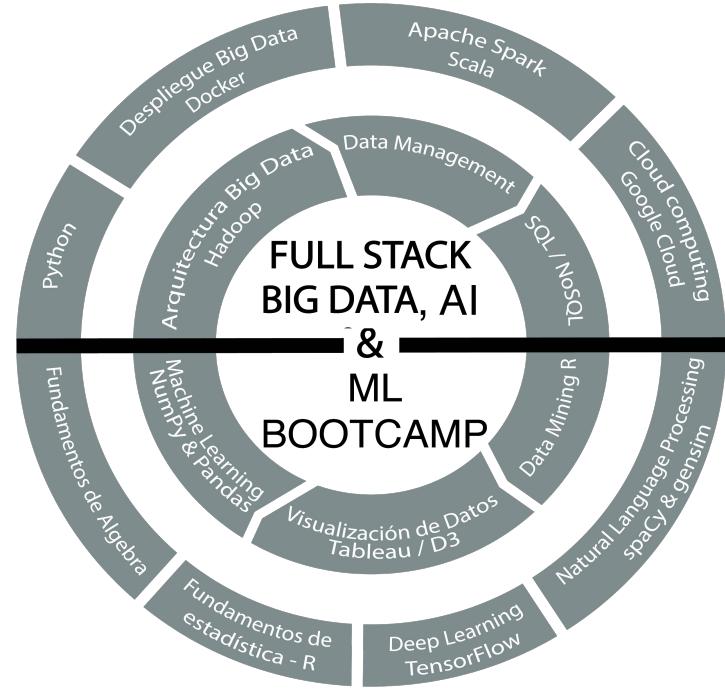


Full Stack Big Data, IA & Machine Learning Bootcamp



Data 101



¿Quiénes somos?



■ ¿Quién soy yo?



<https://www.linkedin.com/in/sandra-navarro-nieto/>



sandra.navarro@contextual.team





Guía





■ Equipo



- i7
- 16Gb RAM
- 512Gb HDD
- Dos pantallas





■ Calendario



URL Pública:

<https://calendar.google.com/calendar/u/2?cid=b3Jibjl1MmFvMmpwZWQ3Z3Vsb2FkcTFoYjRAZ3JvdXAuY2FsZW5kYXluZ29vZ2xILmNvbQ>

<https://calendar.google.com/calendar/embed?src=orbn9u2ao2jped7guloadq1hb4%40group.calendar.google.com&ctz=Europe%2FMadrid>

iCal:

<https://calendar.google.com/calendar/ical/orbn9u2ao2jped7guloadq1hb4%40group.calendar.google.com/public/basic.ics>





■ Slack



- Es nuestro medio normal de comunicación
- Hay que ponerse foto para que todos nos pongamos cara
- Nos tenemos que dar de alta en todos los canales que tengamos disponibles
- Las dudas siempre en los canales de las asignaturas, no en mensajes privados a los profes. “La duda de uno casi seguro que es la de muchos más”
- Los días no laborables por supuesto podéis preguntar dudas, pero tened en cuenta que es posible que el profe no os conteste hasta el siguiente día laborable porque no le sea posible



- Cada canal tiene una función y debemos publicar teniendo esto en cuenta:
- Canales de las asignaturas: para dudas sobre los contenidos de las asignaturas
 - info-oficial: Aquí publicamos información oficial desde keepcoding
 - enlaces-sesiones
 - keynotes-masterclasses: info sobre estas sesiones en particular
 - off-topic: Para hablar de Big Data sin ser contenido de la asignatura
 - frikadas: para hablar de lo que se quiera



- Slack NO es una herramienta de almacenamiento y hay mensajes que pueden desaparecer.
- Zona “chinchetas”: información destacada. Esa información fijada en dichas zonas os permitirá en modo rápido, acceder a todas las publicaciones de interés, evitando que queden imposibles de localizar.



■ Registrarse en Slack

- Nos debería haber llegado una invitación por email.
- Tenemos que conectarnos al espacio de trabajo 8bootcampbdmljun2021.slack.com
- Una vez conectados, tenemos que darnos de alta en todos los canales que tengamos disponibles (pinchando sobre la palabra “Canales”).
- Está disponible para ordenador, tablet y PC





■ Comentarios y sugerencias



- En el mismo momento que la detectéis
- Formulario web:
[https://docs.google.com/forms/d/1rhM8dp5v-2hLJ7wdN-tHPOybvM5Ps8D
bnKD_alzkjl0/viewform?edit_requested=true](https://docs.google.com/forms/d/1rhM8dp5v-2hLJ7wdN-tHPOybvM5Ps8DbnKD_alzkjl0/viewform?edit_requested=true)
- En la encuesta final de cada módulo.
Es importante de cara a poder mejorar, aunque os recomendamos que cualquier mejora que vayáis detectando, se la hagáis llegar rápidamente al instructor para que intente solucionarlo en las siguientes clases que tengáis con él.





■ Módulos y masterclasses





■ ¿Cómo funcionan los módulos?



- Módulos:
 - Tienen clases por videoconferencia
 - Tienen práctica a entregar
 - Tienen “semana de asimilación”, que es un tiempo extra para realizar la práctica

- Masterclasses
 - Son clases extraordinarias como complemento por videoconferencia
 - No suelen tener práctica asociada
 - Según el grado de actualización, pueden ser en vivo o grabadas para que os organicéis mejor el momento en que verlas.



- Plataforma Keepcoding:
 - Clases grabadas: disponibles hasta dos años después de entregar el proyecto final
 - Cursos Online: disponibles hasta la fecha de entrega del proyecto final





■ Gitlab



- En Gitlab se subirán todos los materiales por parte de los profesores.
- La ruta:
<https://gitlab.keepcoding.io/keepcoding-bootcamps/full-stack-big-data-ai-y-ml-viii>





■ Videoconferencias



- Empiezan siempre a las 20:30h.
 - En general duran 4h:
 - 20:30h-22:30h
 - 22:45h-00:30h
 - Si duran menos de 4h siguen empezando a la misma hora pero acaban antes
- Al principio de la videoconferencia pueden estar abiertos todos los micrófonos.
 - Cerradlos según os los abran
 - Cuando queráis hablar simplemente tenéis que activarlo e interrumpir





■ Prácticas de módulo



- Cada Módulo tendrá una práctica final para aplicar los conocimientos adquiridos durante el módulo.



- Su entrega es obligatoria y debe ser hecha y entregada en los plazos definidos por KeepCoding.
- Serán evaluadas individualmente, con APTO y NO APTO, los cuales saldrán impresos en tu certificado final del Bootcamp. Estos resultados serán tenidos en cuenta para la evaluación final.
- Durante el periodo de asimilación del bootcamp se podrán plantear las dudas al profesor en los canales de comunicación establecidos para ello.



- La fecha de entrega de cada práctica estarán definidos en el calendario del Bootcamp y como regla será el último domingo del módulo, a las 23:50 hora de España Peninsular.
- La ÚNICA vía de entrega de prácticas admitida es el formulario diseñado a tal efecto:
<https://forms.gle/VhyqMhp2j3CbvNvx9>
- Cualquier envío hecho a instructores por mail, slack, skype , etc.. no quedará registrado y por tanto quedará como 'no existente' y no se calificará.



- De igual modo, si se obtuviera en alguno de los módulos calificación de NO APTO y el instructor nos solicitara alguna mejora/actualización de proyecto, procederemos a nuevo envío mediante el mismo formulario (sea o no coincidente el link de github suministrado).
- Cuando se realicen entregas, se deje siempre solución a calificar en rama master. En caso de encontrarse en otra rama no se tendrá en cuenta
- Se podrá hacer excepcionalmente una re-entrega, (ya sea por fuera de plazo o por haber obtenido un NO APTO en primera instancia), HASTA LA FECHA LÍMITE DE RE-ENTREGA fijada para el bootcamp. Si está re-entrega entra dentro del plazo será evaluada por el profesor en la fase final del bootcamp.



- La corrección se hace individual y es enviada el alumno por email por el Profesor
- Intentad no dejarlas para el final.
- Si no se dice lo contrario, podéis entregarlas en el lugar que estiméis oportuno: Gitlab, Github, etc.
- No subirlas comprimidas en zip, rar, etc.
- Las prácticas que sean en notebooks, a ser posible entregarlas en google drive



- Las posibles calificaciones para cada uno de los módulos:
- APTO:
 - Práctica entregada y aprobada en fecha de entrega de práctica del módulo
 - APTO EN SEGUNDA REVISIÓN:
 - Práctica entregada en fecha, aunque no superada en el primer intento
 - Aprobada en una nueva entrega antes de la “Fecha de re-entrega de prácticas” (en el calendario, antes del comienzo del período del TFB)
 - APTO FUERA DE PLAZO:
 - Práctica entregada y aprobada antes de la “Fecha de re-entrega de prácticas”



- NO APTO:
 - Práctica entregada y no aprobada en fecha de entrega de práctica del módulo
 - No se produce una re-entrega de la práctica
- NO APTO EN SEGUNDA REVISIÓN:
 - Práctica entregada en fecha, aunque no superada en el primer intento
 - No aprobada en una nueva entrega antes de la “Fecha de re-entrega de prácticas”
- NO APTO FUERA DE PLAZO:
 - Práctica entregada antes de la “Fecha de re-entrega de prácticas” y no aprobada



- NO ENTREGADO

- Si no se produce ninguna entrega de la práctica

- Importante!!

Una entrega de práctica vacía o sin un mínimo lógico se considerará no entregada.





■ Proyecto final



- El proyecto final del Bootcamp será generalmente grupal, formado por alumnos con el mismo nivel de implicación. Esto se determinará en función del número de prácticas APTAS.
- Castas:
 - Guardia Imperial: más del 80% de las prácticas aptas (actualmente 9 módulos)
 - Tropa de asalto: más del 50% de las prácticas aptas (actualmente 6 módulos)
 - Milicia Gungan: menos del 50% de las prácticas aptas (5 o menos)



- El proyecto final del Bootcamp será elaborado bajo una idea propia del equipo.
- El alumno dispondrá de algunas semanas para poder perfeccionar dicho proyecto, una vez terminada las clases.
- El alumno tendrá que tener un mínimo de 60% de las prácticas por módulo entregadas, para que su Proyecto final sea considerado como ENTREGADO.
- En principio se realizará de manera grupal.



- El proyecto final será entregado para su evaluación el día marcado en el calendario para ello y, posteriormente, habrá una presentación online para que sea evaluada por el cuerpo de profesores.
- Posteriormente, podrá ser presentado voluntariamente en el Closing.
- Para aprovechar al máximo el tiempo que tenéis disponible para realizar el Trabajo Final de Bootcamp, sería recomendable haber hecho de la plataforma:
 - Curso de SCRUM y Agile





Diploma



- En el diploma aparecerán tres posibles calificaciones finales en función de las prácticas presentadas y de la entrega del TFB:
- Ha superado con éxito:
 - Aprobadas al menos el 80% de las prácticas
 - Aprobado el TFB
 - Ha realizado:
 - Aprobadas menos del 80% de las prácticas
 - Presentado y aprobado el TFB
 - Ha realizado:
 - Aprobadas al menos el 80% de las prácticas
 - No presentado y/o no aprobado el TFB
 - Ha participado:
 - En el resto de casos





■ Recomendaciones generales



- Sed participativos en las videoconferencias. Interrumpid siempre que lo veais necesario
- La pregunta más tonta puede ser la más interesante
- Exprimid a los profes, tanto en clase como por Slack





■ Recomendaciones del prebootcamp



Para aprovechar al máximo los módulos, sería recomendable haber hecho del prebootcamp:

- Antes de “Exploración y visualización de datos”:
 - Curso de git
- Antes de “Modern exploration and visualization”:
 - Curso de Javascript
 - Curso de Linux
- Antes de “Datawarehouse”:
 - Curso de SQL
- Antes de “Big Data Architecture”:
 - Curso de Python
 - Curso de Docker
- Antes de “Big Data Processing”:
 - Curso de Scala
- Antes del TFB:
 - Curso de Scrum y Agile





Data 101



Motivación



- El big data está teniendo un crecimiento anual superior al 12%
- 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

ANALYTICS

USED
CREATED
NOW
CAPACITY

LARGER
TENS
HUNDREDS
RECORDS

CONTINUES
COMPLEX

SET
USE
INDEXING
CITATION

ORGANIZATIONS
RELATIONAL
SOCIAL
UBIQUITOUS

BURIED
WORLDS
NETWORKS
DATABASES

SEARCH

BIG DATA

The word cloud includes the following words:

- Tools: Disk, Target, Applied, Sensor, Shared, Definition, Current, Zettabytes.
- Internet: Practitioners, Business, Distributed, Sets, Genomics, Complexity.
- Management: Describing, Radio-Frequency, Elapsed, Petabytes, Systems, Tolerable.
- Storage: Storage, Parallel, SAN, Massively, Grow, Quality.
- Technologies: Moving, Wireless, Within, Thought, Capture, Case, Performance, Research, Logs, Related, Computing, Cost.
- Information: Archives, Manage, Biogeochemical, Using, One, Large, Process, Presentations, Combat.

The main title 'Expresiones de moda' is written in large orange letters at the top, and 'Solape' is written in large orange letters at the bottom.



Buzzwords

Término paraguas

Big
Data

Analytics

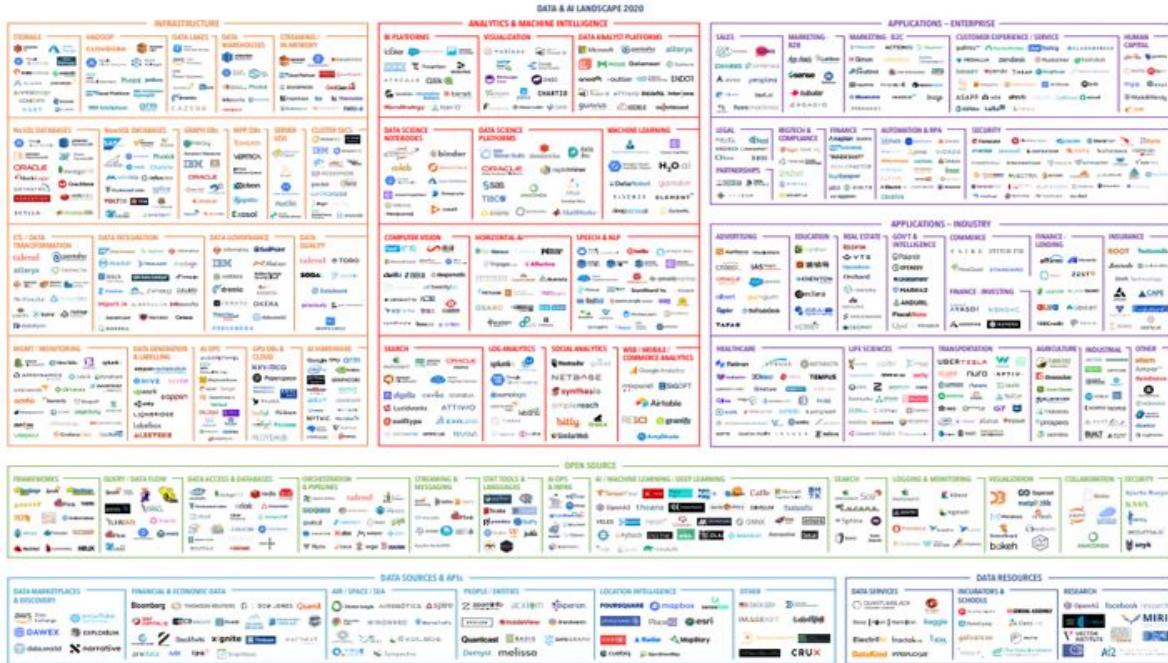


■ Luego...

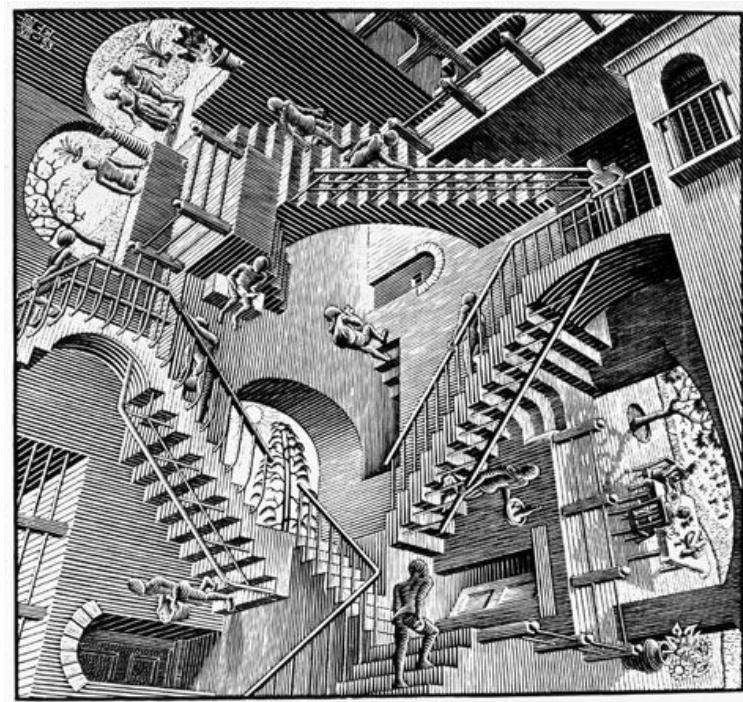
Big Data Analytics



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



■ Me perdí...

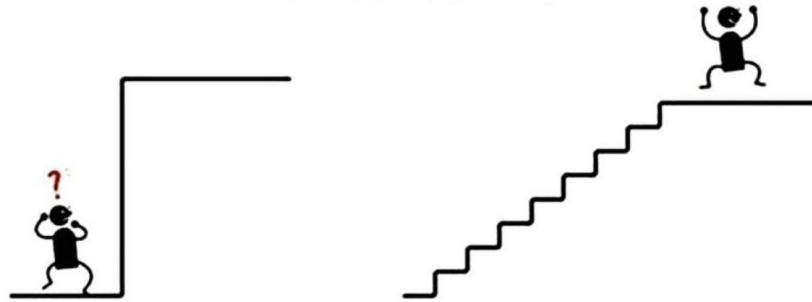


Estás más perdido que un muelle cayendo por las escaleras de Hogwarts



■ ¿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”

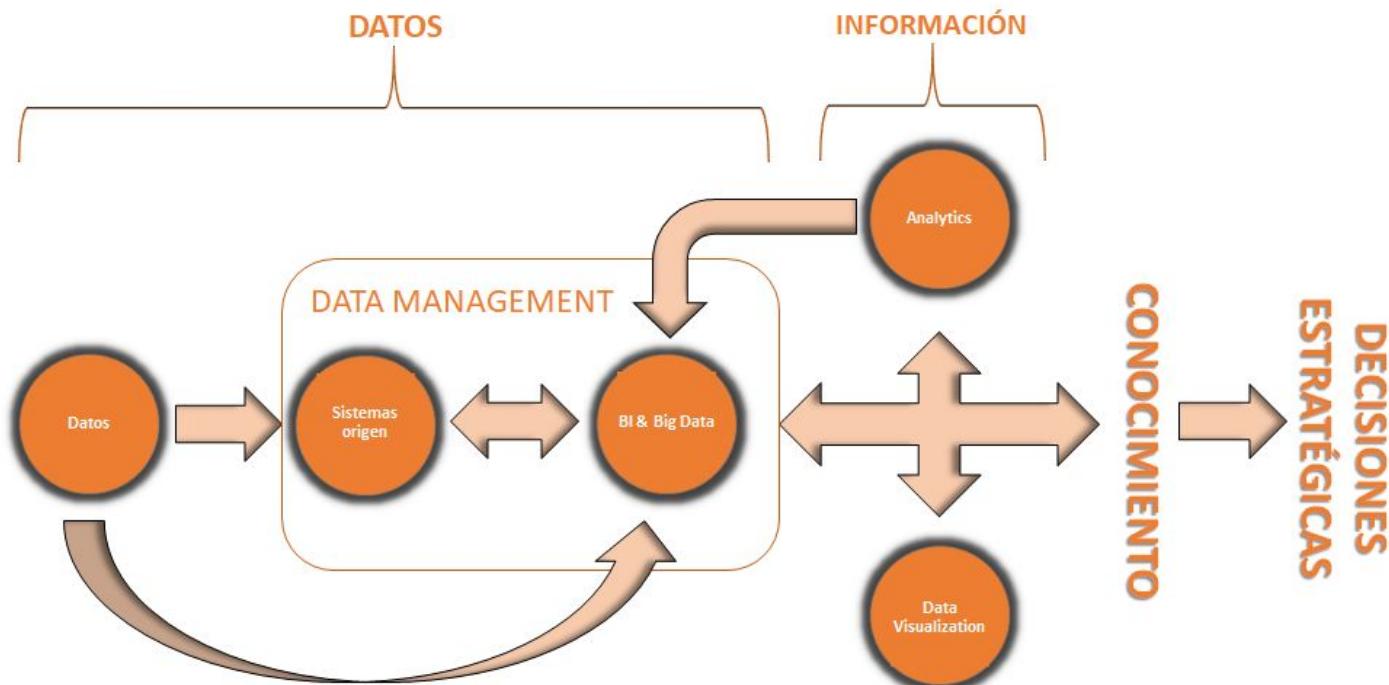




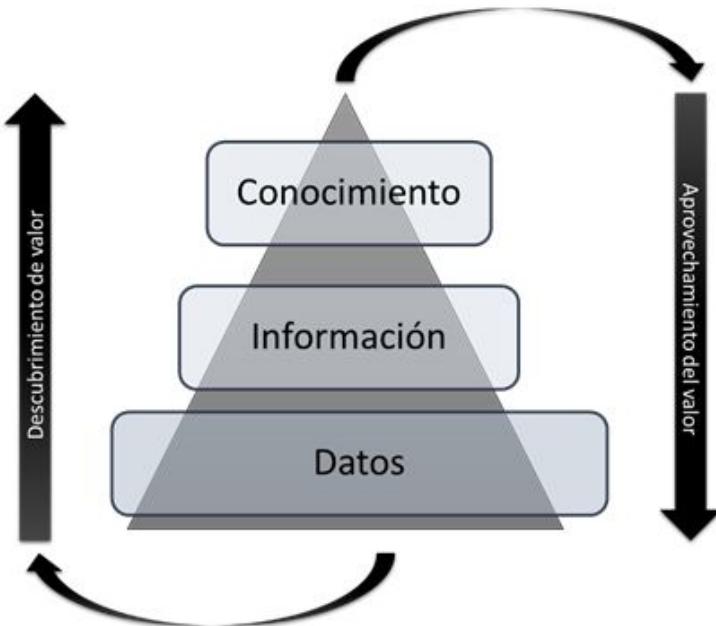
■ De los datos a las decisiones estratégicas



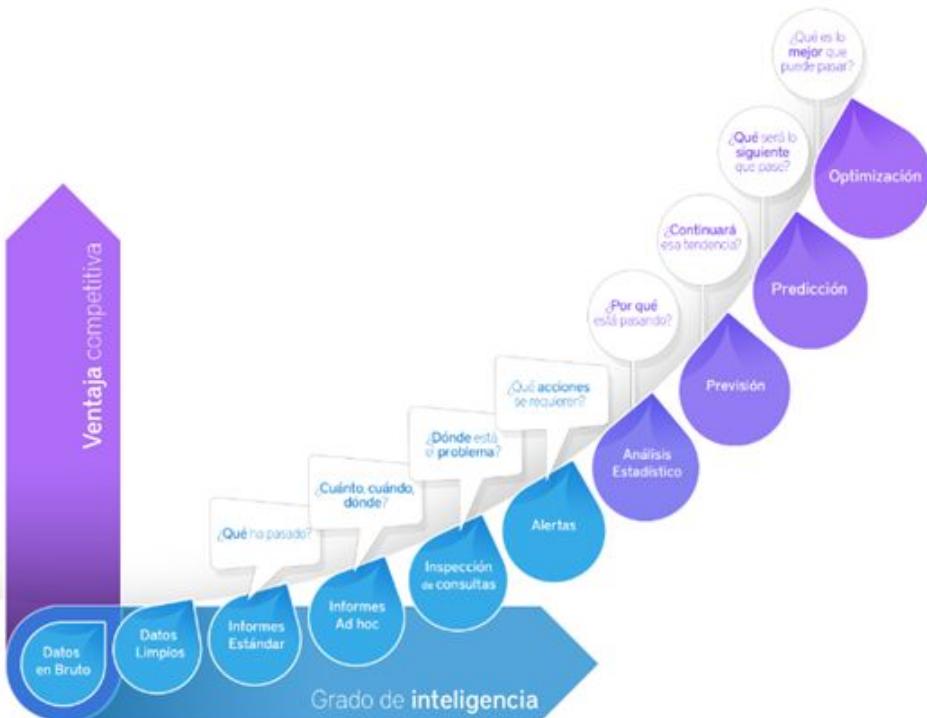
Ciclo de vida del dato



■ 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.



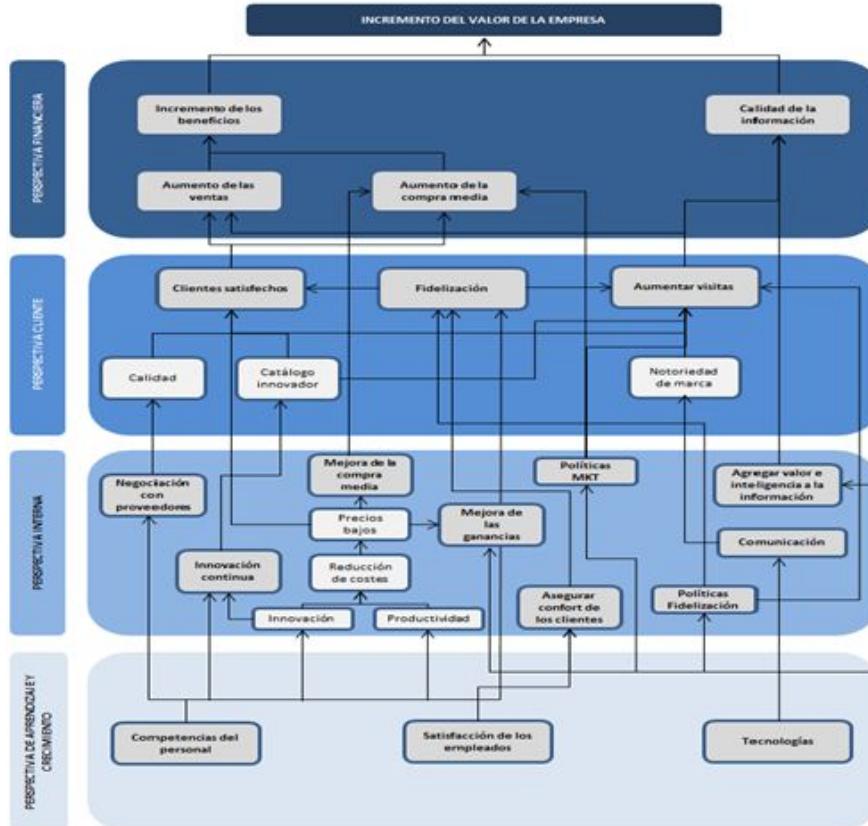
■ Organizaciones basadas en la estrategia (II)

CADENA DE CREACIÓN DE VALOR



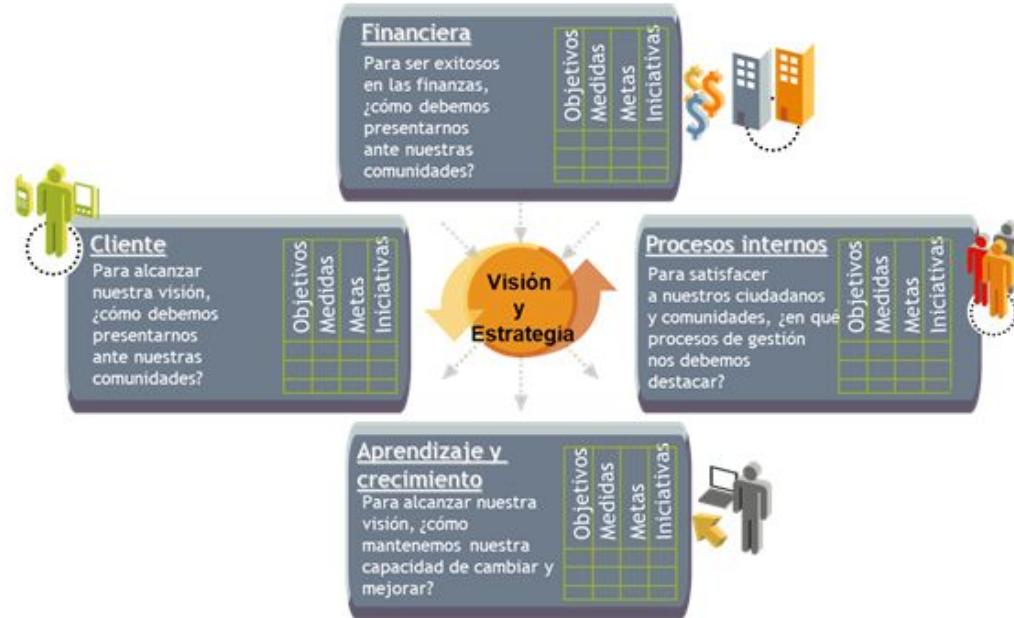
■ Organizaciones basadas en la estrategia (III)

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?





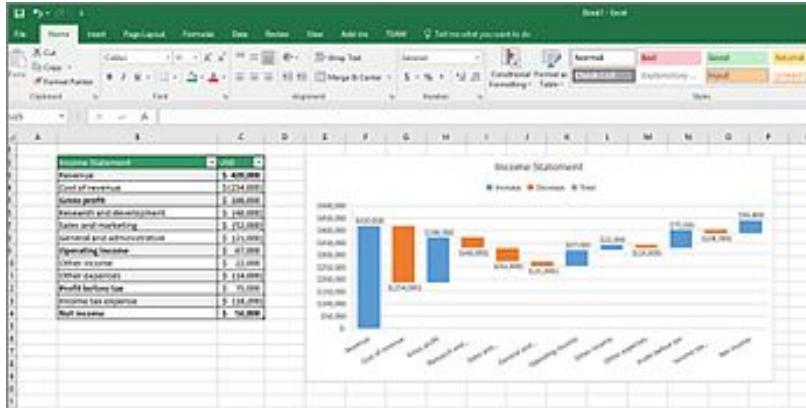
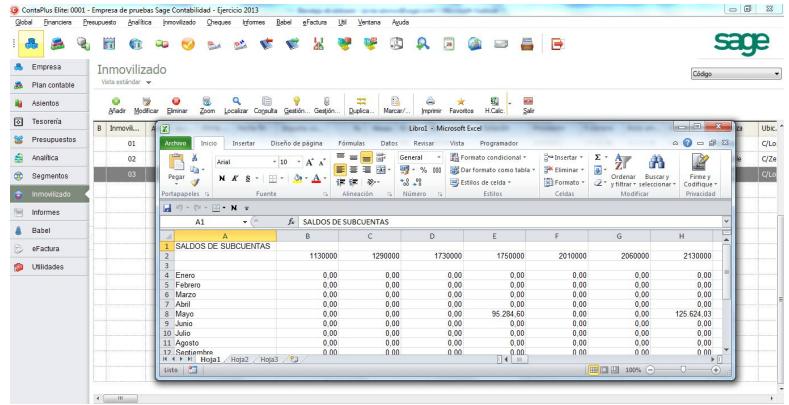
Datos, datos, datos... ¿dónde están?



■ Lo más tradicional



Nivel I



Bases de datos tradicionales



Microsoft®
SQL Server

ORACLE®

ODBC

SQLite



PostgreSQL



MySQL®



TERADATA

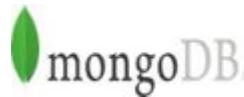


SYBASE®

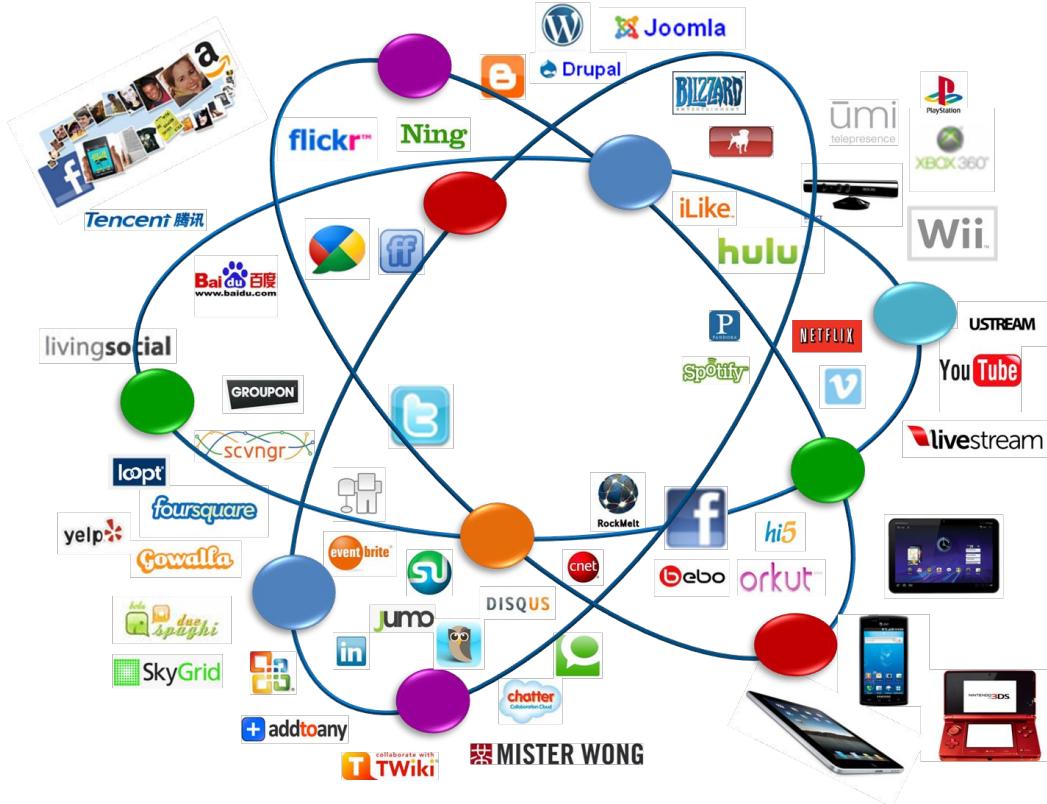
Informix®



Bases de datos NoSQL



Internet





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



Explica
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](#)

Indicador

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. Ofrece el número encuestado, 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.

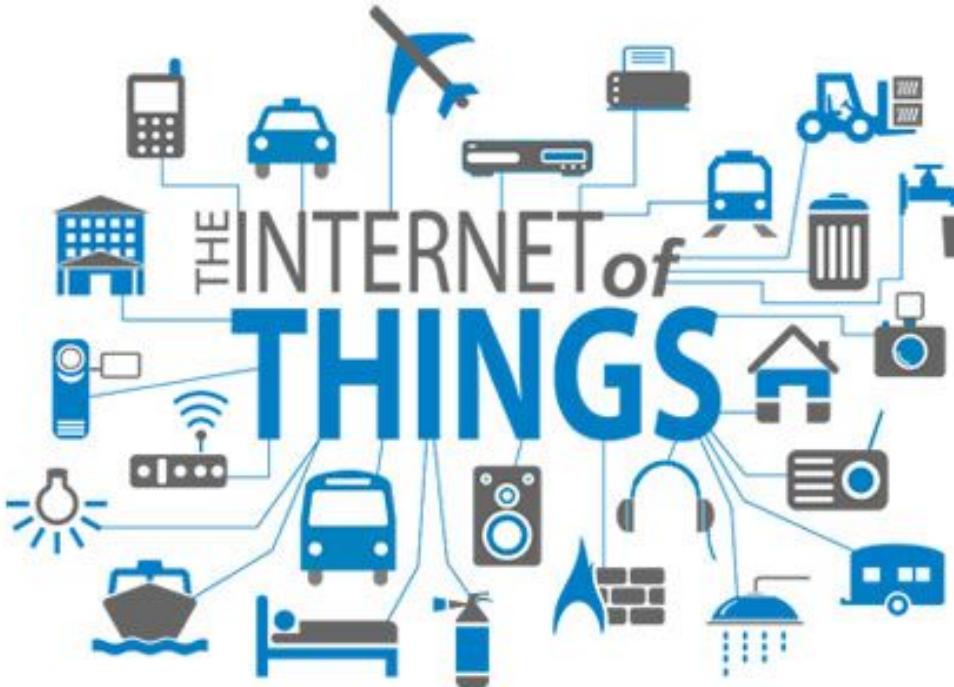


■ Otros datos interesantes



Protección de la identidad



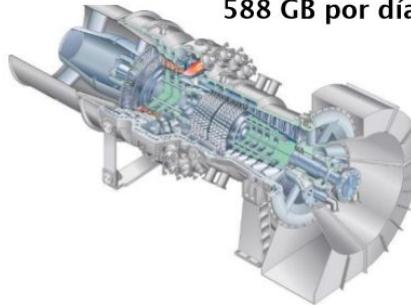


■ Los datos que no están en Internet



270 millones de
usuarios generan
100 GB por día

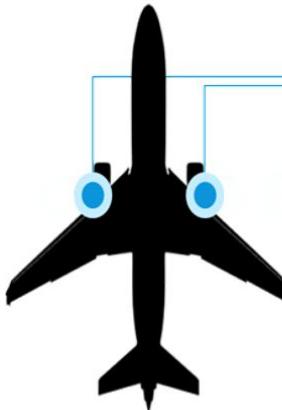
1 turbina genera
588 GB por día



■ 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

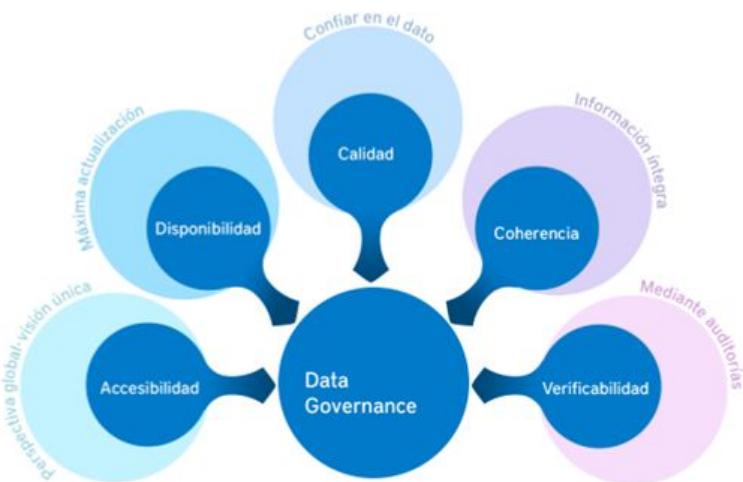


La gestión de los datos, que engloba el conjunto de todas las disciplinas relacionadas con gestionar los datos como un activo valioso, es lo que se conoce como **Data Management**.

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.



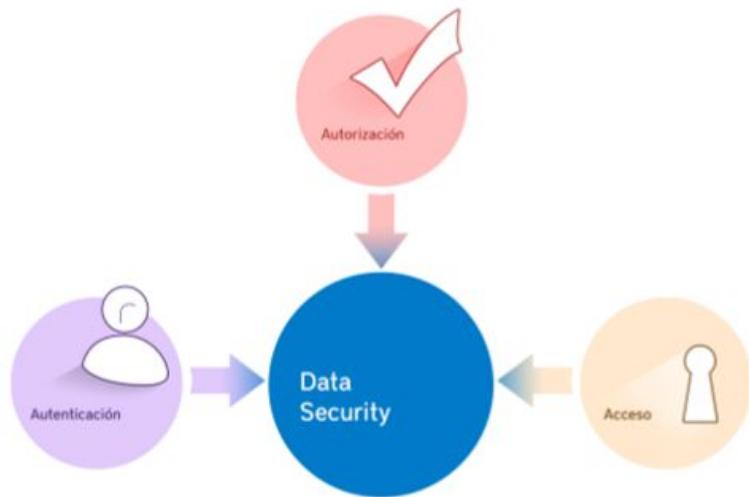
Data Storage



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



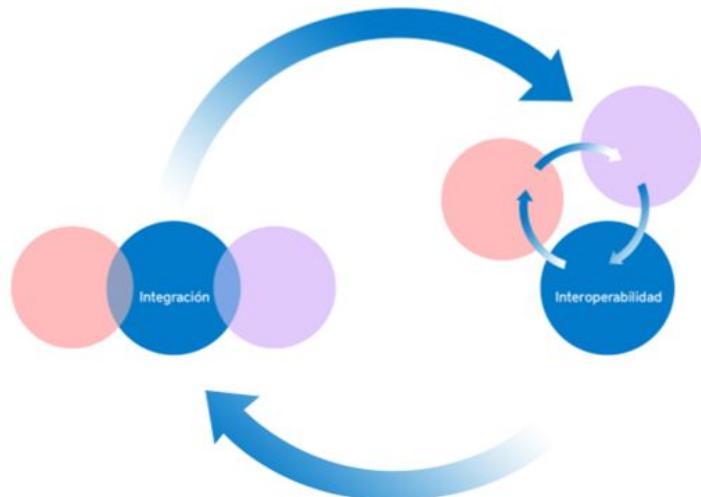
En los tiempos que corren no podemos olvidarnos tampoco del tema de la seguridad.

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

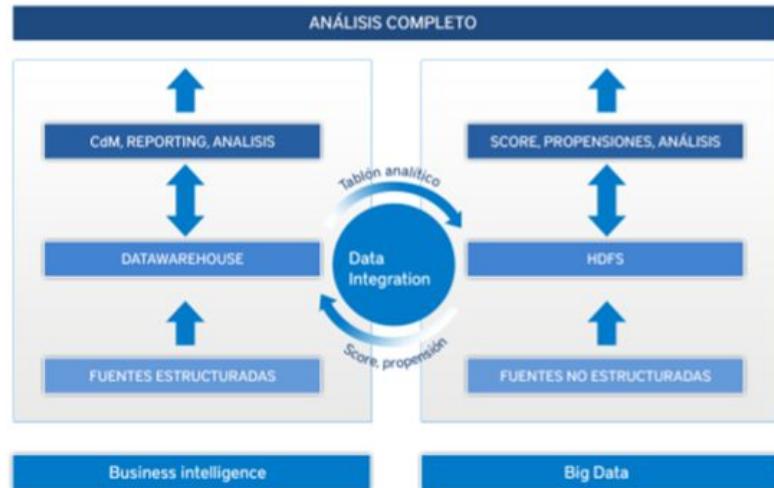


De gran importancia es la disciplina de **Data Quality**, que se encarga de definir, controlar y mejorar la calidad de los datos.

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.



■ ¡A pensar!

Keepcoding os acaba de designar para el puesto de gestor de datos, que es un puesto nuevo que no existía, ya que querías aplicar la metodología del Data Management en sus sistemas de almacenamiento de datos para su correcta gestión.

Cuando te incorporas al trabajo, encuentras la típica situación de empresa con malos hábitos en lo relacionado con los datos:

- Bases de datos relacionales sin conexión entre ellas.
- Otras bases de datos de ficheros planos no relacionales.
- Programas de administrador de bases de datos en usuarios no cualificados con permisos de administrador.
- Poca o nula seguridad en la red interna.
- Nada de calidad del dato.
- Nada de datos maestros con los que homogeneizar todas las bases de datos para una visión global.
- Poca o ninguna documentación sobre las bases de datos y sus contenidos.
- Y cualquier otra mala praxis que te puedas imaginar...

El objetivo es, con la información proporcionada, elaborar un informe para presentar en la que identifiquéis cuáles serían las 5 medidas iniciales que tomaríais para transformar esa mala gestión de los datos en una buena gestión de datos moderna que ayude a mejorar el rendimiento de la empresa, justificando las respuestas para convencerles de llevarlas a cabo.





■ BBDD Relacionales



■ Definición de base de datos

- Una BBDD es una colección electrónica de información diseñada para satisfacer unas determinadas necesidades:
 - > Pueden contener información de diversas fuentes
 - > Proporcionan mecanismos para extraer rápidamente los datos
 - > Permiten compartir información entre los distintos departamentos de una empresa.
- Las BBDD son uno de los pilares de la informática.
- Las bases de datos están por todos lados a nivel informático.



Bases de datos relaciones (I)

- En un principio existían dos modelos de bases de datos: el jerárquico y el de redes, de los cuales el de redes era el que más se adaptaba a las necesidades reales de almacenamiento y clasificación.
- Posteriormente, aparecieron las bases de datos relacionales
- Una BBDD relacional se compone de tablas, que a su vez están compuestas de campos, los cuales están formados por filas y columnas.



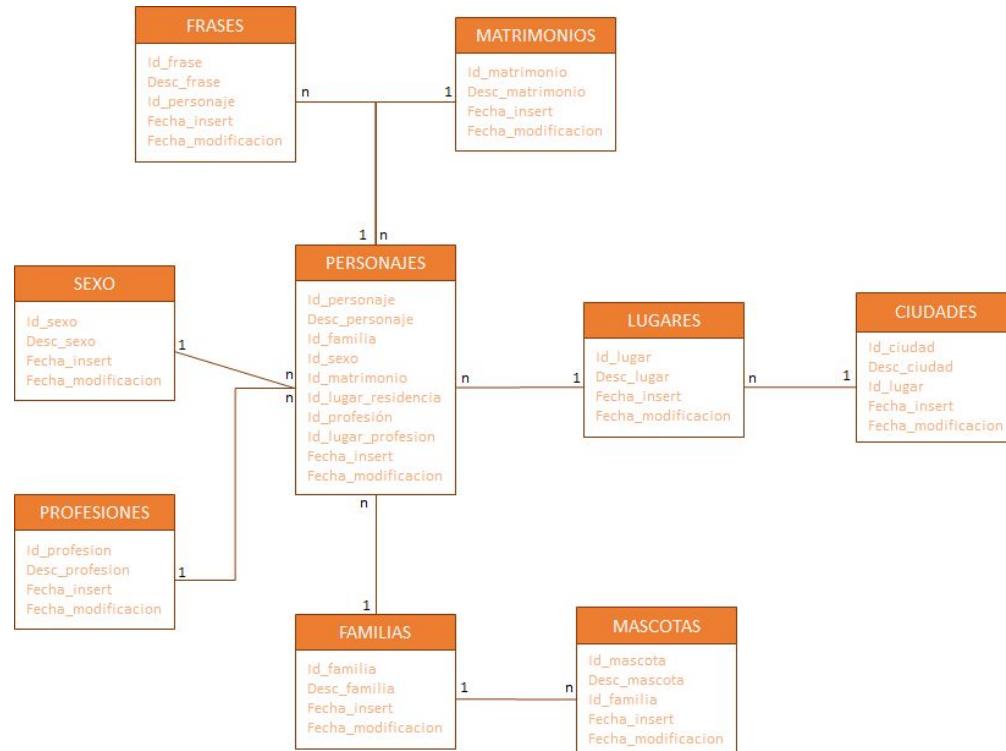
■ Bases de datos relaciones (II)

- Las bases de datos relacionales tienen la ventaja de "relacionarse" entre sí sin la necesidad de duplicar una gran cantidad de información, basadas en un lenguaje estándar llamado SQL (Structured Query Language), el cual es, podríamos decir, la razón para que las bases de datos relacionales tengan un éxito tan arrollador.



Ejemplo

- **Id** → identificador del registro en la tabla
- **fecha_insert** → fecha en la que se insertó el registro
- **fecha_modificacion** → fecha en la que se ha modificado el registro
- **Desc** → descripción de un lugar, una ciudad, el nombre de un personaje, de una mascota, etc.



■ Claves primarias y foráneas

➤ Clave primaria (Primary Key)

- > Consiste en una o más columnas cuyos datos contenidos son utilizados para identificar de manera única cada fila en la tabla.
- > No puede haber duplicados.
- > No puede tener registros vacíos.
- > Sólo puede haber una por tabla.
- > Se almacena en un índice.

➤ Clave foránea (Foreign Key)

- > Es un grupo de una o más columnas en una tabla que referencian a la clave primaria de otra tabla.
- > Puede formar parte de la clave primaria
- > Puede contener duplicados
- > Puede tener registros vacíos.
- > En la misma tabla puede haber varias.



Formas normales

➤ NF1

- > La tabla contiene una clave única
- > No tiene datos repetidos

➤ NF2

- > NF1
- > Todo registro debe depender únicamente de la clave principal
- > Las columnas pueden depender de otras tablas, pero de campos que sean clave en sus respectivas tablas

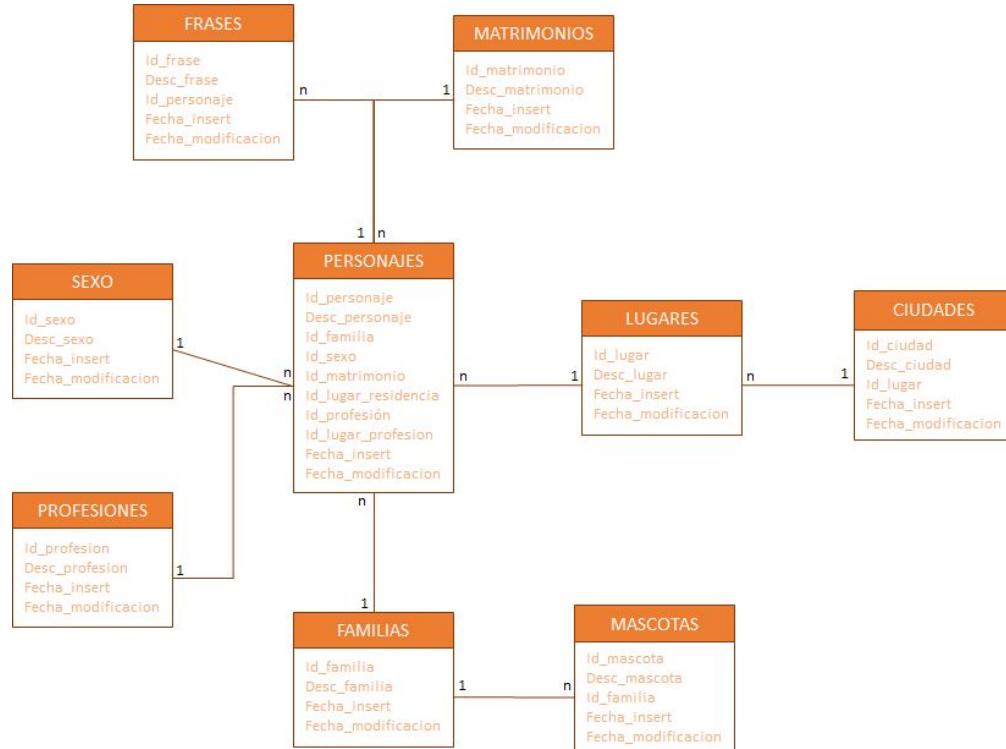
➤ NF3

- > NF2
- > No puede haber columnas que dependen de otras columnas que no sean clave principal

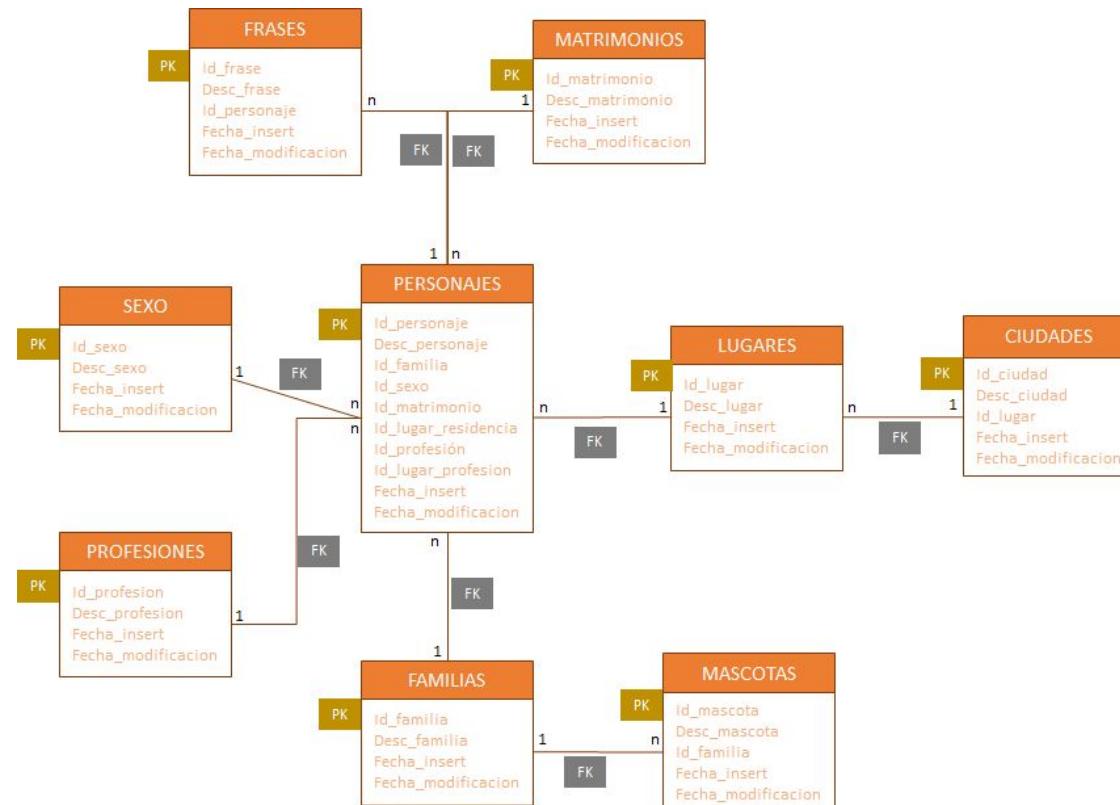


¡A pensar!

- ¿Cuáles serían las PK?
- ¿Cuáles serían las FK?
- ¿La BBDD está en forma normal?



■ No era tan difícil :)



The Hello World Collection

<https://helloworldcollection.github.io/>

The Hello World Collection

"Hello World" is the first program one usually writes when learning a new programming language. Having first been mentioned in Brian Kernighan's [tutorial to the B programming language](#), it became widely known through Kernighan & Ritchie's 1978 book that introduced ["The C Programming Language"](#), where it read like this:

```
main() {
    printf("hello, world\n");
}
```



Since then, Hello World has been implemented in just about every programming language on the planet. This collection includes **567 Hello World programs** in as many more-or-less well known programming languages, plus **76 human languages**.

The programs in this collection are intended to be as minimal as possible in the respective language. They are meant to demonstrate how to output Hello World as simply as possible, not to show off language features. For a collection of programs that tell more about what programming in the languages actually is like, have a look at the [99 Bottles of Beer](#) collection.

The Hello World Collection, started in 1994, was compiled with help from [many people around the world](#). It is the biggest collection of Hello World programs on the Internet, and the only one collecting human languages as well. To contribute, send your program to info@helloworldcollection.de. Begin your contribution with a comment in the respective language. [Real programming languages only](#) please.

Click [here](#) for a list of all contributors and other sources.

Click [here](#) for related links.

Click [here](#) for brief history of the Hello World Collection.

Support the [KDE Education Project](#) with our exclusive [Hello World merchandise](#) — T-shirts, mugs and more!

Last update: Oct 21, 2017.





■ 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.



■ DataWarehousing (I)

Un **DataWarehouse** es una colección de datos:

- Orientados
- Integrados
- No volátiles
- Variables en el tiempo

que ayudan a la toma de decisiones en la entidad en la que se utiliza

Bill Inmon

Un **DataWarehouse** es una copia de las transacciones de datos específicamente estructurada para la consulta y el análisis

Ralph Kimball



■ DataWarehousing (II)

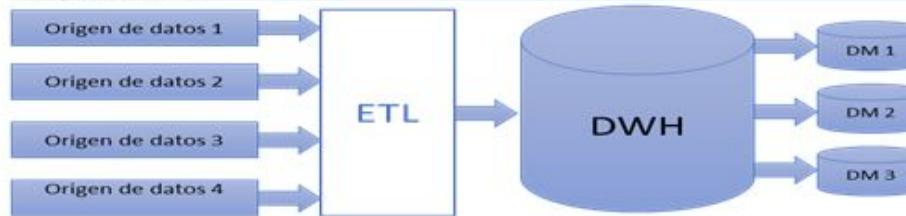
- Un Datawarehouse no se debe comprar, se debe construir
- Implementación por fases
- Implica a varios departamentos de la empresa
- El diseño no es único ni estándar



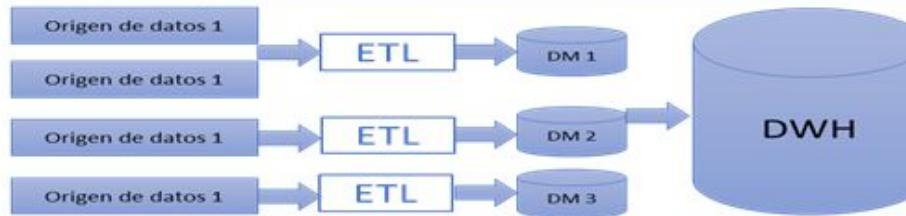
Data Marts

Un **DataMart** es un subconjunto del DataWarehouse que proporciona una visualización personalizada.

Top-Down



Bottom-Up

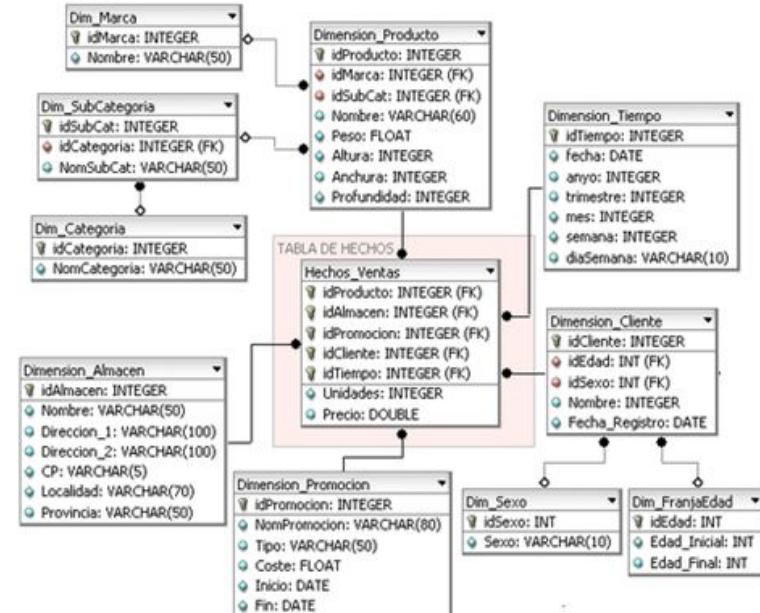


Modelización de la estructura

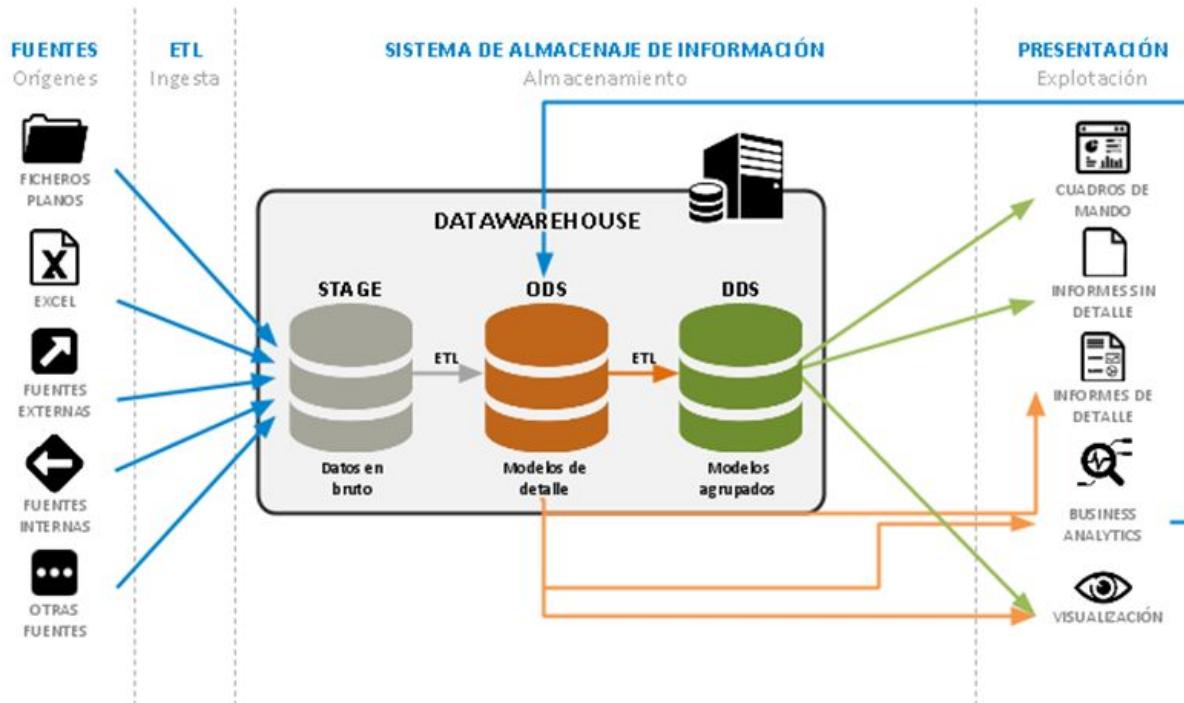
Modelo en estrella



Modelo de copo de nieve



■ Arquitectura de un DataWarehouse



■ Un DataWarehouse no es una BBDD relacional!!!

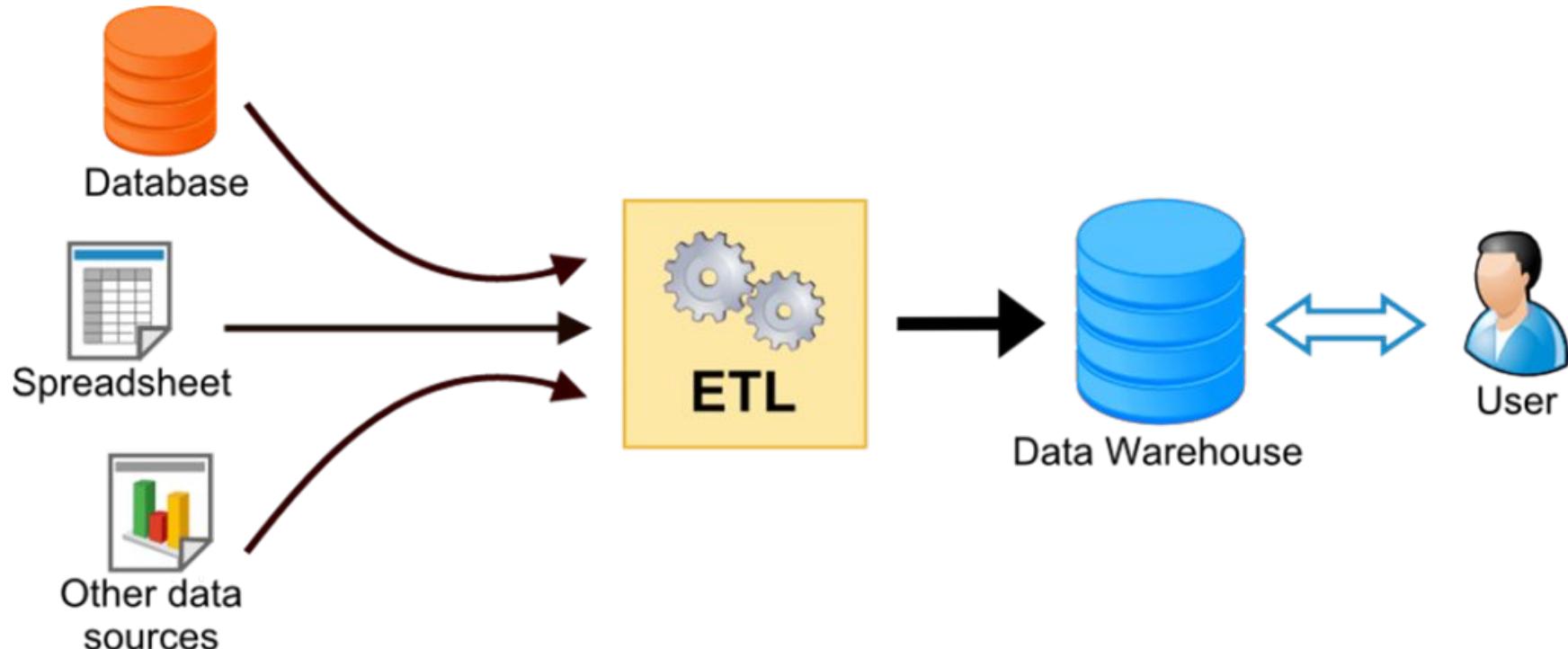
- No podemos tener valores nulos

Tipo de campo	DESCONOCIDO	NO APLICA
Numérico	9999999999999999	9999999999999998
Cadena	DESCONOCIDO	NO APLICA
Fecha	31/12/9999	31/12/9998

- Puede estar desnormalizada



Extract Transform Load



KEEP
CODING

■ Ten siempre en cuenta que:

- Los procesos de BI son vivos
- La actualización de datos debe automatizarse
- BI debe agilizar los procesos para la toma de decisiones posterior





■ 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

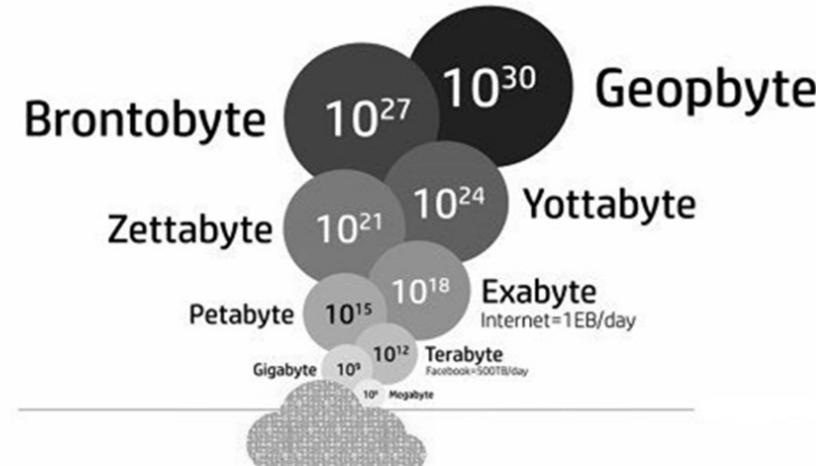
- **Velocidad:** aunque tenga que procesar millones de mensajes en Twitter, por ejemplo, quiero saber “ya” qué están diciendo los usuarios sobre mi empresa, sobre un producto que acabo de lanzar o sobre un acontecimiento que está ocurriendo ahora mismo. O quiero saber en tiempo real si mi cadena de montaje 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

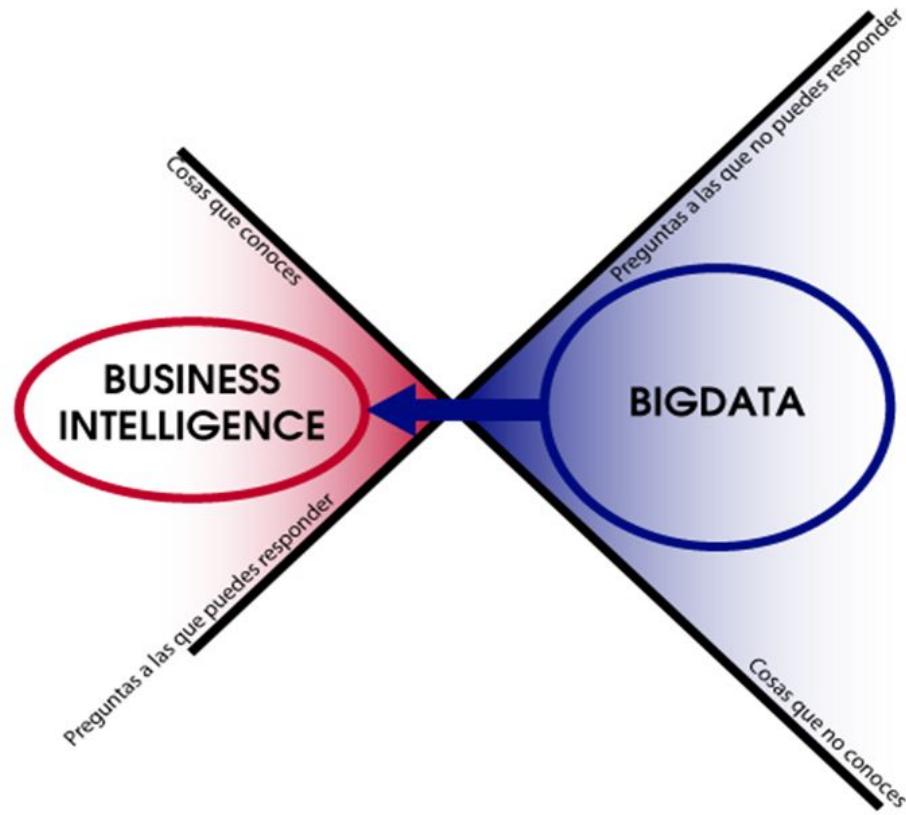


■ BI vs Big Data

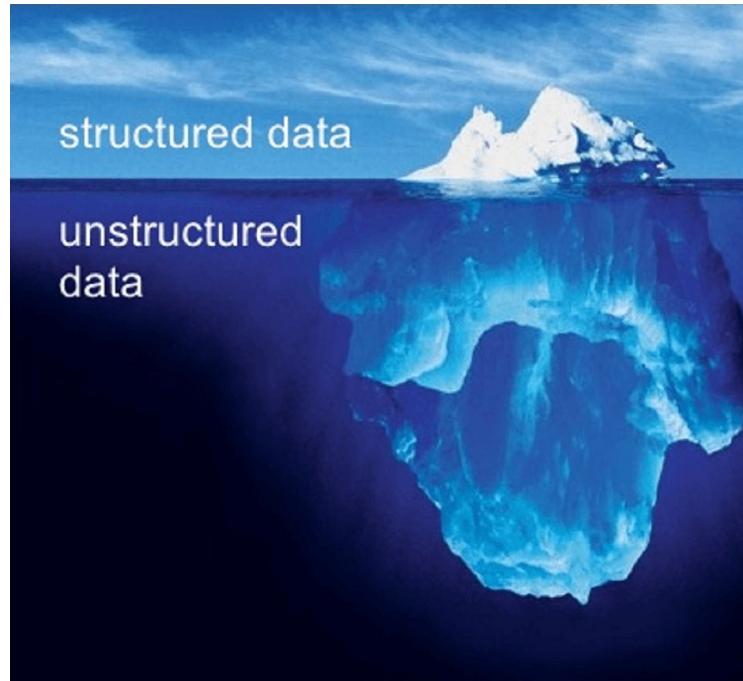
- Islas de información que no permiten vincular la información.
- La recopilación manual que muchas veces se hace.
- Las inconsistencias de los datos que provocan que, por ejemplo, la misma métrica, según dónde se mire, tenga un valor u otro.
- Los informes poco intuitivos.
- El retraso a la hora de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

2012

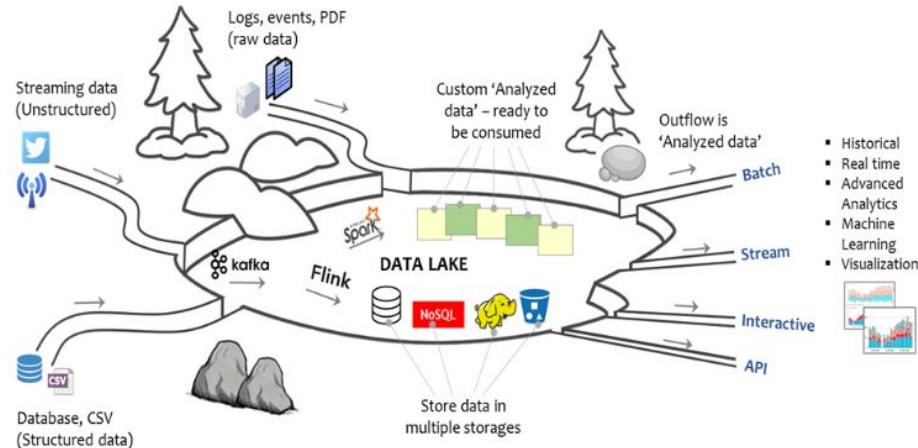
- Fuentes no estructuradas de datos



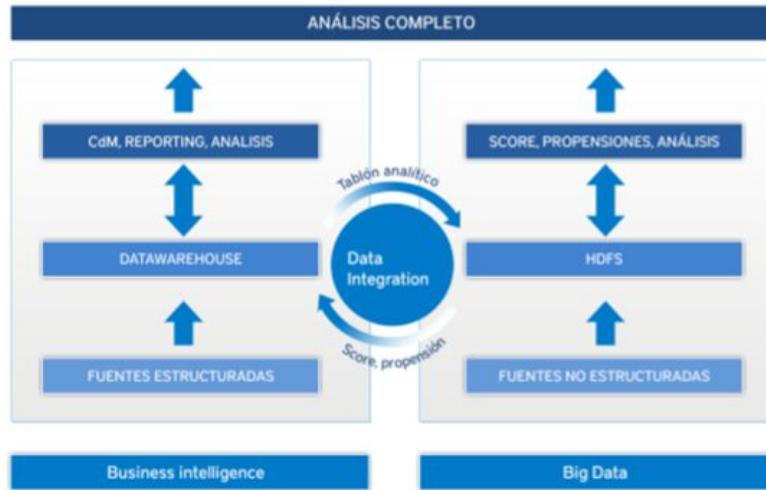
■ BI vs Big Data



Data Lake



■ BI & Big Data coexistiendo





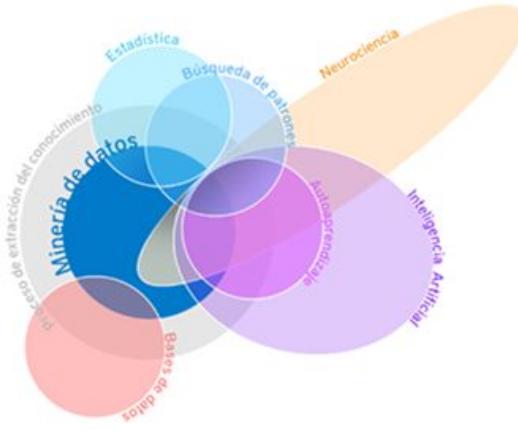
■ Enfoque multidisciplinar



■ La tercera variable



■ Disciplinas científicas



Minería de datos (Data Mining)
Estadística (Statistics)
Búsqueda de patrones (Pattern Recognition)
Aprendizaje automático (Machine Learning)
Neurociencia (Neurocomputing)
Inteligencia artificial (AI)
Bases de datos (Databases)
Proceso de extracción de conocimiento (KDD)

(c) Refinaria SAE Institute

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



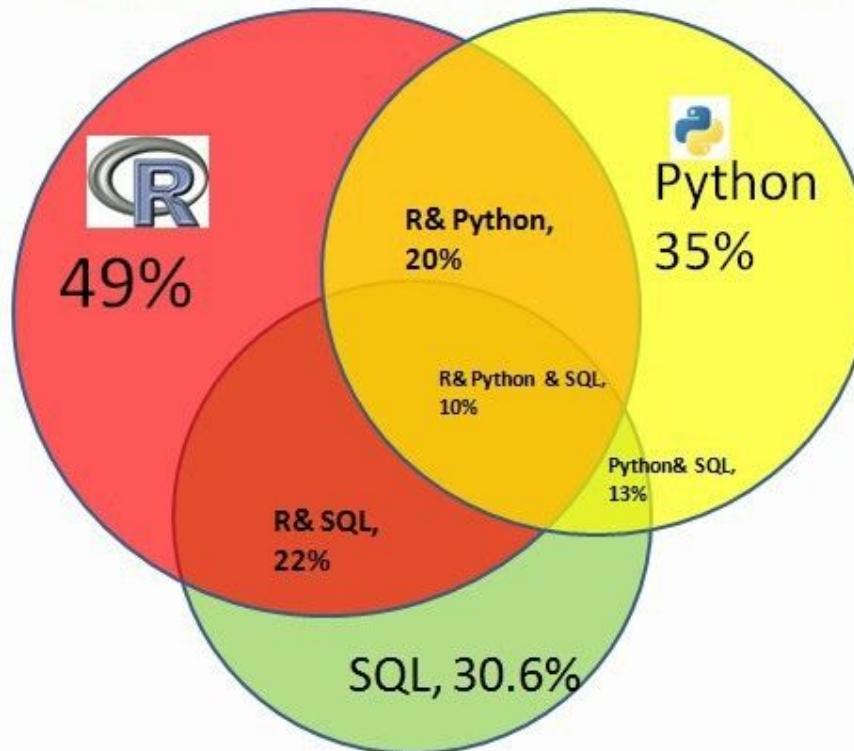
■ El científico de datos

- 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 el fin 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.

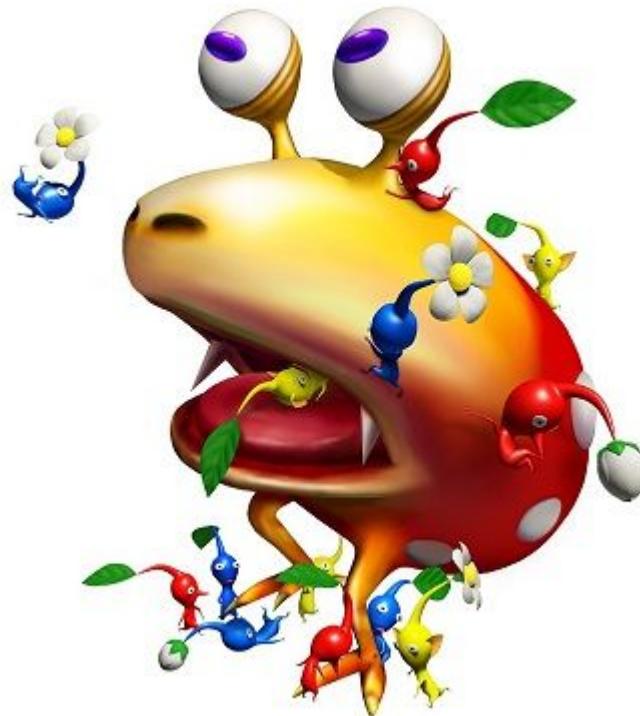
Este directivo que está empezando a incluirse en las organizaciones, pero de un modo bastante lento hasta el momento, la firma de análisis Gartner estima que el **90% de las grandes empresas tendrán un CDO en su organigrama en 2019**.



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





■ Casos de uso



Customer Analytics

- A día de hoy los **clientes** exponen y comparten sus **deseos** a través de **múltiples canales**, y son las empresas las que tienen que escucharles para adaptarse a sus exigencias.
- Sin embargo, la mayoría de las **empresas** están todavía en un **paso anterior**: el cambio de orientación del producto al cliente como centro de su negocio.
- El escenario en el que deberían estar trabajando las empresas a día de hoy es en reforzar su actividad mediante el **diálogo** y la **predicción** sobre lo que van a **desear** los **consumidores** más que sobre la venta, aunque éste sea el fin al que se quiere llegar.
- Para dar este salto estamos obligados a repasar **cuánto sabemos** de nuestros **clientes**. Muchas empresas dirían que los conocen muy bien, pero realizando sencillas preguntas veríamos que no es así.
A día de hoy, en un proceso normal, el cliente realiza una compra y una vez realizada empezamos a conversar con él hasta tener una relación con él que le fidelice. Por tanto, empezamos a tener **información** real del **cliente** en la **tercera etapa** de nuestra interacción con él.
Lo que nos estamos perdiendo es que **nuestros clientes dejan huellas** sobre sus deseos, necesidades, marcas, opiniones, horarios, etc.



Customer Analytics

- A día de hoy lo más importante no es la satisfacción de los clientes, sino la emoción de los mismos, ya que en un mercado tan competitivo donde el consumidor ha tomado las riendas la satisfacción deja de tener sentido.
A esto se lo conoce como la **economía de la intención**. El padre de “La economía de la intención” es Doc Searsi. Su teoría, escrita en un libro que ha recorrido el mundo, se basa en que el cliente va a comprar y la empresa, en vez de venderle, tiene que satisfacer sus necesidades y adivinar sus intenciones para poder ofrecerle productos y servicios personalizados.
- En resumen, quien a día de hoy tiene **fans** en vez de compradores está usando una estrategia basada en Customer Analytics.



■ Customer Analytics (I)



Amazon.com

Amazon comenzó como una plataforma de venta de libros estándar; es decir, tenía personas que aconsejaban a los compradores sobre los libros que debían adquirir en función de sus gustos o necesidades.

Aunque funcionaba bien, la siguiente idea fue personalizar mejor las recomendaciones, añadiendo entradas de datos como las compras que el cliente había realizado anteriormente, o consultas sobre otros libros que el usuario estaba haciendo en la web.

Este conjunto de datos es masivo, y lo va siendo más conforme se tiene éxito en las ventas y se van introduciendo nuevos tipos de productos... En base a la recogida y análisis de estos datos, todos sabemos hoy en día qué es la plataforma de venta por internet de amazon, la diversidad de productos que ofrece y el volumen de ventas que genera.



■ Customer Analytics (II)



Macy's y precios en tiempo real

Macy's es uno de los comercios minoristas más importantes de los Estados Unidos, que destaca por su e-commerce. Utilizando la tecnología de SAS Institute ha conseguido mejorar sus ingresos y la experiencia del usuario.

Gracias a la velocidad de análisis y los informes obtenidos con esta nueva tecnología, han reducido en 500.000 dólares el gasto anual de analítica. Macy's sabe hoy perfectamente el impacto de sus newsletters y notificaciones y conoce mejor a los clientes más satisfechos, lo que les gusta y lo que no...

Hoy, el uso de estos datos, les permite segmentar al máximo sus envíos, de manera que envían menos emails, pero con mucho más impacto y han conseguido reducir las desuscripciones hasta en un 20%.

Gracias a la utilización de un algoritmo y al control de la demanda y el inventario, pueden lanzar ofertas cruzadas, ajustar precios y hacer rebajas casi en tiempo real para sus 73 millones de artículos a la venta.



Customer Analytics (IV)



Aumentar la fidelidad del cliente, Avis Budget

Los clientes de hoy en día son más exigentes y volubles que nunca. Mantener o aumentar la cuota de mercado exige a las empresas entender a sus usuarios tanto como sea posible, mejorar continuamente sus productos y servicios y estar dispuestos a adaptar sus modelos de negocio para reflejar las necesidades reales de sus clientes.

Avis Budget, empresa dedicada al alquiler de vehículos implementó una estrategia destinada a aumentar la cuota de mercado que ha resultado en cientos de millones de dólares de ingresos adicionales.

La iniciativa pretendía lograr valor a través de la implicación de los clientes, mediante su segmentación y la oferta de incentivos por niveles para mejorar su fidelidad.

Para ello, su socio tecnológico CSC aplicó un modelo sobre Big Data que calculaba el valor de la base de datos de clientes de Avis y, a continuación, lo validaba mediante una campaña de marketing multicanal con análisis simultáneo.

Los datos de valoración del cliente ahora se combinan con otros, incluyendo el historial de alquiler, los problemas del servicio, la demografía, la afiliación corporativa y la retroalimentación de los clientes.

Además, Avis también está recopilando y analizando los datos de redes sociales. Está utilizando el Big Data para pronosticar la demanda regional en cuanto a necesidades en sus flotas de vehículos y la política de precios.



Customer Analytics (V)



Reducción de portabilidades de T-Mobile

T-mobile consiguió reducir a la mitad el número de portabilidades gracias a la aplicación de técnicas sobre Big Data.

Las operadoras de telefonía móvil e Internet tienen un número impresionante de datos sobre sus clientes: la cantidad de llamadas que realizan, las horas en las que tienen lugar, sus números favoritos, el número de llamadas que se cortan por problemas de cobertura y un larguísimo etcétera.

Con todos estos datos en la mano y analizando las interacciones de sus clientes en redes sociales, en T-Mobile se propusieron rebajar sustancialmente el número de portabilidades hacia otros competidores en Estados Unidos.

Para ello la empresa utilizó tres herramientas básicas: sus propios sistemas de cobro (billing systems), herramientas de monitorización social, además de Splunk y Tableau Software para analizar la información y presentarla de una forma visual.

Combinando toda esta información en T-Mobile descubrieron que las expectativas de portabilidades pueden determinarse a través del análisis de tres factores:

- Facturas
- Llamadas que se cortan debido a mala cobertura
- Conversaciones de los clientes: positivas, negativas o neutrales

Todos estos factores fueron asociados a la influencia o reputación en medios sociales de cada uno de sus clientes, partiendo de la hipótesis de que clientes con un gran número de seguidores o influencia podrán tener un efecto positivo o negativo (según las circunstancias) en otros potenciales clientes de la marca.

La combinación de todos los aspectos mencionados anteriormente llevó a T-Mobile a calcular para cada cliente un 'Customer Lifetime Value', un valor monetario individual según las expectativas de negocio y permanencia. Esta información era transmitida en tiempo real a cada agente de la compañía para presentar a los clientes ofertas personalizadas en función de su valor personal.

De esta forma la empresa pasó de casi 100.000 portabilidades en un trimestre, pasaron a tan sólo 50.000 en el siguiente trimestre, una reducción del 50% gracias a un buen aprovechamiento del Big Data y de todos los datos e información que la operadora tiene de sus clientes.



Customer Analytics (VI)



Fidelidad de clientes en banca

Un gran banco quería incrementar los resultados de sus campañas de fidelización de clientes, y mejorar el éxito de sus campañas de venta, profundizando en el conocimiento de sus clientes. Esto es un proyecto de Big Data, que se basó en datos mensuales del orden de 36 Terabytes.

El banco había detectado que el comportamiento de sus clientes había estado cambiando en los últimos años, y que tendían a cambiar de banco cuando hacían grandes compras, como una vivienda.

Se generaron nuevos insights para dar soporte a esta situación. Se detectaron eventos importantes en la vida de los clientes: nacimiento de un hijo, casarse, comprarse un coche, o comprarse una vivienda. Se trazaron métricas de vinculación entre estos hechos y transacciones de productos con el banco; es decir, se observó al contrario de lo habitual, que era ofrecer un producto y ver si el cliente respondía. Ahora se observa qué le ocurre al cliente, y cómo actúa con su banco en consecuencia.

En base a este análisis se hizo una nueva segmentación de los clientes, y un modelo predictivo que pueda ver venir qué productos van a querer los clientes en función de lo que vayan a hacer en sus vidas. Y también la amenaza que supone esto hacia la posibilidad de que cambien de banco.

Se identificaron en total 200 momentos relevantes distintos en la vida de los clientes. Diariamente se ejecutan los algoritmos que tratan de prever estos momentos, y esto es lo que ha dirigido la estrategia de las campañas de producto del banco durante el 2016.



User experience

- Hoy en día las **personas** tienen más **acceso** a la información y se han vuelto más **críticos**, sobre todo se informan a través de internet en las **redes sociales**, **blogs**, etc.
- Además, con el desborde de información existente, se han creado efectos como la **ceguera al banner** o la **alergia a las campañas del siglo XX**: 2x1, ¡Llame ya!, ¡Oferta!, ¡Promoción!,...
Esto afecta negativamente sobre todo a las empresas más tradicionales. Otras se han adaptado, pero aún no han alcanzado el retorno de la inversión hecha, pero también hay otras que esta evolución ya les está dando sus frutos, poniendo al cliente como centro de todo.
- Al cliente a día de hoy hay que provocarle emociones y una de las maneras de hacerlo es con la **experiencia de usuario (User Experience)**.
- Si una experiencia es muy positiva será recordada y ritualizada para la posteridad.



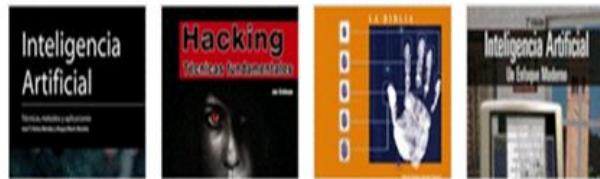
User experience

- La experiencia involucra múltiples factores, algunos incontrolables. Lo que se busca con la metodología de la **experiencia de usuario** son aquellos que se pueden controlar con tal de crear una experiencia positiva en relación al consumo de algo. De hecho, va más allá, porque intenta descubrir más allá de lo que las personas quieren, llegar a conocer lo que las personas aman.
- Si generamos una **experiencia de usuario positiva** no sólo nos servirá positivamente con nuestro cliente, sino que como vivimos en una era interconectada, muchas más personas sabrán la magnífica experiencia que se obtiene con ese consumo y seguramente se volverá viral.
- La experiencia de usuario **aplica tanto a productos como a servicios**.
- Además, promover internamente en la empresa la cultura de que el **cliente** es el **centro** de todo y nuestro objetivo de generarle emociones positivas.



User Experience (I)

LOS CLIENTES QUE VIERON ESTE LIBRO TAMBIEN VIERON



Canal Online, Casa del libro

En el caso del canal online aplicado al sector Editorial, La Casa del Libro es un buen ejemplo de cómo aprovechar la información derivada del Análisis de la Cesta de la Compra en tiempo real.

Si hacemos la prueba, y seleccionamos cualquier libro de toda su oferta, debajo de la descripción y los detalles del mismo nos encontramos un mensaje del tipo "los usuarios que consultaron este libro también vieron..."

Esto se llama motor de recomendación, y su funcionamiento está basado en el análisis en tiempo real de las visitas de los usuarios a la plataforma online y su patrón de navegación por la web.



User Experience (II)



Canal físico, Walmart

Realizaron un estudio para saber qué productos se vendían con mayor frecuencia junto a los pañales de bebé. Y sorprendentemente encontraron que era la cerveza.

Profundizando en el estudio, identificaron que el patrón de compra más frecuente era el viernes por la tarde y el consumidor habitual hombres de entre 25 y 35 años.

La explicación que dieron a este descubrimiento es que como el paquete de pañales es un producto voluminoso, y

suele ser responsabilidad del hombre encargarse de esta compra, que habitualmente realiza el viernes, y de paso aprovecha para comprar cervezas para el fin de semana.

El supermercado actuó en consecuencia y colocó la cerveza al lado de los pañales, potenciando su consumo. El resultado fue que los consumidores que ya cumplían el patrón intensificaron la compra de cerveza, mientras que los padres que antes no la compraban empezaron a hacerlo por la proximidad de ésta.

Conclusiones: se transforman datos en resultados.

Como vemos, el Análisis de la Cesta de la Compra proporciona tres resultados muy específicos:

Beneficios tangibles de los datos, rentabilizando la información interna

Optimización de los recursos, mejorando ratios de productividad

Aumento de la facturación anual y del margen de beneficios

De forma global, la aplicación de la Minería de Datos a cualquier área de una compañía se transforma en resultados tangibles en la medida en la que es capaz de resolver un problema de negocio.



User Experience (III)



Pokemon Go

Salvo que hayas permanecido encerrado en una sala completamente aislada del exterior durante los últimos meses, probablemente hayas oido hablar de Pokémon Go. Este videojuego ha dado mucho que hablar: en 30 días había logrado 200 millones de dólares en ingresos, con más de 10 millones de descargas en la primera semana. Pero, ¿cómo?

En los últimos años, el mercado de los videojuegos móviles ha ido relegando a un segundo plano los PCs y las consolas de videojuegos. También el modelo de negocio es diferente en ambos casos: mientras que en PC y consolas el propio videojuego tiene un coste que puede ser elevado (en torno a 60€); los juegos con más descargas en plataformas móviles suelen adecuarse al modelo de negocio denominado "freemium", en el que el juego en sí es gratuito, pero el jugador puede realizar compras dentro del juego de contenido adicional: objetos, vidas, dinero virtual, etc.

Si tenemos en cuenta este modelo de negocio, en el que el jugador no desembolsa ningún dinero antes de comenzar a jugar, es evidente que a las compañías les interesa retener a los jugadores e incitarles a que realicen alguna compra.

Por este motivo, es importante para una compañía saber qué están haciendo sus usuarios, pero es aún más interesante saber qué van a hacer en el corto plazo.

El primer paso para poder analizar el comportamiento de los usuarios y poder realizar un estudio predictivo de sus próximas acciones consiste en monitorizar todas las acciones que estos realizan sobre el juego. Esto implica guardar todos y cada uno de los eventos que realiza el jugador: iniciar sesión, vencer a un enemigo, invitar a un conocido, pasarse un nivel, perder una vida y, por supuesto, realizar una compra.

Mediante este proceso, obtendremos grandes cantidades de datos proveniente de todos nuestros usuarios. Dado el carácter temporal de la información, el comportamiento de cada usuario puede establecerse como una secuencia de acciones.

De este modo, la pregunta que nos debemos plantear es la siguiente: si conozco las últimas acciones que ha llevado a cabo un determinado jugador, ¿puedo saber cuáles son las posibles siguientes acciones y con qué probabilidad las llevará a cabo?

En términos generales, el problema anterior puede considerarse un problema de clasificación, por lo que se sitúa dentro de los problemas denominados "de aprendizaje supervisado" dentro de la rama del aprendizaje automático.

Se trata de ir almacenando secuencias de acciones que hace un usuario, y entrenar al sistema para que aprenda estas secuencias.

Ahora bien, si tenemos datos de un nuevo usuario, ¿qué debemos hacer para predecir su comportamiento futuro? En este caso, basta con tomar las "n" últimas acciones que ha realizado y buscar en el modelo este antecedente.



User Experience (IV)



Ulabox - e-commerce y Big Data

Es una empresa que quiere dar respuesta a la insatisfacción con la experiencia de compra en los supermercados tradicionales y sus variantes online. Ulabox te permite hacer tu compra online a partir de un catálogo que dispone de más de 13.000 productos.

Para el éxito de este negocio se han centrado en la experiencia del cliente para evitar así el abandono del carrito y han llegado a alcanzar un ratio de conversión de hasta el 17%.

La logística de un supermercado online es muy compleja ya que implica a una enorme cantidad de líneas y se precisan diferentes tratamientos a nivel de temperaturas. El caso de Ulabox es especialmente complejo, pues realiza más de 10.000 envíos al mes. Su estrategia logística para dar paso a tanto crecimiento se basa en dividir su modelo en dos:

Barcelona y su área metropolitana: Representa un 60% de su mercado.

El resto de España: Donde solo entregan las 11.000 referencias de seco.

Todo esto exige un procesamiento de rutas en cada parte del modelo, de manejo de las miles de líneas de productos, y de gestión de una cantidad masiva de envíos. Aquí Big Data significa procesamiento de gran cantidad de información, en función de las interacciones de los usuarios en tiempo real.



User Experience (V)



Promofarma

Su objetivo es unir la poco desarrollada oferta en Internet de parafarmacia en España y la venta online. Su modelo de negocio combina la venta online con el valor y la confianza que genera en los consumidores el hecho de que los productos sean preparados por las mismas farmacias.

El funcionamiento es el siguiente: El cliente tiene acceso a un extenso catálogo de artículos propuestos por las diferentes farmacias que lo componen. Realizan su pedido y Promofarma utiliza un algoritmo matemático para identificar la farmacia que ofrece un mejor precio para la cesta de productos demandados por el cliente, incluyendo los costes de transporte.

Aquí es donde hablamos de Big Data: procesamiento de una entrada ingente de datos, su comparación, la ejecución de un algoritmo complejo, ordenación masiva de datos, y su presentación de forma amigable al usuario.

Para este modelo de negocio, la logística representa el 50% del éxito del eCommerce.

También han sido necesarios importantes desarrollos tecnológicos y una coordinación muy precisa por parte del partner de transporte para recoger y entregar los pedidos de clientes, preparados por los clientes de farmacias con las que colaboran, a tiempo, en el lugar correcto, haciendo la ruta óptima... otros de los casos de uso que contamos es justamente sobre una de estas empresas de transporte.

Por eso se dice que un ecosistema Big Data acaba posibilitando el movimiento entero de un mercado o de un sector, no sólo de una compañía aislada hacia sus clientes; también con sus proveedores, partners, y por supuesto para conocer e influir en su competencia.

De cara al futuro, Promofarma pretende optimizar los algoritmos de última milla para posibilitar entregas en bicicleta o vehículos eléctricos.

Por otro lado, otros algoritmos y procesamientos avanzados podrían reducir los plazos de entrega para el día siguiente o incluso el mismo día, o incluyendo plazos más reducidos de una y dos horas para las urgencias.



User Experience (V)



Netflix

Netflix es la plataforma más grande del mundo de series y películas online. La compañía supo detectar el gran poder del Big Data para conocer los gustos de sus usuarios y su éxito se debe, entre otros factores, a las recomendaciones que hace a sus consumidores en base a los patrones de consumo de contenido.

¿La evidencia de su éxito? la serie House of Cards. Una producción propia que Netflix creó en base a los gustos de los usuarios.

A partir de los patrones de consumo obtenidos de más de 40 millones de consumidores, detectó que lo que más atraía eran contenidos que incluyeran drama, política, sensualidad y poder y, en base a ello, crearon la serie.



Business Analytics

¿Qué puedo mejorar dentro de mi empresa?

- Cualquier empresa necesita clientes para existir, pero no sólo depende del comportamiento de los clientes que la empresa vaya bien.
- Los **procesos internos** de la empresa también son importantes para asegurarse la optimización de los recursos disponibles, la capacidad de reacción ante una adversidad y la corrección de errores en los procedimientos internos.
Sin esto, por muchos clientes que la empresa tenga, está predestinada al fracaso en algún momento, ya sea porque aumente el número de clientes y por cómo está definida no tenga capacidad de darles servicio a todos, o bien porque le falla alguno de sus recursos y no lo tenía previsto con el consiguiente retraso en dar servicio, o porque sus procedimientos no están optimizados y cuando tenía pocos clientes no se notaba y al aumentar el número de clientes llegan los retrasos en dar sus servicios,...
- Por otro lado, nos encontramos que también es muy importante para el buen funcionamiento de la empresa tomar **decisiones estratégicas** adecuadas y, para que así ocurra, es realmente útil apoyarse en los resultados que se obtienen con los métodos analíticos existentes.



Business Analytics

- El objetivo de cualquier empresa debería ser convertirse en una **empresa inteligente**. Hay que entender que una empresa es un ente más complejo que un sistema o una única persona. El conseguir una empresa inteligente implica la gestión eficiente de la información: almacenamiento, reporting y explotación analítica para la toma de decisiones. Adicionalmente, disponer de las habilidades de la inteligencia humana: aprender, resolver problemas, tomar decisiones óptimas, etc.
- Nuestro objetivo es subir en el gráfico del **DIKW** cuanto más alto mejor.
Es importante primero tener **informes** de seguimiento **fiables** con métricas que midan la **bondad** de los procesos internos y otras métricas que midan si la estrategia que se está siguiendo es adecuada y está dando sus frutos.
Estos informes nos sirven también para ver si ha ocurrido algo y en qué circunstancias ha pasado. Con esta información, buscar cuál es el problema y determinar qué acciones tomar.
Después, ver por qué ha pasado y ver si es algo puntual o que tiene tendencia a crecer, mantenerse o disminuir hasta desaparecer.
Con todo esto, predecir qué será lo siguiente que pase y, usando la analítica más avanzada, obtener qué acciones se deberían realizar.
- Conseguir el **DIKW** completo es un proceso complejo, no alcanzable ni viable para todas las empresas. Cada empresa tiene que ver en qué punto se cubren sus necesidades, aunque si bien es cierto, en cualquier empresa se puede conseguir todo, aunque es cuestión de tiempo en implementarlo y dinero para financiarlo.



Business Analytics (I)



Precios de inmuebles

Una de las consultoras más grandes de España ha desarrollado un sistema basado en algoritmos analíticos para calcular automáticamente el precio de venta más adecuado para viviendas.

El objetivo es ahorrar tiempo a los equipos comerciales, y obtener precios de venta más precisos. Es más barato trabajar con un sistema informático delante de un ordenador, que ir a visitar un inmueble, con su desplazamiento incluido, hacer el peritaje o valoración, etc...

Al mismo tiempo, se obtienen variables como la sensibilidad al precio para el éxito de la venta, o el tiempo esperado de venta de cada propiedad. Se puede saber qué inmuebles se venderán demasiado pronto, con lo cual se recibirán pocas ofertas, cuáles se venderán en su tiempo adecuado, y cuáles no se venderían nunca. Y poder conocer cómo cambia este período de venta en función del cambio que se haga en el precio propuesto.

Para esto, se toma información del inmueble como su localización, metros cuadrados, características, antigüedad y otras muchas.

Y se toma información externa al propio inmueble como el precio de venta de propiedades las propiedades colindantes o cercanas, y cómo van cambiando estos precios a lo largo del tiempo. Esta información se obtiene a través de webs públicas de compra venta de pisos entre particulares.



Business Analytics (II)



DHL y riesgos en logística

DHL implementa en su operativa una solución específica para la gestión de riesgos en la cadena de suministro basada en la analítica de Big Data, Resilience360, y proporciona a sus clientes una visión clara de potenciales disruptores en sus cadenas de suministro.

Otras soluciones como 'DHL Parcel Volume Prediction' o la herramienta 'DHL Geovista' se encuentran actualmente en la fase final de desarrollo para ser implantadas como piloto.

Se combina el estudio de la evaluación de los riesgos con instrumentos para la monitorización de la cadena de suministro que operan prácticamente en tiempo real.

Gracias a esto, las cadenas de suministro podrían ser más resilientes, o sea resistentes a riesgos, con lo que se disminuye la prevención de paradas en la producción y también la pérdida de ingresos.

Se analiza la correlación que existe entre factores como las condiciones meteorológicas, las epidemias de gripe y las tendencias de compra online de los usuarios.

El modelo 'DHL Parcel Volume Prediction', por ejemplo, puede facilitar el análisis previo del volumen de cargas de reparto que habrá que gestionar. En este caso, los modelos Big Data ayudan a optimizar los procesos y a mejorar el servicio al cliente.



Business Analytics (III)



General Electric

La potencia industrial está transformando esta compañía en una empresa de software y servicios de IoT enfocada en mejorar la eficiencia, la productividad y los flujos de ingresos de los clientes. Vamos a ver ejemplos.

Fallos en las tuberías: Pequeños sensores de ultrasonidos sujetos a una sección de tubería pueden detectar la corrosión en las paredes interiores de la tubería en tiempo real. Los sensores se pueden montar de distinta forma dependiendo de lo que una empresa de servicios públicos o de petróleo y gas quiere medir. Las lecturas pasan a un nodo analítico local y luego a la nube y mediante un mapa se muestra dónde y cuánto está dañada la tubería.

Ascensores: Un ascensor al subir y bajar crea viento que puede depositar polvo y suciedad en el mecanismo de la puerta. Schindler Group, que gestiona más de 1,5 millones de ascensores y escaleras mecánicas en todo el mundo, se ha asociado con General Electric para desarrollar sensores que puedan detectar esta acumulación mediante vibración.

Es uno de los elementos que se pueden utilizar para el mantenimiento predictivo, junto con la ejecución de algoritmos predictivos, en los ascensores, evitando interrupciones inesperadas.

Control de turbinas

General Electric ha aplicado IoT a las turbinas eólicas, utilizando sensores y análisis de datos para alimentar decisiones sobre qué turbinas funcionan en qué momentos. Si no hay suficiente viento, el coste del funcionamiento de una turbina puede ser mayor que la potencia que podría generar. Los registros meteorológicos, la ubicación de cada turbina, los precios en tiempo real y la esperanza de vida prevista de las piezas basadas en los índices de fallos históricos son importantes en las tomas de decisiones.



■ Business Analytics (IV)



Google y la predicción del virus N1H1

El gobierno de los Estados Unidos recogía datos por todos sus centros de salud, en base a los casos de enfermos que se atendían, y se procesaban esos datos en conjunto para tratar de tener prevista la incidencia de esta enfermedad.

Conseguían hacer predicciones acertadas, pero el problema es que siempre llegaban 2 semanas tarde, es decir la enfermedad ya se había propagado hacia 2 semanas en el lugar en el que ellos predecían que se iba a producir, por lo tanto llegaban siempre demasiado tarde.

En esta situación, google recopiló los términos de búsqueda más utilizados en las zonas y en las épocas pasadas de máxima incidencia de la enfermedad.

Entonces, cuando en el buscador empezaban a introducirse estos términos de forma insistente en una zona concreta, en tiempo real, google fue capaz de predecir dónde se iba a dar un brote de esta enfermedad de forma inmediata, con un porcentaje de éxito altísimo y en el momento en el que el brote estaba comenzando.



RRHH Analytics

- Un 96% de los directores de RR.HH. considera que el Big Data permitirá avanzar para predecir tendencias y lo que pasará en el futuro, de manera que los **departamentos** de **RR.HH.** podrán **anticiparse** a las **necesidades** de la **empresa**, en cuestiones de formación, recursos, promoción, etc...
- Un buen análisis de Big Data puede ayudar a **detectar** quiénes tienen **intención** de **Marcharse** de la empresa. Con esos datos el manager puede anticipar acciones para retener a esos profesionales.
- Si además se analizan datos internos se puede **detectar** a los que son más **innovadores**, líderes, los empleados que están con riesgo de marcharse de la compañía porque están descontentos, etc...



RRHH Analytics

- El **empleado de éxito** presenta características que no están en los criterios de selección actuales, dicen los expertos en recruiting. Ya no es el que posea mejor CV ni mejores notas, ni el que ha causado la mejor impresión en la entrevista personal, sino que tiene que ver más con el Big Data en las redes sociales y el concepto de reclutamiento social: cada vez se reclutan más candidatos a través de las redes y sobre perfiles de LinkedIn y otras plataformas.
- Una de las técnicas usadas es la **gamificación**. A través de juegos se busca identificar las **capacidades** de aprendizaje, de adaptación y de innovación de los candidatos. Además, ver su grado de **frustración** con el que se enfrentan ante determinados problemas y su capacidad analítica. Con ello se pretende conocer sus patrones de conducta para determinar qué tipo de personas son.
- Otras disciplinas usadas son la **inteligencia artificial** para identificar ciertos gestos del lenguaje no verbal del candidato y el **análisis semántico** del mapa que se origina al exponer al candidato a situaciones que se le plantea resolver.
Esto tiene una ventaja para el candidato, ya que se elimina la subjetividad del entrevistador en este tipo de pruebas.
- Este tipo de pruebas están dando muy buen resultado en la búsqueda de candidatos con gran potencial.



■ RRHH Analytics

No hay que olvidarse de que siempre debe existir una **entrevista presencial**. Estos mecanismos ayudan a hacer filtros muy buenos de candidatos, pero a día de hoy, es necesario pasar un último filtro con especialistas en RRHH para asegurar el perfil y, una vez pasada ésta, si el puesto es para el trabajo dentro de un equipo, es recomendable una entrevista o reunión con sus futuros compañeros para ver si el feeling es positivo.



RRHH Analytics (I)



El Big Data en el béisbol, caso MoneyBall

Las decisiones en el mundo de los deportes siempre han estado basadas en dos tipos de factores: personales o subjetivos, y monetarios. Sin embargo Billy Beane, manager general de los Oakland Athletics de la Major League Baseball estadounidense, decidió poner fin a estas limitaciones.

Billy Beane utilizó una serie de métodos estadísticos propios de los mercados financieros para determinar la valía de sus jugadores y de otros potenciales. Llegó a la conclusión de que en el mundo del baseball se prestaba demasiada atención a una serie de estadísticas y se dejaba de lado otras muchas que tenían un gran valor a la hora de seleccionar jugadores, como por ejemplo los porcentajes 'on-base' o 'slugging', que tienen que ver con errores en el bateo y otros conceptos del juego.

Esta aplicación poco convencional del Big Data en el mundo de los deportes llevó a Billy Beane a sentar cátedra entre sus compatriotas, a crear una nueva escuela de pensamiento en el mundo del deporte; sobre esto se hizo la película Moneyball que fue un éxito de masas, con Brad Pitt.

Pero los éxitos no se quedaron aquí, ya que el extrovertido directivo de los Oakland llevó a su equipo a competir con otros como los New York Yankees con un presupuesto mucho más ajustado: \$45 millones de dólares frente a \$125 millones.

Una utilización efectiva del Big Data que muestra que con recursos limitados también pueden obtenerse grandes resultados.



■ RRHH Analytics (II)



Mejorando el rendimiento deportivo

La mayor parte de deportistas de élite están ya adoptando técnicas de análisis de grandes volúmenes de datos.

En tenis se lleva mucho tiempo utilizando la herramienta SlamTracker (basada en la tecnología de IBM SPSS de análisis predictivo) en los torneos más prestigiosos del mundo (Wimbledon, Roland Garros, Open de Australia).

La plataforma lleva registrados más de 8 años de datos de Grand Slams (unos 41 millones de data points) para determinar patrones y estilos de jugadores ganadores.

Aplicando análisis a las grabaciones de vídeo de cada jugador de fútbol y tecnología de sensores en equipamiento deportivo como balones o palos de golf, podemos obtener retroalimentación de los datos generados, y mostrar los resultados para mejorar el rendimiento, o para mostrar una visión enriquecida de la retransmisión a los espectadores.

Muchos equipos de élite realizan ya seguimiento de sus atletas fuera del entorno de competición, usando dispositivos inteligentes para monitorizar desde la nutrición, la preparación física y el sueño, hasta las conversaciones en redes sociales en las que participan para controlar el bienestar emocional.



■ RRHH Analytics (III)



RRHH en Amazon

Después de varias denuncias públicas sobre las condiciones laborales por parte de los trabajadores en algunos centros de Europa, finalmente la compañía decidió actuar. El año pasado este conocido ecommerce inició un proyecto de Big Data para conocer cómo se sentían sus empleados en sus puestos de trabajo.

Mediante un sistema interno llamado "Amazon Connections", consultan periódicamente a sus trabajadores sobre aspectos como la satisfacción laboral, el liderazgo o las oportunidades de formación. Toda la información es analizada y posteriormente compartida con los directivos de la compañía a través de informes diarios.

Las respuestas de los empleados no son anónimas, pero solo los responsables del proyecto de Big Data tienen acceso a ellas. Los informes que llegan a los directivos de la compañía únicamente contienen los datos globales de las respuestas.



Text Analytics

- Las organizaciones necesitan saber lo que sus **usuarios opinan**. Por ello se hace necesario el realizar **análisis de textos masivos**.
- Las primeras dificultades que nos encontramos fueron cómo almacenar semejante cantidad de información y cómo procesarla con cierta rapidez. Esto se solucionó con la llegada del Big Data.
- Luego nos encontramos con el problema de que el análisis de texto no estructurado presentaba **problemas** con el **significado** de las **palabras** ya que una misma palabra puede cambiar de significado según el contexto y que no todos utilizamos la misma palabra para el mismo significado.

Ante esto, el análisis de textos se limitaba a construir **nubes de palabras** y los más osados a realizar **análisis de sentimiento** distinguiendo contenidos positivos de negativos (previamente etiquetados por investigadores de una manera subjetiva).

Sin embargo, el análisis de sentimientos era bastante **poco fiable**, ya que las personas no usamos el lenguaje siguiendo las normas correctamente muchas veces y los análisis se veían superados ante esta circunstancia.

Por eso, los análisis automatizados se centraban principalmente en las nubes de palabras o en buscar correlaciones entre palabras, asumiendo el ruido del lenguaje.



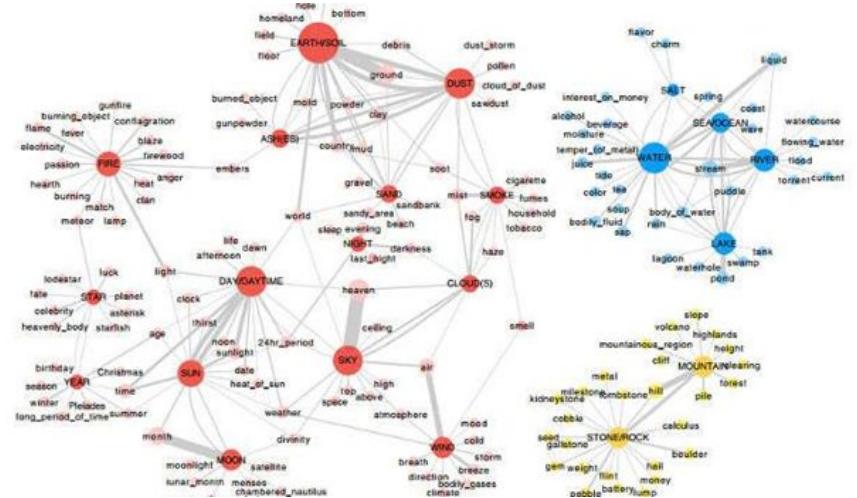
Nubes de palabras

A word cloud centered on the word **ANALYTICS**. The words are arranged in a roughly circular pattern around the center, with larger words indicating higher frequency. The words include:

- VOLUME
- SKILLS
- RESEARCH
- INFORMATION
- BUSINESS
- REPORTING
- BEHAVIOR
- DATA
- ANALYTICS
- OLAP
- SOFTWARE
- CLOUD
- VISUALIZATION
- INTELLIGENCE
- MARKET
- ON-SITE
- COLLECTION
- PROCESS
- RESOURCES
- ESTIMATING
- COMPLEX
- PRACTICE
- METRICS
- INVESTIGATE
- TOOLS
- INTEGRATION



■ Redes semánticas



■ Text Analytics

- Los actuales sistemas de análisis de textos analizan conjuntamente los textos junto con otro tipo de información, como el sexo de quien lo escribe, la hora a la que lo hace, etc.
A día de hoy los sistemas permiten establecer si hay diferencias entre el contenido generado por mujeres y hombres, si ante una queja se genera un contenido específico, etc.
- Realizando este tipo de análisis más avanzado podemos predecir pautas de comportamiento y de consumo de nuestros clientes.



■ Text Analytics (I)



La reelección de Obama

Tras su primer mandato, el presidente de los EEUU, Barack Obama, decidió utilizar Big Data para su reelección en 2012. Un centenar de personas trabajaron en el departamento de analítica de la campaña.

50 estaban fijos en las oficinas centrales, otros 30 se movilizaron a lo largo y ancho de las distintas sedes del país, y 20 estaban únicamente centrados en la interpretación de los datos recibidos.

Tras un primer análisis, los esfuerzos de la campaña se enfocaron en tres aspectos: registro (recoger datos de los votantes convencidos), persuasión (dirigirse a los dudosos de una forma eficaz) y voto del electorado (asegurarse de que los partidarios fueran a ejercer el voto sí o sí).

Y, por primera vez, los tres equipos más importantes de las campañas electorales: el de campo, el digital y el de comunicación, trabajaron con una estrategia unificada con los respectivos datos de cada uno.

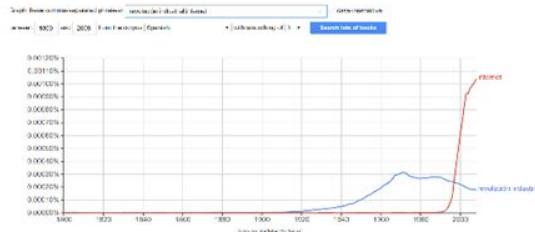
El motor de todo, la plataforma inteligente utilizada fue HP Vertica. Entre las acciones más efectivas que permitía esta plataforma estaban: recoger datos a pie de campo y realizar un feedback muy rápido vía notificaciones email por parte del equipo online (se mejoraba en tiempo y eficiencia); o detectar los nichos en los que funcionaría mejor la publicidad en TV cruzando datos de los votantes con otros demográficos, audiencias, precios de publicidad, programas... (se mejoró en impacto y segmentación).

Con su analítica, el equipo de Obama optimizó la comunicación y mejoró la respuesta del electorado afín, permitiendo no malgastar recursos, tiempo y dinero en los votantes que no eran partidarios de su partido.



■ Text Analytics (II)

Revolución industrial e internet



Culturomics

Este ejemplo viene por un aprovechamiento de los datos almacenados, que no era el objetivo principal de la idea inicial.

Google dispone de un servicio para digitalización de libros, que en primera instancia estaba pensado para preservar libros antiguos que puedan deteriorarse con el tiempo.

En este almacenamiento se dispone de 95 billones de líneas de texto en múltiples idiomas. Pues Culturomics nace como el análisis estadístico de todo este texto. Su objetivo es el estudio del comportamiento humano y tendencias culturales a partir del estudio de estos textos digitalizados.

En la dirección <https://books.google.com/ngrams> se puede utilizar esta herramienta, que está publicada para su uso gratuito.

Por ejemplo, buscamos las palabras “corrupción política” para el idioma castellano, y vemos claramente cómo su uso está disparado de un tiempo reciente a esta parte.

O ponemos revolución industrial, e internet. Se aprecia con toda claridad cómo los libros hablaban mucho de la revolución industrial como tema de estudio hasta hace 30 años, pero a partir del 1990 el término Internet se dispara en su uso, siendo además en valores absolutos mucho más utilizado.

Vemos que con esta herramienta podemos estudiar algunos aspectos de una sociedad, a través de las palabras que se utilizan en los libros.



■ Realidad virtual

- La **realidad virtual** es una **simulación** que reproduce un **mundo** creado en un **ordenador**. Es una actividad en la que el usuario, utilizando dispositivos como cascos, gafas, sensores o mandos, puede **interactuar** con esta realidad alternativa y con todo aquello que la compone.



■ Realidad virtual

- La realidad virtual no es una ciencia nueva. En su momento tuvo una gran aceptación, pero se quedó un poco en hibernación quizás porque la tecnología del momento no era lo suficientemente potente.
- A día de hoy se empieza escuchar mucho de nuevo hablar sobre la realidad virtual y, aunque la mayoría la asociamos a los videojuegos, es mucho más que eso.
- La **visualización** es algo **vital** en los procesos **Big Data**, el cómo muestro la información obtenida para que el usuario pueda usarla para obtener el conocimiento necesario para lleva a cabo la toma de decisiones.
- Éste es uno de los grandes retos porque disponemos de **muchísima información** y hay que conseguir que la mente humana la proceсе de un **modo sencillo**.
- Con la realidad virtual podemos conseguir que el usuario tenga una **experiencia** más **cercana** que le permita identificar y visualizar lo que necesite de un modo más sencillo y rápido.



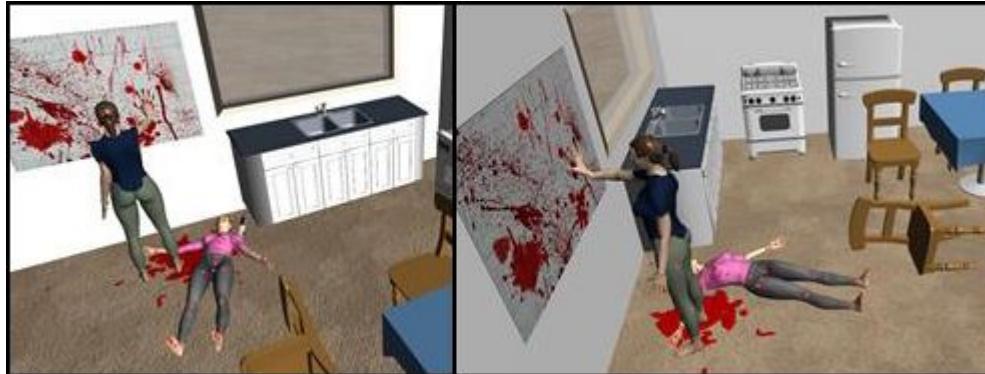
■ Realidad virtual

- No hablamos sólo del **usuario interno** que necesita analizar la información disponible para llevar a cabo decisiones estratégicas del negocio, sino también de los **clientes**, a quienes se les puede mostrar lo que necesiten de un modo más amigable, lo que se traduce también en una mejor experiencia de usuario que puede traducirse en una venta y una mejor opinión de la empresa.
- Uno de los sectores pioneros fue el de la **salud**. Crean simulaciones para la detección y diagnóstico de enfermedades, la usan como herramienta terapéutica ante fobias o como técnica de relajación para procesos de estrés, entre otra infinidad de usos.



■ Realidad virtual

- En **ingeniería** se usa para la creación de prototipos dónde poder inspeccionarlos antes de implementarlos físicamente.
- En el sector del **turismo** se está apostando fuerte por crear entornos virtuales que permitan visualizar los destinos sin estar allí.
- Quien haya visto la serie *CSI* conoce de primera mano, aunque en su momento fuera ciencia-ficción, lo que la realidad virtual puede facilitar la reconstrucción de escenarios de un **crimen** para analizarlos.



■ Realidad virtual

- En los **entrenamientos especiales**, como los militares, espaciales o de emergencias, que permiten realizar simulaciones para saber cómo actuar antes de que se vean involucrados en ellas en la vida real.
- En el **cine** se están planteando que la persona deje de ser un simple espectador para formar parte de la película.
- En el **e-commerce** y en el **retail** la posibilidad de que no navegués simplemente por un catálogo, sino que puedas ir por la tienda virtual para realizar la compra como si de una tienda física se tratase.



■ Realidad virtual

- Sin embargo, a día de hoy es una ciencia muy **cara** de implantar y, aunque avanzamos rápidamente en **hardware** cada día más **potente**, todavía es necesario que madure más para que lleguemos a verla más asentada en nuestro día a día.



■ Robótica

- Según Gartner, “**la explosión del Big Data llegará con los robots**”.
- A día de hoy, ya sabemos de la importancia que están tomando los **robots** en el sector **industrial** y en el sector de **manufacturas**.



■ Robótica

- Gracias a la gran cantidad de datos que tenemos a día de hoy, el que los robots estén cada vez más inmersos en otros sectores e incluso en nuestra vida cotidiana se está volviendo cada vez más sencillo.
- Cuantos más datos tenemos, más capaces son los robots de **imitar** el comportamiento **humano**.
- En los **aviones** en la actualidad se disponen de sistemas de despegue y aterrizaje automático sin necesidad de presencia humana. Éste es un primer paso en la evolución hacia aviones no tripulados, en los que los accidentes provocados por errores humanos se minimicen.
Hay que decir que ya existen aviones no tripulados, lo que conocemos como drones, aunque de momento son aviones no de pasajeros, su uso es para toma de imágenes en tiempo real al sobrevolar determinados escenarios e incluso llevar a cabo alguna acción que se le transmita de manera remota por humanos, al menos de momento.



■ Robótica

- Otro escenario en el que los robots empiezan a tener especial importancia es como **analistas de negocio**, ya que son capaces de analizar mucha información pasada y presente y aportar distintas alternativas para ayudar a la toma de decisiones.
- En la **medicina** vemos como estos robots ya están ayudando por ejemplo en operaciones complicadas. Aunque su uso va más allá, ya que, conociendo el historial médico del paciente, sus características e incluso su ADN y comparándolo con el de otros pacientes, se puede llegar a diagnosticar y prescribir los medicamentos más adecuados para cada persona.



■ Robótica

- Con los avances en el reconocimiento del lenguaje natural, cada vez está más cerca que muchos **Call Centers** empiecen a utilizar robots para responder a las llamadas por su capacidad de análisis y aprendizaje.



■ ¿Opiniones?

- ¿Los robots nos quitarán el trabajo?



Otros casos de uso (I)



El Bot Mitsuku

Mitsuku es una joven de 18 años, o mejor dicho, esa es la personalidad de este bot conversacional que ha conseguido ser el mejor del año 2016 haciendo pasar por un humano.

Los bots están muy de moda, y más que lo estarán, así que es mejor que vayamos acostumbrándonos a ellos porque quizás la próxima vez que compres una pizza a través de una aplicación, en la otra parte haya un bot entendiendo todo lo que le dices y, por supuesto, siguiendo la conversación con total naturalidad.

Un chatbot es justamente eso, un bot con el que puedes hablar, un bot conversacional, un programa que ha sido creado para ser capaz de mantener una conversación natural con un humano, analizando sus palabras y contestando en consecuencia.

Mitsuku habla inglés, así que nosotros también debemos escribirle en ese idioma. Le puedes contar lo que quieras, aunque ella también te hará algunas preguntas; es capaz de mantener conversaciones naturales y fluidas. Este bot conversacional habla cada día varios a miles de personas, ya no sólo a través de su web, sino que también está integrado en Pandorabots y en Kik Messenger, la aplicación de mensajería. Y aprende de cada una de las conversaciones que mantiene.

Por ejemplo es mejor hablarle con frases cortas y bien escritas, es capaz de recordar datos e información que anteriormente has compartido con ella y si crees que Mitsuku se ha equivocado en una respuesta, puedes corregirla.

Eso sí, no es un aprendizaje automático, todas esas sugerencias deben ser aprobadas por su creador, que estudiará si incluir, o no, esas respuestas en su memoria (es obvio por qué lo hace así, en Internet hay mucho bicho raro... tenemos ejemplos con algunas respuestas inadecuadas que aprendió Siri en sus primeras versiones).

Mitsuku, el resto de bots, el aprendizaje con redes neuronales y la inteligencia artificial como otro capítulo del mundo Big Data es un campo en plena expansión y podemos ir preparándonos porque ya están dando el salto a servicios reales.



Otros casos de uso (II)



Watson y la NASA

El sistema de computación cognitiva Watson ayuda a investigar a la NASA, tamizando montañas de datos digitales para recoger rápidamente información útil e incluso ayudar en la toma de decisiones importantes.

Watson ha desarrollado enormemente su capacidad de procesamiento del lenguaje natural. Muchos descubrimientos suceden por pura casualidad porque alguien lee un artículo justo después de leer otro y conecta los puntos.

Entonces Watson lee rápidamente un gran conjunto de documentos técnicos, documentos y publicaciones, y une los puntos que un investigador con tiempo y recursos limitados puede no ser capaz de conectar por su cuenta.

La NASA e IBM también están trabajando para desarrollar un sistema que podría servir como asesor de operaciones de vuelo. Watson leería una gran colección de documentos relacionados con la aviación y se convertiría en asesor de los centros de operaciones de las aerolíneas comerciales. La NASA también quiere usarlo como asesor de las tripulaciones de vuelo en tiempo real.



Otros casos de uso (III)



Cortana será Einstein

Hay un proyecto de Inteligencia Artificial para que Cortana forme parte de algo más grande.

Según publicó Reuters, Microsoft ya está trabajando en un nuevo asistente de voz que incorporaría lo mejor de Cortana, pero iría aún más lejos:

El proyecto, llamado Einstein, desarrolla una tecnología que permite adelantarse a las peticiones del usuario en función de su ubicación física y el tiempo concreto del día. Sería un salto cualitativo en los asistentes de voz: una tecnología proactiva, no sólo reactiva.

La idea que hay detrás de Einstein es conseguir un asistente de voz que sea capaz, con técnicas de aprendizaje automático, saber dónde está el usuario y qué quiere hacer, qué busca o qué necesita en ese momento. Pura Inteligencia Artificial.



Otros casos de uso (IV)



Coche autónomo de google

El análisis de big data está ayudando a máquinas y dispositivos a ser más inteligentes y autónomos. Un ejemplo que ya es una realidad, es el coche autopilotado de Google.

Los coches que usan para el proyecto están equipados con cámaras, GPS, conexión a internet, y un abanico de procesadores y sensores que permiten al vehículo circular de forma segura por la vía pública sin necesidad de intervención humana.

También se usan herramientas de análisis de big data para optimizar las redes de energía a partir de datos de los medidores inteligentes.

El tráfico a tiempo real también es información a la que tienen acceso estos vehículos, que disponen de dos estilos de conducción. Uno de ellos es más cauto y en una situación de duda cedería el paso a otros coches, mientras que el segundo es más agresivo y está programado para salir primero.

El director de producto de Google ha dicho que en 5 años la tecnología estará más que lista y probada, pero que el problema del plazo de su implantación está en los aspectos legales.



Otros casos de uso (V)



Tráfico y Big Data

Investigadores de la Universidad de Granada (UGR) han desarrollado un nuevo sistema informático, basado en técnicas de 'big data', que permite predecir cuándo se va a producir un atasco de tráfico y ayuda a prevenirlo. Este proyecto se denomina 'Predicción del Estado del Tráfico' o PETRA.

La investigación ha sido financiada por la Dirección General de Tráfico (DGT) y llevada a cabo por el Grupo de investigación GeNeura del departamento de Arquitectura y Tecnología de Computadores de la Universidad de Granada.

El objetivo del proyecto PETRA ha sido crear un sistema que permita ofrecer información completa y en tiempo real del tráfico en carreteras interurbanas, para poder predecir la densidad del tráfico en diferentes marcos temporales, y ofrecer esa información como datos abiertos para su consumo por parte de aplicaciones para el móvil.

Actualmente existen diferentes medios que permiten conocer, en tiempo real, el estado de ciertos tramos de las carreteras, como por ejemplo las cámaras de seguridad, o los aforadores.

Para el proyecto PETRA, los investigadores desplegaron una serie de dispositivos en distintas zonas urbanas e interurbanas de Granada, Málaga y Almería, que recogían información de los dispositivos móviles (mediante Bluetooth) para realizar seguimiento de los vehículos que pasaban cerca de esos nodos, almacenando su identificador y una marca de tiempo.

Los resultados de ese prototipo fueron una serie de mediciones, tales como paso de vehículos por días, horas, y velocidades medias, en un área limitada que demostraban la viabilidad de la propuesta.

Predicción a partir de los datos: Este proyecto ha permitido obtener datos nuevos que puedan compararse y ser analizados con otros medios de obtención de datos, y realizar análisis de predicción sobre ellos.

El análisis de los datos ha servido para proporcionar información relativa a predicción de atascos, uso de las carreteras, velocidades medias o información geolocalizada sobre el estado de las carreteras.

Asimismo, se ha propuesto la extensión de los sistemas de difusión de los datos obtenidos usando servicios web, creando una serie de datos de acceso público para que futuros investigadores puedan realizar sus análisis, facilitando así el Open Data.

Se puede llevar el análisis un paso más allá y mirar con más detalle los movimientos diarios de masas: de norte a sur de una ciudad, de este a oeste en determinada región... y con más detalle, el movimiento de las personas de una misma empresa que trabajan en la misma sede... ¿se les podría proponer compartir vehículo? Esta línea de trabajo daría mucho juego...

Con este proyecto, se ha demostrado que es posible ahorrar tiempo, dinero y emisiones de CO₂ si se controlan adecuadamente los flujos de tráfico, los cruces, los semáforos, las rotundas, los aparcamientos y el alumbrado nocturno, por citar sólo los ejemplos más evidentes.

En este proyecto cobraban especial relevancia: la alta disponibilidad, la seguridad de la información, la capacidad de cómputo, las posibilidades de Big Data y de Open Data.



Otros casos de uso (VI)



Dinámicas de turismo en la ciudad de Madrid

Bajo este nombre, el Ayuntamiento de Madrid analizó el comportamiento de los turistas en cuanto a su actividad comercial.

Entre los muchos resultados, el estudio sirvió para cuantificar el impacto económico de eventos concretos en diversas zonas de la ciudad como el Orgullo Gay, una final de la copa del rey, o determinados congresos de gran afluencia en el recinto ferial.

En alguno de estos eventos se detectó que el gasto comercial aumentó un 24% respecto a la misma semana del mes anterior.

Además, se conocen otros datos interesantes como los turistas que más gastan, en qué gastan, por dónde se mueven, etc. Lo cual puede dirigir decisiones en cuanto a incrementar los servicios públicos de transporte entre determinadas zonas, la seguridad ciudadana, los servicios de limpieza, etc...



Otros casos de uso (VII)



Aristóteles y las energías renovables

Kaiserwetter es una empresa alemana que ha diseñado un sistema de Big Data basado en el internet de las cosas, llamado Aristóteles.

Aristóteles combina el análisis del Big Data, el internet de las cosas y una instalación digital en la nube.

Internet de las Cosas permite la agregación de datos desde prácticamente cualquier lugar del mundo, además de realizar análisis de las desviaciones a partir de los datos esperables para determinadas condiciones meteorológicas; se pueden analizar los datos de rendimiento y los datos financieros de cada unidad de producción (por ejemplo cada molino de viento) a través de distintos niveles de agregación de datos.

La idea es que las empresas de energía, los inversores, los bancos y las aseguradoras podrán reemplazar prácticamente las funciones de los departamentos de Controlling (simplificando así sus estructuras empresariales) y concentrarse en su negocio; es decir que los directivos pueden consultar toda la información en las pantallas de sus ordenadores de manera compacta, transparente y en tiempo real.

Esta empresa dice literalmente que «la transición energética constituye el mayor proyecto informático de todos los tiempos, aparte de que abre puertas a nuevas áreas de negocio».



Otros casos de uso (VIII)



Prevención del delito

El procesamiento de grandes volúmenes de datos se ha convertido en una herramienta para predecir cuándo y dónde tendrá lugar un delito. Así, la policía envía patrullas a la zona para disuadir a los potenciales delincuentes de cometerlo.

La Policía de Memphis (EE UU) afirma que gracias al big data ha reducido los delitos serios en un 30%, con un descenso del 15% en sucesos violentos.⁶ En Richmond, la capital del estado de Virginia, la policía logró una caída de casi el 30% en crímenes violentos y homicidios, para una determinada área y en un periodo de 12 meses.

También en Estados Unidos, la ciudad de Reading, en Pensilvania, registró una disminución de delitos tras instalar un software predictivo. En 14 meses los robos se redujeron un 23%, pese a que la policía contó con menos efectivos. La clave estuvo en el algoritmo de PredPol, una startup californiana que usa big data para predecir cómo se comportará la delincuencia.

El modelo tiene en cuenta la predicción de delitos y los recursos de la policía, para optimizar la ruta de las patrullas y cubrir los espacios potencialmente más problemáticos.

Las predicciones se hacen para las próximas diez o doce horas. La información que se usa son los datos históricos: qué tipo de delito es, dónde ocurrió y cuándo. También hay que combinar el contenido de informes policiales, a veces escritos a mano, con información meteorológica (porque la criminalidad varía si llueve o nieva) y con ciertas coyunturas, como si es día de pagos (porque en Estados Unidos muchas empresas aún utilizan cheques).

También se trabaja con listas de sospechosos. En Miami Bay reducen las listas de sospechosos o incluyen algunos nuevos que ellos no habían pensado.

No es un Minority Report, no se trata de predecir quién va a cometer un crimen o se va a convertir en un criminal: El algoritmo no incluye información personal, como quiénes son los delincuentes, su aspecto, su raza o su condición socioeconómica.

El software nunca deja de aprender, nunca descansa, por ejemplo si la policía está teniendo impacto en el índice de crimen de un área y este empieza a subir en otra, el algoritmo adaptará sus predicciones para afrontar los nuevos patrones.





■ 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=

