** haciendo que la interacción del usuario sea perfecta en cualquiera de los canales, sin que note diferencias que hagan que un canal u otro lo considere mejor.
- **Responder sin demora** a la demanda del usuario en el momento que le surja, porque si no otro se nos adelantará.
- **Personalizar** la experiencia del cliente al máximo en función de sus intereses
- **Simplificar** al máximo los procesos para que al cliente le resulte fácil la interacción.



El cliente pasa a convertirse en el centro de todas las estrategias.

Los datos, su almacenamiento y su estudio cobran especial relevancia, puesto que es imprescindible para conocer al cliente y seguir sus interacciones



■ ¿Empresas en la que soléis comprar que utilicen la estrategia de la omnicanalidad?





Datos, datos, datos... ¿dónde están?



■ Lo más tradicional



Nivel I

Contaplus Elite: 0001 - Empresa de pruebas Sage Contabilidad - Ejercicio 2013

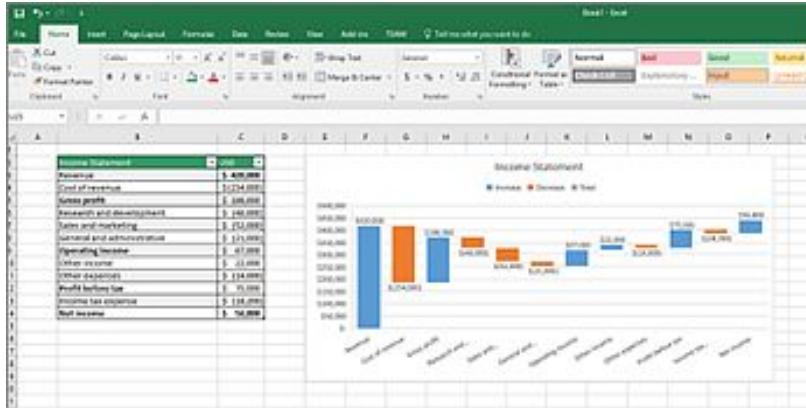
Sage

Empresa
Plan contable
Asientos
Tesorero
Presupuestos
Analitica
Segmentos
Inmovilizado
Informes
Babel
eFactura
Unidades

Inmovilizado

Saldo de subcuentas

	A	B	C	D	E	F	G	H
1	SALDOS DE SUBCUENTAS	1130000	1290000	1730000	1750000	2010000	2050000	2130000
2								
3								
4	Enero	0,00	0,00	0,00	0,00	0,00	0,00	0,00
5	Febrero	0,00	0,00	0,00	0,00	0,00	0,00	0,00
6	Marcio	0,00	0,00	0,00	0,00	0,00	0,00	0,00
7	Abri	0,00	0,00	0,00	0,00	0,00	0,00	0,00
8	Mayo	0,00	0,00	0,00	95.284,60	0,00	0,00	125.624,03
9	Junio	0,00	0,00	0,00	0,00	0,00	0,00	0,00
10	Julio	0,00	0,00	0,00	0,00	0,00	0,00	0,00
11	Agosto	0,00	0,00	0,00	0,00	0,00	0,00	0,00
12	Septiembre	0,00	0,00	0,00	0,00	0,00	0,00	0,00
13	Octubre	0,00	0,00	0,00	0,00	0,00	0,00	0,00
14	Noviembre	0,00	0,00	0,00	0,00	0,00	0,00	0,00
15	Diciembre	0,00	0,00	0,00	0,00	0,00	0,00	0,00



Bases de datos tradicionales



ORACLE

ODBC



PostgreSQL



TERADATA

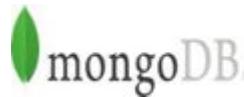


SYBASE®

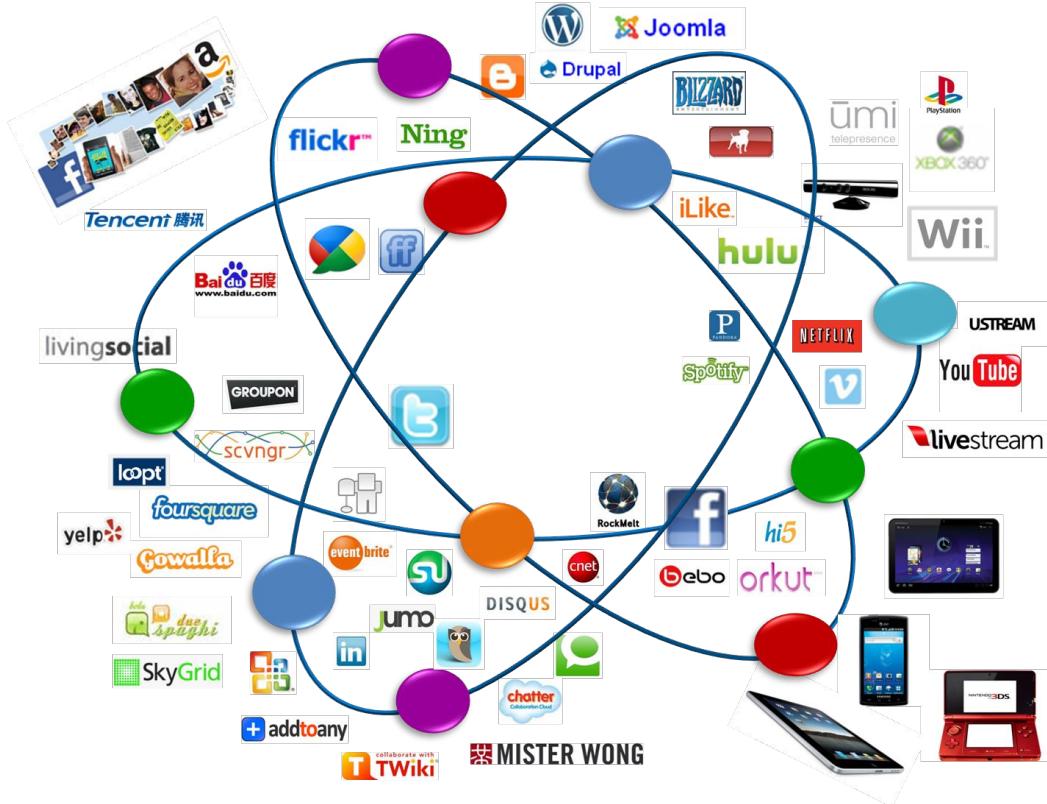
Informix®



Bases de datos NoSQL



Internet





A A A | ☰ ☆ ☰



Censo electoral

Sede electrónica

 Buscar

ES | E

iria Colaboradores en encuestas del INE



INEbase La información estadística

- * Agricultura y medio ambiente
- * Ciencia y tecnología
- * Demografía y población
- * Economía
- * Industria, energía y construcción
- * Mercado laboral
- * Nivel y condiciones de vida (IPC)
- * Servicios
- * Sociedad
- Síntesis estadística
- Información estadística europea
- Indicadores económicos. FMI



másINÉ
Revista digital



Estadísticas territoriales

El IPC en un clic

Apellidos y nombres

¿Cuántos habitantes...?

Elecciones al Parlamento de Cataluña de 21 de diciembre de 2017



Acceso a todos nuestros vídeos e infografías

Última hora

- 23 Nov 17. Estadística del taxi
- 23 Nov 17. Coyuntura Turística Hotelera. CTH
- 21 Nov 17. Entrada de Pedidos en la Industria. IEP
- 21 Nov 17. Índices de Cifras de Negocios en la Industria. ICN
- 21 Nov 17. Indicadores de actividad del sector servicios. IASS
- 20 Nov 17. Mujeres y hombres en España

Más noticias

Indicador	Periodo	Valor	Variación (%)
IPC	2017M10	102.668	1,6
EPA. Ocupados (miles)	2017T3	19.049,2	2,82
EPA. Tasa de paro	1 2017T3	16,38	-2,53
PIB	24 2017T3	—	3,1
Población total (miles)	3 2017	46.528,9	0,19

1. Valor en %. Variación: diferencia respecto a la tasa del mismo periodo del año anterior.

2. Índice de población encuestada, ref. 2010. Datos corregidos de efectos estacionales y de calendario.

3. Cifras de población a 1 de enero de 2017. Datos provisionales.

4. Datos avance

Aquí se muestra un gráfico en formato Flash

1. Para poder visualizarlo, se debe utilizar [este Visor Flash](#)
2. A continuación, en la ventana emergente que aparecerá, pulsar sobre el botón "Permitir".
3. La página se volverá a cargar mostrando el contenido en ese formato.

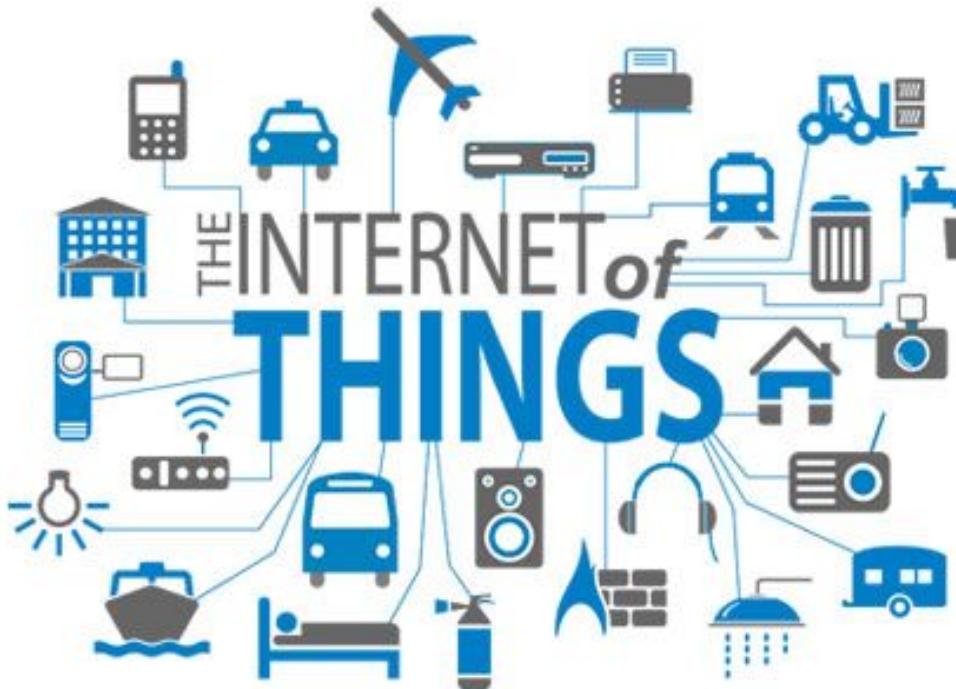


■ Otros datos interesantes



Protección de la identidad



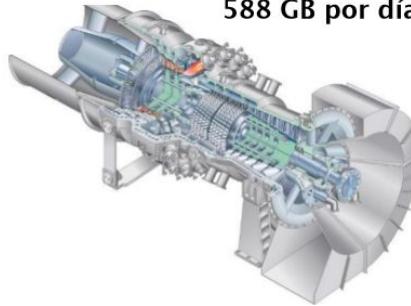


■ Los datos que no están en Internet



270 millones de
usuarios generan
100 GB por día

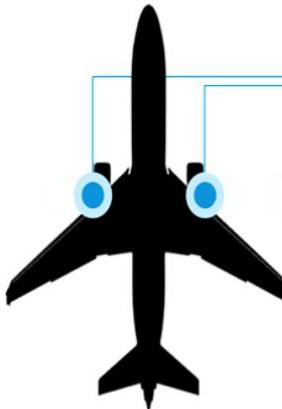
1 turbina genera
588 GB por día



■ Los datos que no están en Internet



Sensor data from a cross-country flight



20 TB

20 terabytes of
information per
engine every hour

×

2

×

6

×

28,537

×

365

twin-engine
Boeing 737

six-hour, cross-
country flight from
New York to Los
Angeles

of commercial
flights in the sky in
the United States on
any given day.

days in a year

= **2,499,841,200 TB**

(2,5 ZB al año)

Fuente: HP



■ Resumiendo

- Bases de datos propias
 - Internet
 - Máquinas y sensores
 - Dispositivos móviles
 - Administraciones públicas
 - Asociaciones y organizaciones
 - Open Data
 - ...
- ¿Alguno más?





■ ¿Orígenes de datos “singulares”?



■ La saliva!



■ ¿Por qué me vigilan, si yo no soy nadie?



<https://youtu.be/NPE7i8wuupk>





■ Data Management



Data Management

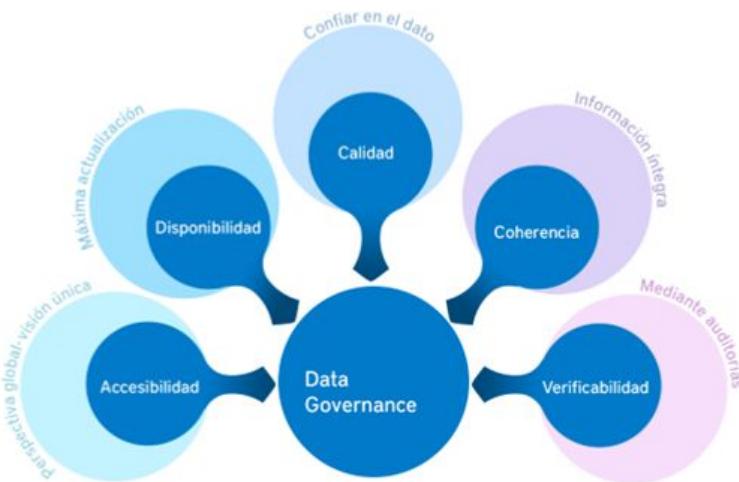


La gestión de los datos, que engloba el conjunto de todas las disciplinas relacionadas con gestionar los datos como un activo valioso, es lo que se conoce como **Data Management**.

El fin del Data Management es tener una visión unificada de todos los datos de la empresa así como de su entorno de negocio.



Data Governance



Cada vez disponemos de más información en las organizaciones, pero la manera de gestionarla no ha evolucionado al mismo ritmo en la mayoría de las ocasiones, lo que provoca que las organizaciones muchas veces **no saben ni qué datos tienen ni dónde encontrarlos**. Por tanto, se convierte en algo fundamental a día de hoy el ser capaz de obtener la **trazabilidad** de cualquier dato para saber qué caminos sigue la información en su flujo de negocio. De esto se hace cargo la disciplina del **Data Governance**.



Data Architecture



Los datos hay que **almacenarlos**, pero no de cualquier manera, sino siguiendo unas **reglas específicas**. De ello se encarga el **Data Architecture**, que describe los procesos, sistemas y organización humana necesarios para almacenar, acceder, mover y organizar los datos.



Data Modeling & Design



Una vez establecidas las reglas de cómo almacenar los datos, el siguiente paso es llevarlas a cabo.

De esto se encarga la especialidad de **Data Modeling & Design**, que modela y diseña las bases de datos, se encarga de su **implementación y soporte**, de manera que los datos puedan ser usados como recursos.



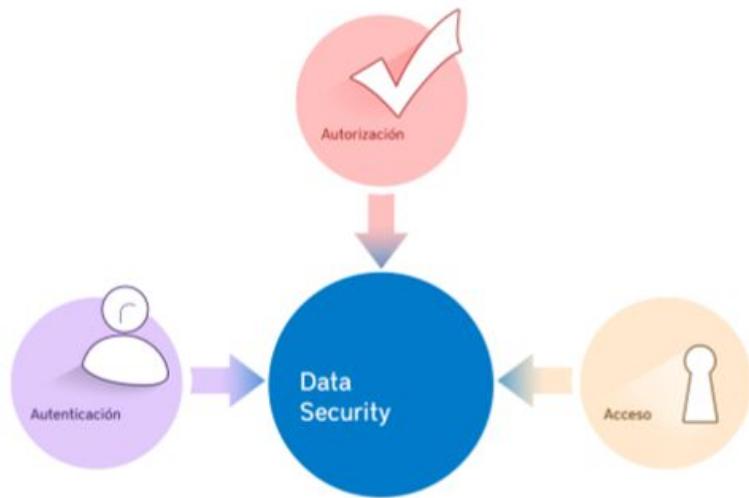
Data Storage



Un punto importante que hay que tener en cuenta en todo lo relacionado al almacenamiento es que hay que **evitar** que sea cada aplicación la que **decida** cómo guardar los datos y para ello el **Data Storage** tiene que ser el que vele por controlar cómo, cuándo y qué se almacena en cada uno de los sistemas.



Data Security



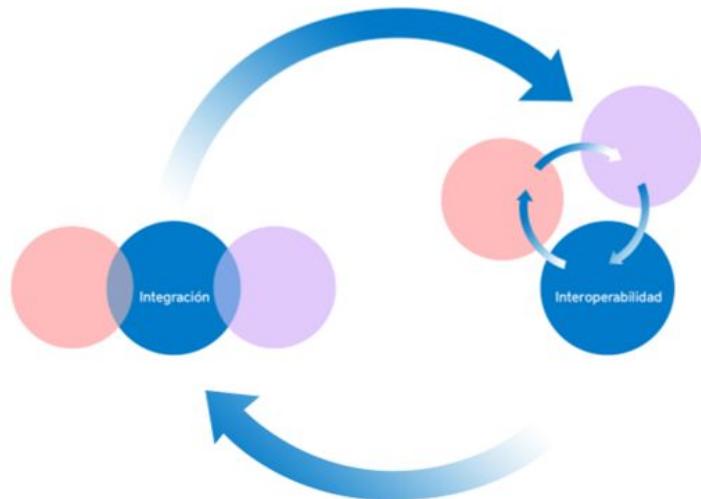
En los tiempos que corren no podemos olvidarnos tampoco del tema de la seguridad.

El módulo de **Data Security** se encarga de la **protección** de los **datos** contra el acceso, visualización, modificación o eliminación no autorizada, ya sea accidental, intencionada o maliciosa.

Se llevan a cabo todos los mecanismos, acciones y políticas necesarias para **garantizar** la **seguridad** de los datos en el entorno de la empresa.



■ Data Integration & Interoperability



En una organización **coexisten muchas aplicaciones y sistemas**, pero si no existe una comunicación entre todas ellas se está perdiendo eficacia y visión, lo que conlleva que la capacidad de decisión se vea disminuida.

La **integración** implica entender cómo la información se guarda en los distintos sistemas para que, al **interactuar** entre ellos, no sólo se conecten, sino que se entienda.



■ Master Data



Los **datos maestros** son utilizados en varios procesos de la organización, por lo que es de vital importancia **estandarizarlos** en los **diferentes sistemas** donde se almacenan.

El negocio necesita tener una visión de 360º sobre estos datos para tomar decisiones alineadas a la visión de la empresa.



■ Meta-Data



El módulo de **Meta-Data** es el que se encarga de integrar, controlar y proporcionar metadatos.

Los **metadatos** describen, etiquetan y caracterizan los datos a los que se refieren, haciendo más fácil su **interpretación** y **utilización**.



Data Quality

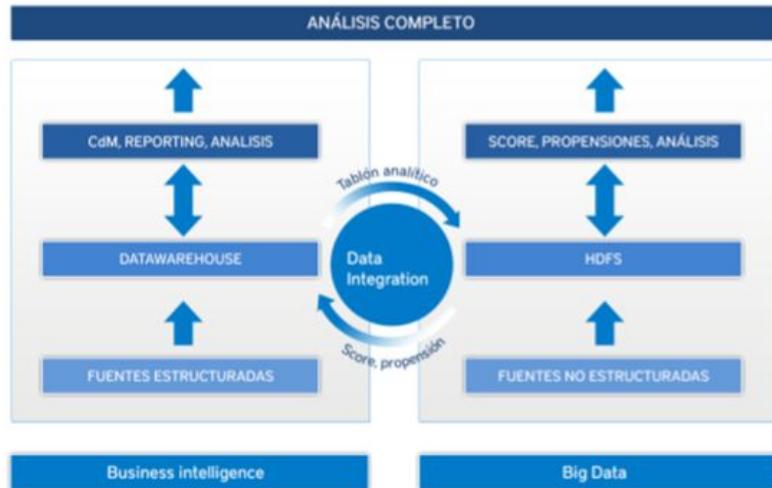


De gran importancia es la disciplina de **Data Quality**, que se encarga de definir, controlar y mejorar la calidad de los datos.

Es necesario que los **datos** sean de calidad, esto es, que sean **fiables, precisos, consistentes** y que proporcionen una **visión única**.



Business Intelligence & Big Data



La disciplina del **BI & Big Data** tiene como objetivo proporcionar una **visión integrada** de la información de la empresa para dar **apoyo** a la **toma de decisiones**.

Se ocupa de los datos tanto históricos como actuales y de las analíticas oportunas.





■ Business Intelligence



■ Contexto en el que surge

- Entorno que cambia rápidamente
- Cada vez hay más datos y menos tiempo
- Hay que hacer las preguntas adecuadas
- Se encargan los sistemas de información
- Soporte a las transacciones
- Distribuyen la información
- Soporte a la toma de decisiones



■ Business Intelligence

El **Business Intelligence** es un término paraguas que abarca los procesos, las herramientas y las tecnologías para convertir datos en información, información en conocimiento y planes para conducir de forma eficaz las actividades de los negocios.

The DataWarehouse Institute



■ El Business Intelligence se encarga de...

- Analizar la información de forma continua en el tiempo
- Explorar para comprender qué sucede
- Descubrir relaciones entre variables, tendencias, patrones.
- Almacenar información en tablas interrelacionadas en un DataWarehouse.
- Objeto de análisis y objetivo concreto.
- Comunicar los resultados y efectuar los cambios.



Bases de los sistemas de información

- **Percibir** lo que está ocurriendo en la compañía
- **Recordar** lo que ya ocurrió
- **Aprender** de lo recordado
- **Actuar** en base a lo aprendido



Sistemas operacionales vs informacionales

	SISTEMA OPERACIONAL	SISTEMA INFORMATICO
Fuente de datos	Datos operacionales, constituyen las fuentes originales de datos.	Datos consolidados y ficheros externos
Propósito del dato	Ejecución y control de las principales tareas de gestión.	Ayuda en el proceso de análisis, planificación y toma de decisiones.
Inserciones y actualizaciones	Inserciones y actualizaciones cortas y rápidas realizadas por los usuarios finales.	Actualizaciones periódicas y extensas para el refresco de datos.
Consultas	Estandarizadas y simples, con respuestas típicas de unos pocos registros.	Normalmente consultas complejas que requieren de agregaciones, cruces y filtrado de datos.
Velocidad de procesamiento	Normalmente muy rápidos.	Dependiente de la cantidad de datos almacenados y la complejidad de las consultas. El refresco del cálculo de algunas de ellas puede suponer horas.
Necesidades de espacio	Puede ser relativamente pequeño si los datos históricos se archivan.	Normalmente mayores necesidades de espacio, debido al almacenamiento de datos históricos.
Diseño de la base de datos	Altamente normalizada, con gran cantidad de tablas.	Normalmente desnormalizada, con menor número de tablas: modelos de datos en estrella, copo de nieve, etc.
Backup y recuperación	Backup imprescindible, ya que la pérdida de datos puede conllevar pérdida de dinero y consecuencias legales.	Backups realizados con cierta regularidad. En ocasiones la recuperación consiste en la nueva recarga de datos desde los sistemas OLTP.



Oportunidad de aprendizaje

- Cada interacción con el cliente, cada transacción, cada venta, cada acción comercial es una **oportunidad de aprendizaje**.
- Para aprender es necesario almacenar de manera conjunta, organizada, consistente y útil las fuentes de datos. Es lo que se conoce como el **Datawarehousing**.
- El Datawarehousing permite a las compañías recordar lo que se ha percibido.





■ Big Data



■ Las 4 V's del Big Data

- **Volumen:** hoy en día hablamos de procesar, por ejemplo, todo el contenido de Twitter porque quiero saber si en Twitter alguien está hablando bien o mal de mi marca. Estamos hablando de petabytes, que son miles de terabytes, que son millones de gigabytes. Hace muy pocos años este volumen de análisis era inabordable
- **Variedad:** trabajamos de forma masiva con texto escrito en redes sociales, con valores medidos por sensores de cualquier tipo, con imágenes grabadas por cámaras de vigilancia, etc. y esto hablando sólo de datos no estructurados, además de los datos estructurados en bases de datos de toda la vida.



■ Las 4 V's del Big Data

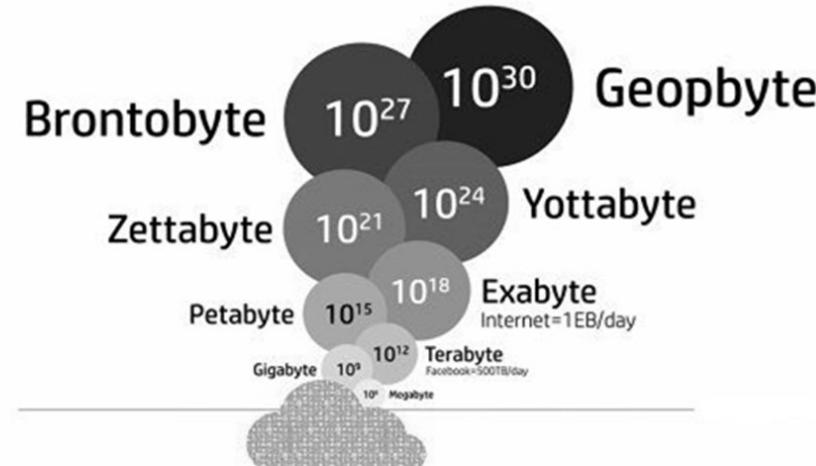
- **Velocidad:** aunque tenga que procesar millones de mensajes en Twitter, por ejemplo, quiero saber “ya” qué están diciendo los usuarios sobre mi empresa, sobre un producto que acabo de lanzar o sobre un acontecimiento que está ocurriendo ahora mismo. O quiero saber en tiempo real si mi cadena de montaje de coches robotizada está funcionando perfectamente, procesando la información de muchos miles de sensores funcionando al mismo tiempo.
- **Variabilidad** o datos “vagos”: no todos son datos concretos como los de un sensor o una medida técnica en unidades concretas. ¿Qué pasa con el texto que escribimos las personas, con sus dobles sentidos, bromas, lenguaje inventado, ironía y chistes? Eso también hay que saber interpretarlo, sobre todo en idiomas como el español; por cierto, los usuarios escriben y hablan en docenas de idiomas distintos, pero lo que quiero saber de ellos es siempre lo mismo, qué opinan de mí o si van a comprar mi producto o el de la competencia.



■ Las V's del Big Data



Datificación

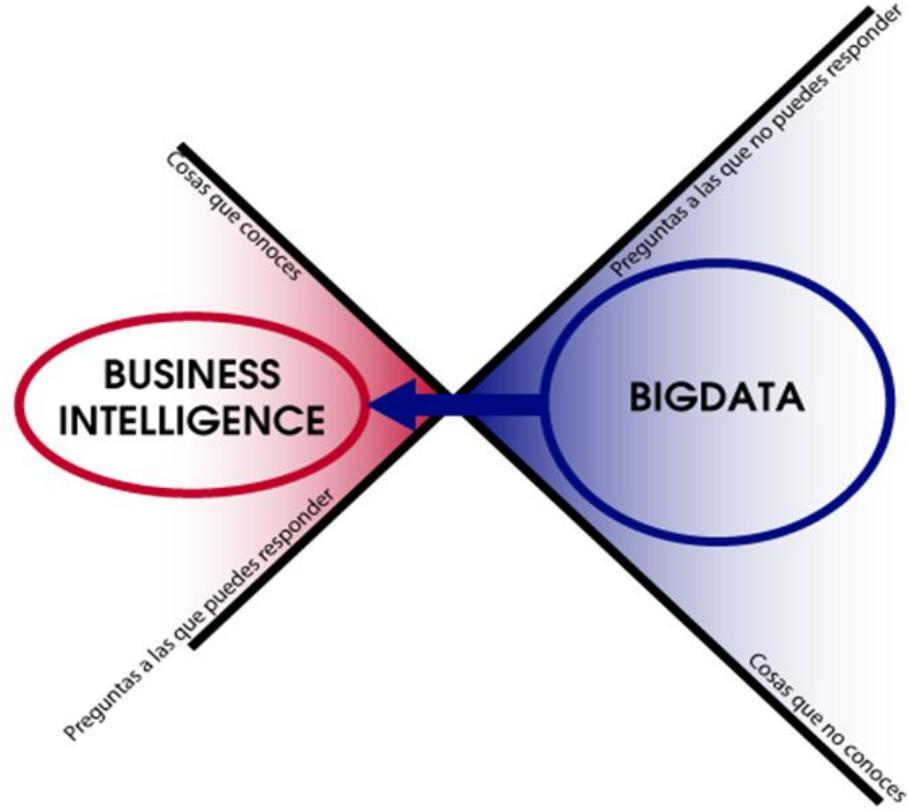


■ BI vs Big Data

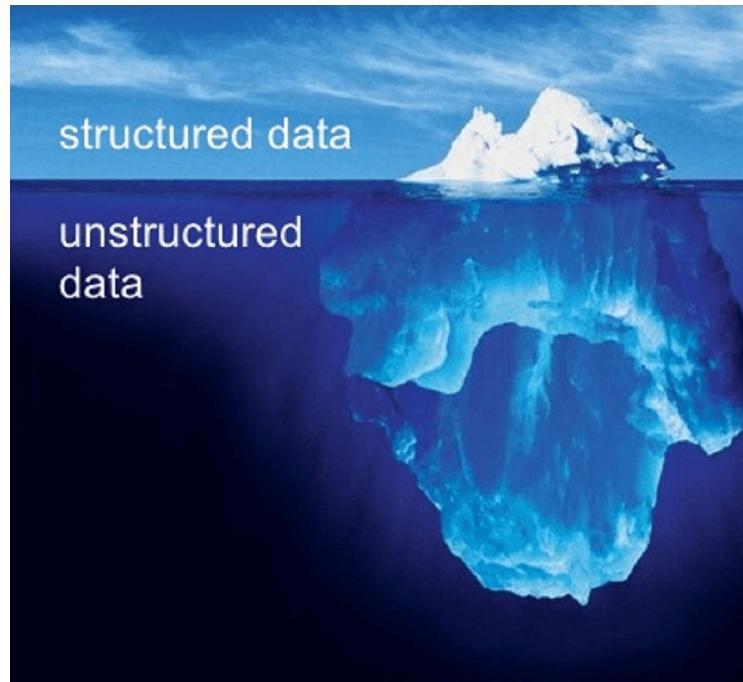
- Islas de información que no permiten vincular la información.
- La recopilación manual que muchas veces se hace.
- Las inconsistencias de los datos que provocan que, por ejemplo, la misma métrica, según dónde se mire, tenga un valor u otro.
- Los informes poco intuitivos.
- El retraso a la hora de obtener la información que se necesita.
- etc.

2012

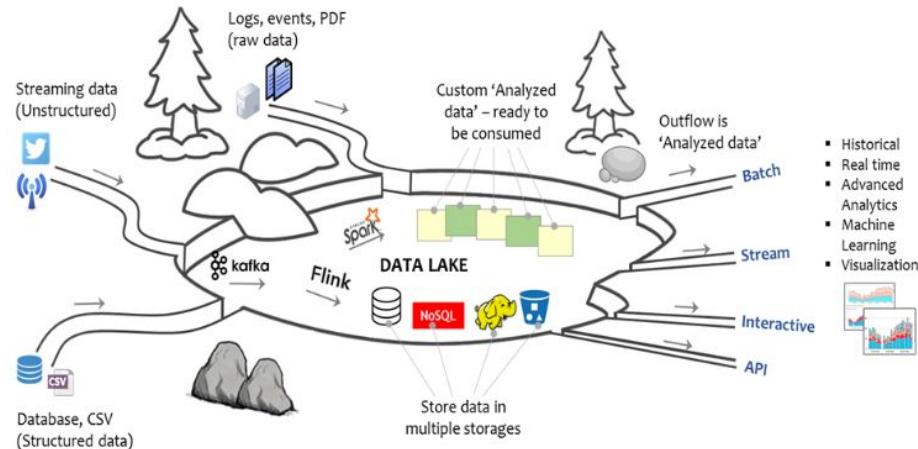
- Fuentes no estructuradas de datos



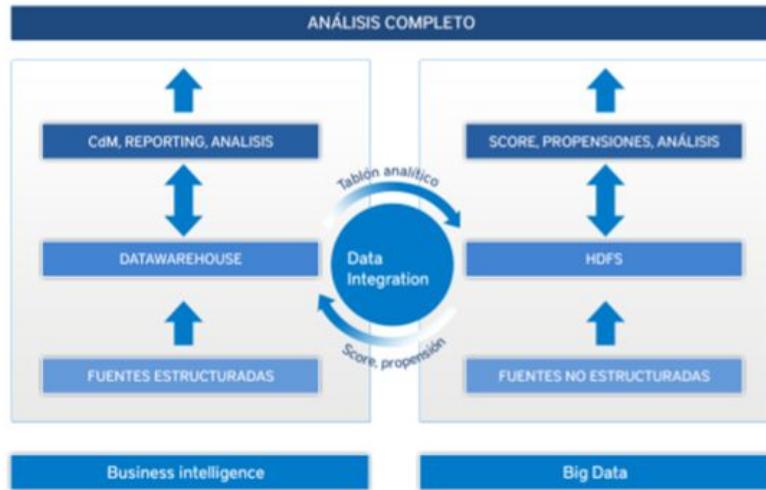
■ BI vs Big Data



Data Lake



■ BI & Big Data coexistiendo





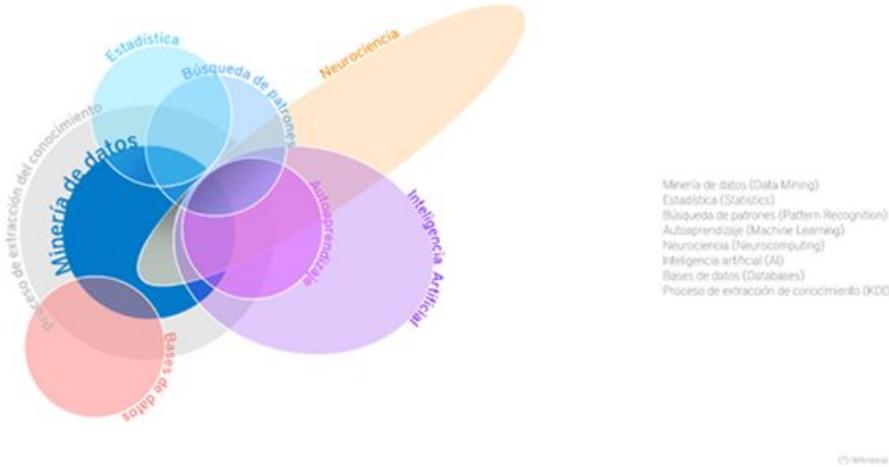
■ Enfoque multidisciplinar



■ La tercera variable



■ Disciplinas científicas



Se trata de muchas disciplinas que se entrecruzan entre sí y al final, del uso de algunas de ellas y sobre todo del uso de todas ellas como conjunto, obtenemos información de gran valor para la empresa



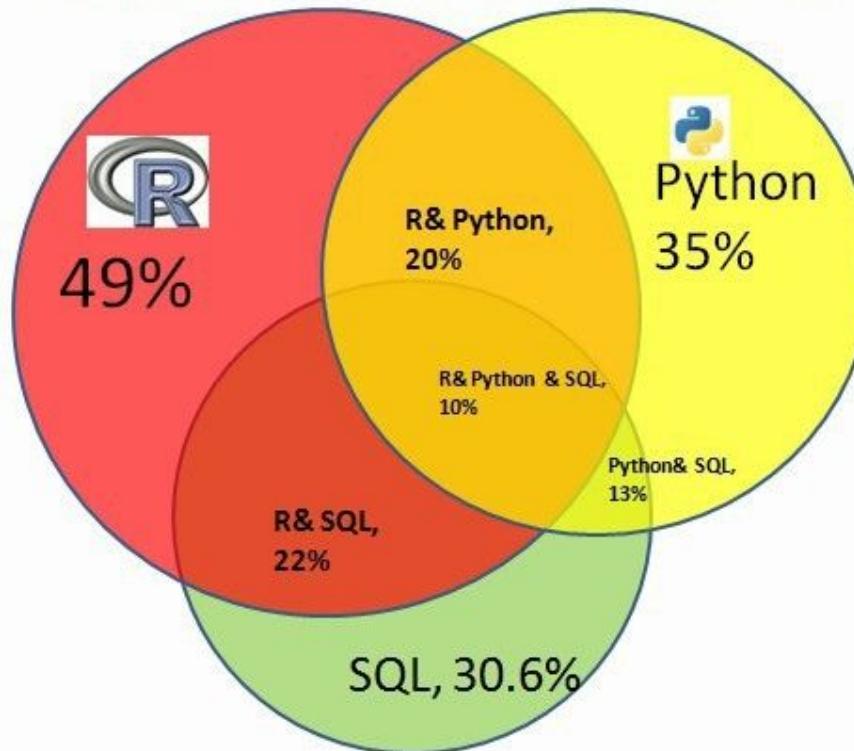
■ El científico de datos

- Popularmente: “Es un estadístico que trabaja en San Francisco”
- Josh Wills: “Persona que sabe más de estadística que cualquier programador y que a la vez sabe más de programación que cualquier estadístico”
- Sea como sea, es un profesional dedicado a analizar e interpretar grandes bases de datos, algo muy importante en la era digital en la que estamos inmersos
- Según Burtch Works:
 - El 32% de los científicos de datos en activo vienen del mundo de las matemáticas y la estadística
 - El 19% de la ingeniería informática
 - El 16% de otras ingenierías



Lenguajes de los Data Scientists

KDnuggets 2014 Poll: Languages used for Analytics/Data Mining



■ Las matemáticas...

- El Big Data tiene un lado “desagradecido”, el cómo almacenar y tratar los datos.
 - La fuerza del Big Data está en optimizar procesos muy complejos, cuanto mayor es el sistema, mayor valor podrá aportar.
 - Hay que construir la infraestructura flexible y escalable para que vaya creciendo a nuestra medida.
- Su lado más “agradecido”, y a la vez más peligroso, está basado en las matemáticas. Son la matemáticas de la efectividad, ya que si tenemos muchos y diversos datos y los combinamos con las reglas adecuadas, podremos encontrar los patrones, los valores atípicos que nos ayudarán a gestionar de forma potente nuestra estrategia de negocio.
- ¡Big Data entendido como que no se trata de los datos sino del impacto que tiene en el negocio!



■ ¿Cómo lo hacemos?

- Un primer paso del análisis puede ser usar datos pasados para predecir la posibilidad de qué va a pasar después. También el utilizar los datos del presente por medio de realtime para poder hacer un análisis predictivo potente en base a lo que ya conocemos.
- La idea es tener un sistema que aprenda a la vez que nosotros
- Se trata de encontrar los patrones, los perfiles, los modelos, las relaciones dentro de nuestros datos que nos permitan optimizar al máximo nuestro negocio.
- La anticipación y la pronta alerta será más y más importante para ganar ventaja competitiva. Estamos en un mundo en el que la ganancia marginal es la clave que nos diferencia de la competencia.
- El problema es que se hacen los análisis predictivos pero no se acompañan de las hipótesis para tomar acción. Se utilizan para entender comportamientos, para entender cómo funcionan nuestros esfuerzos, pero no como base de las acciones.



Perfiles profesionales Big Data Analytics



■ Big Data Dream Team

El éxito de un proyecto depende de:

- Equipo (talento)
- Tecnología
- Gestión del proyecto
- Modelo de negocio
- Alineamiento



■ Big Data Dream Team

- Hay escasez de profesionales
- Muy difícil encontrar al equipo ideal
- ¿Qué perfiles son necesarios?



■ Jefe de proyecto (Project Manager)

Facilita la innovación y el crecimiento del equipo.

Gestiona las expectativas de los interesados y asegura la calidad.

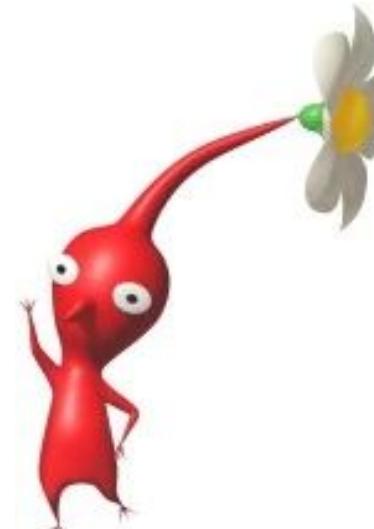
- Conocimientos TIC
- Metodologías ágiles. DataOps
- Experiencia en gestión de equipos



■ Arquitecto Big Data (Big Data Architect)

Diseña y optimiza las infraestructuras de forma que sean escalables y tolerantes a fallos.
Tiene que asegurar la disponibilidad y el correcto funcionamiento.

- Conocimientos TIC
- Arquitecturas DataWarehouse
- Administración SO
- Administración redes
- Experiencia en Cloud
- Experiencia en bases de datos SQL y NoSQL



■ Desarrollador Big Data (Big Data Developer)

Desarrolla e integra software que se ejecuta de un modo eficiente y procesa grandes cantidades de datos en arquitecturas Big Data.

- Conocimientos TIC
- Conocimiento de lenguajes de programación
- Experiencia en bases de datos SQL y NoSQL
- Experiencia con Hadoop y/o Spark
- Experiencia en cloud
- Aplicaciones en tiempo real



■ Científico de datos (Data Scientist)

Extrae el valor de los datos al aplicar algoritmos eficientes.

Descubre los insights que hay tras los datos.

- Experiencia en Data Mining
- Resolución analítica
- Buena base matemática/ estadística
- Experiencia en Aprendizaje automático
- Conocimientos de Inteligencia Artificial



■ Ingeniero de ciberseguridad (Cybersecurity engineer)

Capaz de mantener la seguridad de la información en entornos Big previniendo ataques.

- Experiencia en Seguridad
- Seguridad Front / Back
- Estudios especializados en seguridad informática
- Deseable conocimientos de Hacking ético



■ Analista de negocio (Business Analyst)

Identifica las necesidades del negocio de sus clientes y usuarios con de ayudarlos a determinar las soluciones a sus problemas.

- Expertise en consultoría
- Conocimientos TIC
- Deseable MBA



Citizen Data Scientist

- El **Citizen Data Scientist** es un híbrido entre un analista de negocio y un científico de datos.
- Gartner augura que el crecimiento de la demanda de este tipo de profesionales crecerá 5 veces más rápido que el de los científicos de datos.



Chief Data Officer (CDO)

Las principales **funciones** del CDO son:

- Crear una estrategia de gestión de la información de la empresa alineada a la estrategia del negocio y basada en la tecnología disponible en la organización.
- Trabajar de la mano del CIO y del CEO de la empresa.
- Adoctrinar al personal de la organización sobre la importancia de los datos para el éxito de la empresa.
- Asegurar el aprovechamiento de los datos mediante métricas u otros mecanismos.

Este directivo que está empezando a incluirse en las organizaciones, pero de un modo bastante lento hasta el momento, la firma de análisis Gartner estima que el **90% de las grandes empresas tendrán un CDO en su organigrama en 2019**.



Chief Protection Officer (CPO)

El mayor responsable del cumplimiento de las leyes de protección de datos.

Actualmente está cobrando especial importancia dada la cantidad inmensa de datos que se manejan en las organizaciones.



■ Otros involucrados...

- Cúpula directiva
- Expertos BI
- IT
- Propietarios de los datos
- Usuarios
- Abogados digitales
-



■ Perfiles profesionales

	Analista de negocio	Jefe de proyecto	Arquitecto Big Data	Desarrollador Big Data	Científico de datos	Ingeniero de Ciberseguridad
Modelo de negocio	●	●	●	●	●	●
Gestión del proyecto	●	●	●	●	●	●
Arquitectura	●	●	●	●	●	●
Desarrollo	●	●	●	●	●	●
Analítica	●	●	●	●	●	●
Ciberseguridad	●	●	●	●	●	●





■ Ciclo de vida del Big Data



Ciclo de vida del Big Data





■ Smart Data y Smart Visual Data



Smart Data

Con el Big Data, las empresas empezaron a guardar todos los datos disponibles a su alcance y luego trabajar sobre esos datos para extraer información.

El **Smart Data** es dar un paso más, es una evolución del Big Data. Ya no se trata de almacenar todos los datos disponibles, sino sólo aquéllos que seleccionados y bien procesados aporten el valor que la empresa necesita para llevar a cabo sus objetivos en el mínimo tiempo posible.



Smart Visual Data

Tras una selección inteligente de los datos de valor, lo interesante es poder mostrarlos de una forma visual y comprensible, como se viene haciendo desde siempre con los datos que ha habido en cada momento.

El Smart Visual Data siguen siendo dashboards o paneles de datos en tiempo real, como ocurría con el Big Data, y desde mucho antes con el Business Intelligence, pero ahora enfocado a los datos de valor, aunque con el mismo objetivo: que las personas que los manejen entiendan e interpreten mejor y más rápido los resultados que se obtienen.

Sin embargo, no nos equivoquemos pensando en los típicos informes tediosos y aburridos que se hacen desde hace años. Ahora hablamos de informes gráficos, con llamativos diseños y formados principalmente de gráficos / imágenes que son mucho más simples de entender a pesar de que tienen la misma información.



■ Smart Data y Smart Visual Data





■ Digital Transformation



■ Transformación Digital en la empresa

- La Transformación Digital es la reinención de las empresas a partir de las tecnologías digitales
- Es una transformación profunda dentro de las empresas, que afecta a toda la organización, algunas con parte del camino andado, y otras que aún no han empezado.
- Es una transformación a todos los niveles, en la que se busca aprovechar todo el potencial de la tecnología digital para encontrar oportunidades e impactar en la sociedad de un modo más eficiente.



■ Transformación Digital en los hogares y la sociedad (I)

- **Ahorro:**
 - Al hacer que todos los aparatos de nuestro hogar estén conectados a la red podremos monitorizarlos y saber qué está encendido y qué no en cada momento, así como encenderlos o apagarlos remotamente desde nuestro móvil.
 - Podremos consultar el consumo de cada uno en cada momento. Distribuyendo una serie de sensores y procesadores, podríamos automatizar el control de las ventanas, la temperatura del hogar, las luces, etc.
 - En definitiva, al estar conectado todo a internet sería posible controlar de forma inalámbrica lo que se nos antoje.



■ Transformación Digital en los hogares y la sociedad (II)

- **Seguridad:**
 - Hoy en día hay infinidad de compañías de seguridad que vigilan su hogar o lugar de trabajo.
 - Esto no se sustituirá, pero sí ayudará a perfeccionar el sistema con los objetos cotidianos y sensores que estarán conectados a la red.
 - En los últimos años compañías como Apple han desarrollado soluciones domésticas como HomeKit, que nos permite integrar interruptores, luces o enchufes inteligentes con el iPhone y gestionarlos incluso con Siri u otras 200 aplicaciones.
 - Hoy en día puedes programar las luces si te vas de vacaciones. Con este sistema podrías perfeccionar esa programación y ganar en tranquilidad.
 - Asimismo, ayudaría a detectar posibles robos ya que podríamos programar reacciones ante entradas no autorizadas, o incluso avisos si hay un pequeño incendio.



■ Transformación Digital en los hogares y la sociedad (III)

➤ **Confort:**

- Podemos controlar las persianas, luces, el aire acondicionado o cualquier aparato que esté conectado a la red.
- Gracias al smartphone o a la tablet estas personas verían cómo su autonomía aumenta con creces.



■ Industria 4.0

- IoT tendrá un efecto definitivo en la forma en que se producen las cosas, provocando junto a la impresión 3D, la IA y Big Data / Analytics una revolución industrial hacia fábricas inteligentes.
Con la primera revolución siendo la introducción de la máquina de vapor, la segunda trayendo la producción en serie en la línea de ensamblaje y la tercera el establecimiento de la robótica y automatización industrial, es ahora el momento para la cuarta revolución industrial.
- En estas fábricas inteligentes tendremos máquinas y componentes interconectados con sistemas embebidos, que llevarán a cabo funciones muy específicas monitorizadas en tiempo real, permitiendo una producción automática y autónoma:
 - Conocer en tiempo real y de forma concisa los costes y la demanda, pudiendo realizar una optimización activa
 - Monitorización de los dispositivos y elementos para su integración en un único entorno de control.



■ Fintech (I)

- **Marketing Automation:** La conexión entre máquinas y el flujo de información constante permitirán la generación de ofertas y campañas personalizadas para cada usuario en tiempo real y totalmente automatizadas, ahorrando en tiempo, recursos y costes.
- **Innovación continua:** Los servicios financieros dejarán de atender las demandas de los usuarios y comenzarán a adelantarse a sus propios deseos individuales. La posibilidad de monitorizar 24/7 el rendimiento de los servicios y la actividad de los clientes permitirá generar nuevas herramientas que atiendan a necesidades recién descubiertas, creen nuevas gamas de productos y permitan la mejora continua
- **Microproductos:** Las compañías aseguradoras y bancarias dejarán de ofrecer paquetes de productos y se centrarán en los llamados microproductos. Así, el permanente intercambio de datos permitirá ofrecer seguros de coche sólo para trayectos frecuentes o coberturas antirrobo solo para ubicaciones concretas, por ejemplo.



■ Fintech (II)

- **Sistemas de recompensa y aumento de la fidelización:** La información que aportemos a través de los dispositivos hiperconectados será clave para mejorar tanto los productos como sus tarifas. Así, por ejemplo, un estilo de vida saludable y ligado al cumplimiento de objetivos monitorizados a través de un wearable puede contribuir a una bajada en el importe de la prima, transformando a la aseguradora en un servicio tanto de prevención como de solución.
- **Fin de las contraseñas:** No hay duda de que toda esta revolución plantea grandes retos a la seguridad y a la privacidad de la información, pero también va a suponer una mejora sustancial a la cada vez más complicada gestión de contraseñas. Las herramientas de reconocimiento fácil y de voz eliminarán la necesidad de memorizar códigos y preguntas de seguridad y a la vez aumentarán la seguridad, convirtiéndonos a nosotros mismos en las únicas personas capaces de desbloquear un password.





■ Internet of Things



- El Internet of things (uno de los términos más populares de los últimos años en la industria tecnológica) es un concepto que se basa en la interconexión de cualquier producto con cualquier otro de su alrededor. Desde libros, termostatos, botiquines, paquetería, refrigeradores, etc. Tanto en entornos domésticos como en entornos urbanos, rurales e industriales.

- Un ejemplo práctico: gracias al IoT, nuestros botiquines serían capaces de medir las medicinas restantes en su interior y encargar al proveedor especificado las medicinas agotados. Así, nunca tendremos que preocuparnos de comprar las medicinas que necesitemos. El botiquín medirá continuamente el stock y actuará independientemente en consecuencia.



- También se pueden aplicar de forma más profunda al hogar. Distribuyendo una serie de sensores y procesadores, podríamos automatizar el control de la temperatura del hogar, las luces, apertura de puertas, etc. Y, al estar conectado todo el sistema a internet, también sería posible controlar de forma inalámbrica.

- Otro de los principales campos de aplicación del Internet of Things es las ciudades, haciéndolas más inteligentes y eficientes. Por ejemplo: evitaría esperar ante un semáforo en rojo, ya que estos podrían conectarse a un circuito de cámaras distribuidas por la ciudad e identificar el nivel de tráfico y de movimiento de personas, evitando así esas absurdas esperas en las zonas de escaso movimiento.



- Este mundo donde todos los objetos cotidianos están interconectados y pueden intercambiar información será el motor de un cambio disruptivo para las sociedades actuales. La trascendencia de IoT será mayor incluso que la era digital.
- No obstante, los requisitos tecnológicos suponen un gran reto a solventar, y aunque el interés actual es máximo la adopción general será en unos 5 años.
- Para suplir la demanda actual el internet de las cosas debería codificar de 50 a 100,000 mil millones de objetos y seguir el movimiento de estos; puesto que se calcula que todo ser humano está rodeado de por lo menos de 1,000 a 5,000 objetos. Aún estamos muy lejos de estas cifras, diferentes estudios estiman que en 2020 habrá en el mundo entre 26 y 30 mil millones de dispositivos inalámbricos conectados a Internet.





Big Data vs Analytics vs Transformación Digital vs IoT



■ Para terminar...Dejemos volar la imaginación



https://www.youtube.com/watch?v=qSOY1I_Nr2Q&t=

