



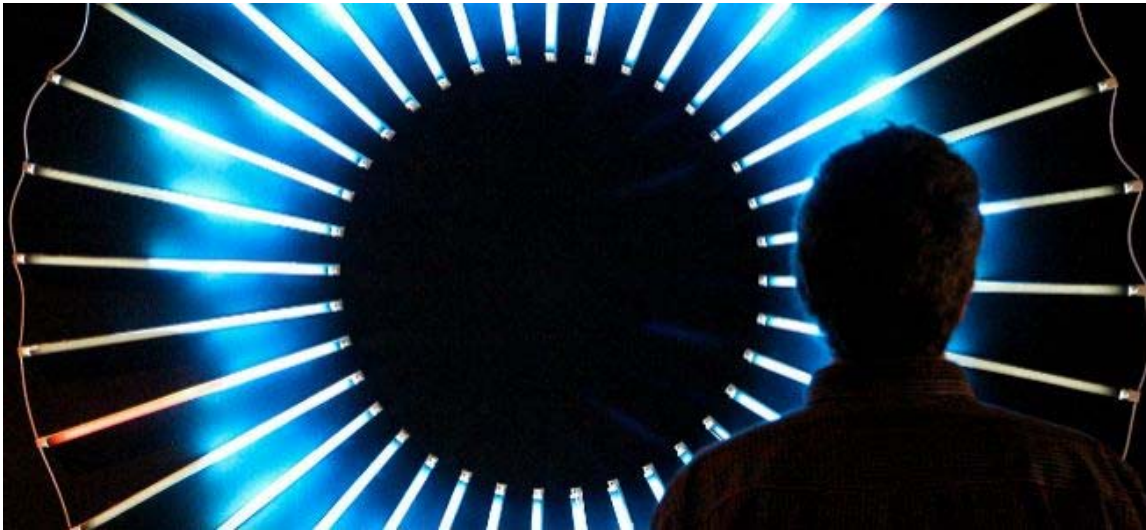
Big Data

M E C D

Carmen Cabanillas Serrano
Subdirectora General Adjunta
Tecnologías de la Información y las Comunicaciones
Ministerio de Educación, Cultura y Deporte

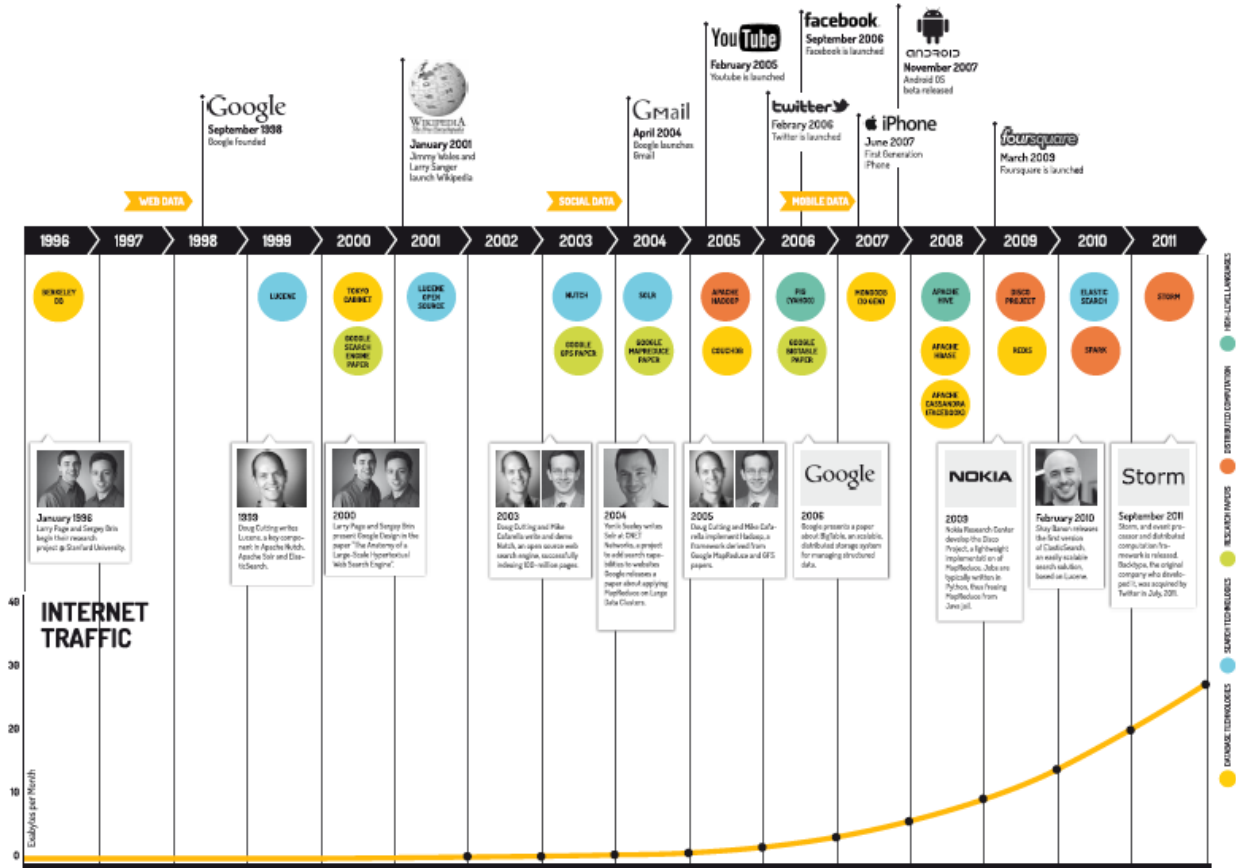
15 de noviembre de 2016

Índice



Historia

BIG DATA A BRIEF HISTORY



Historia

COMUNICACIÓN



1^a mitad
siglo XX

Uno comunicaba y
todos escuchaban



2^a mitad
siglo XX

Varios comunicaban y
todos escuchaban

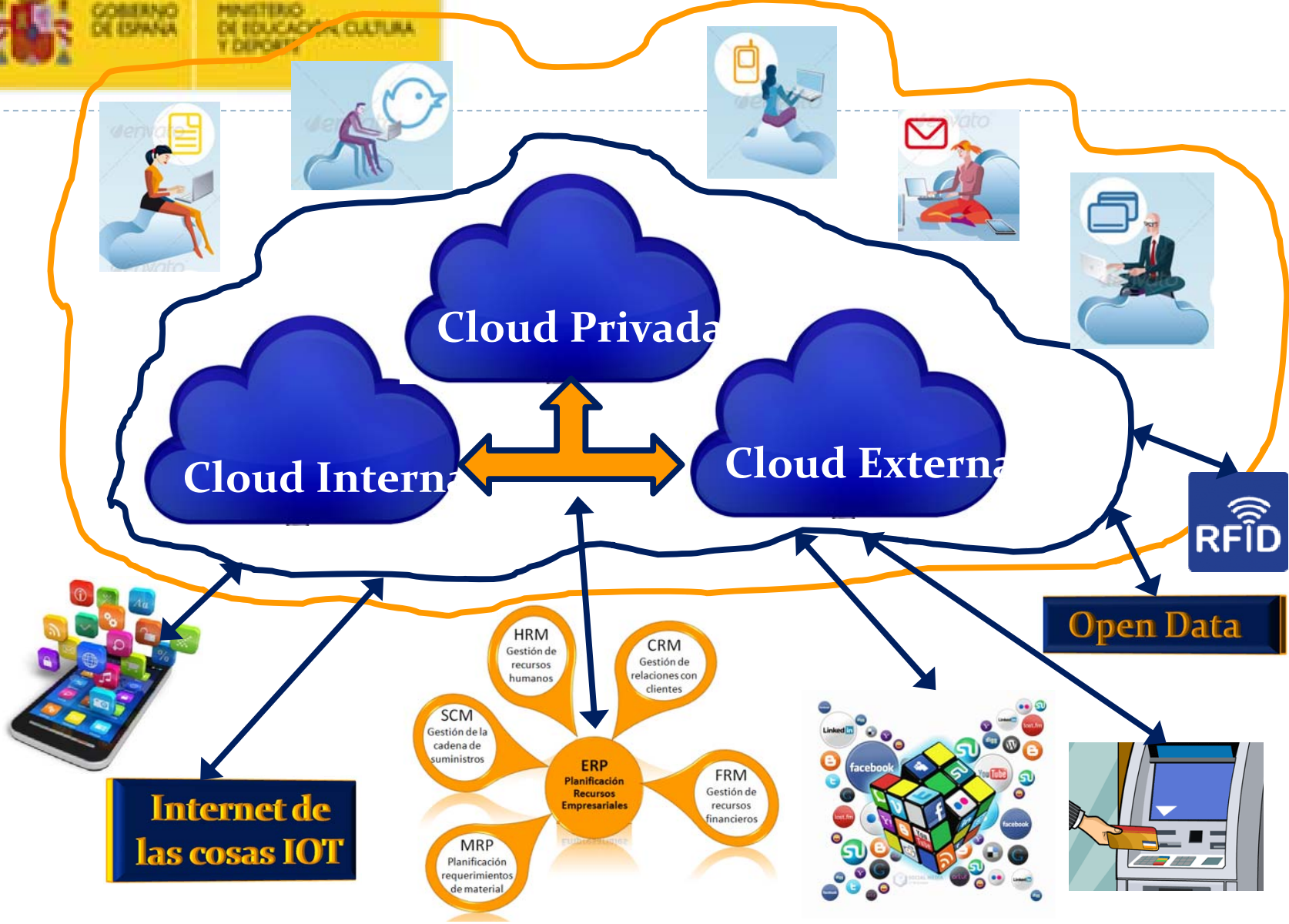


Hoy

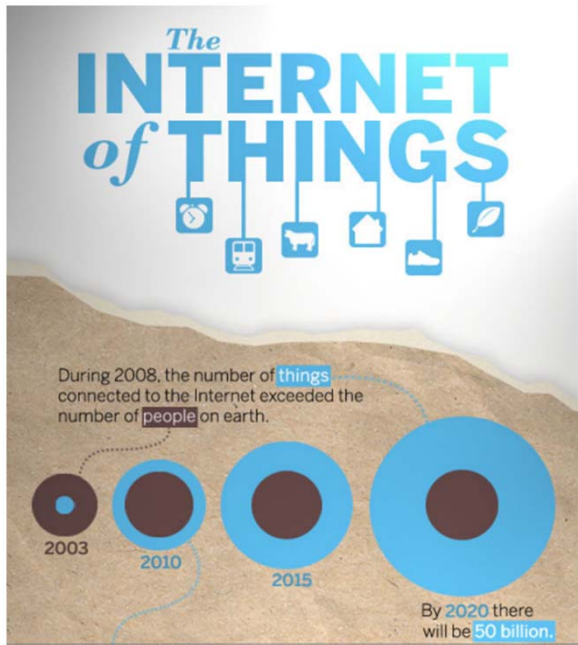
Muchos comunican y
muchos escuchan



Fundamentos

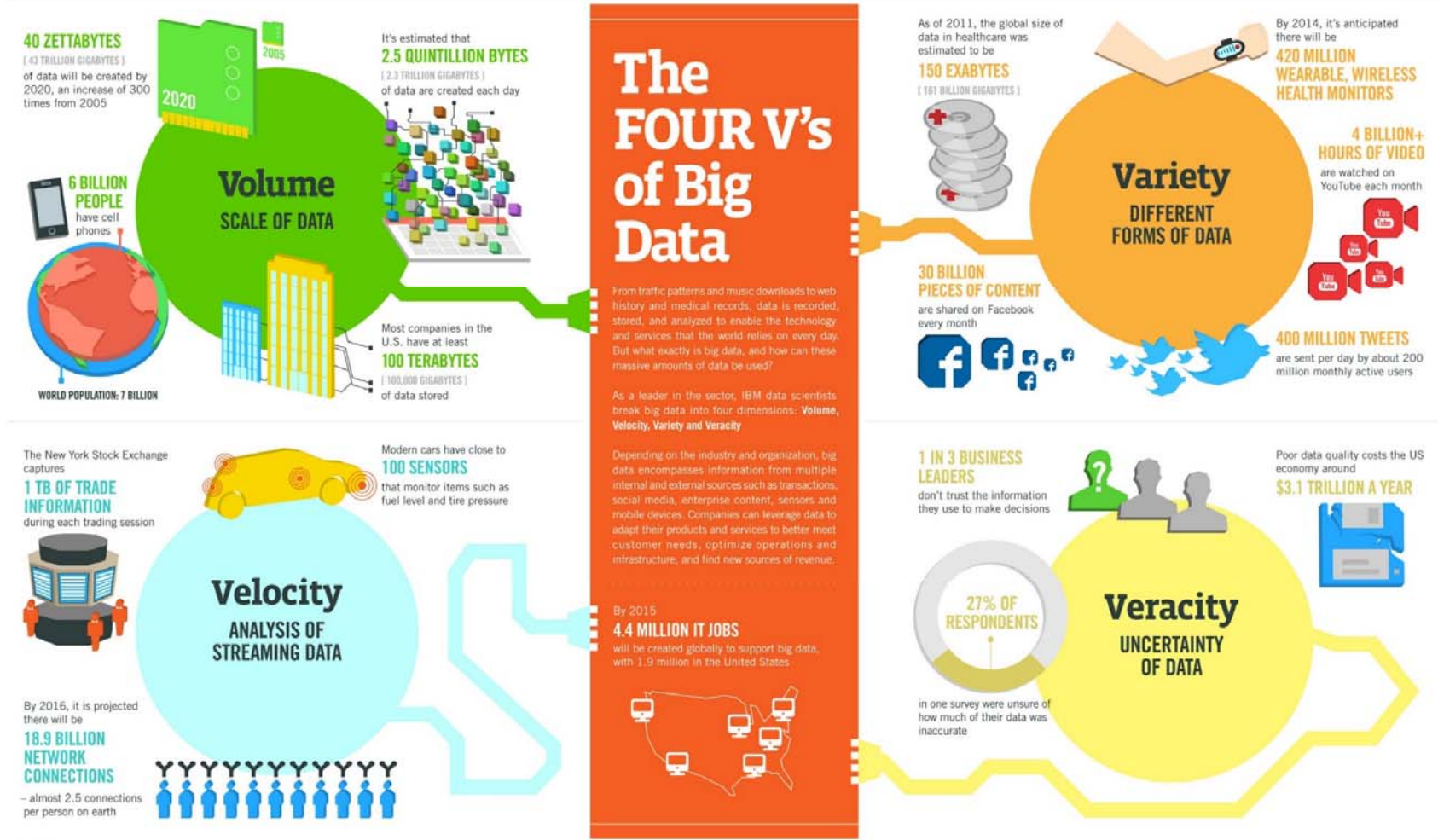


Fundamentos



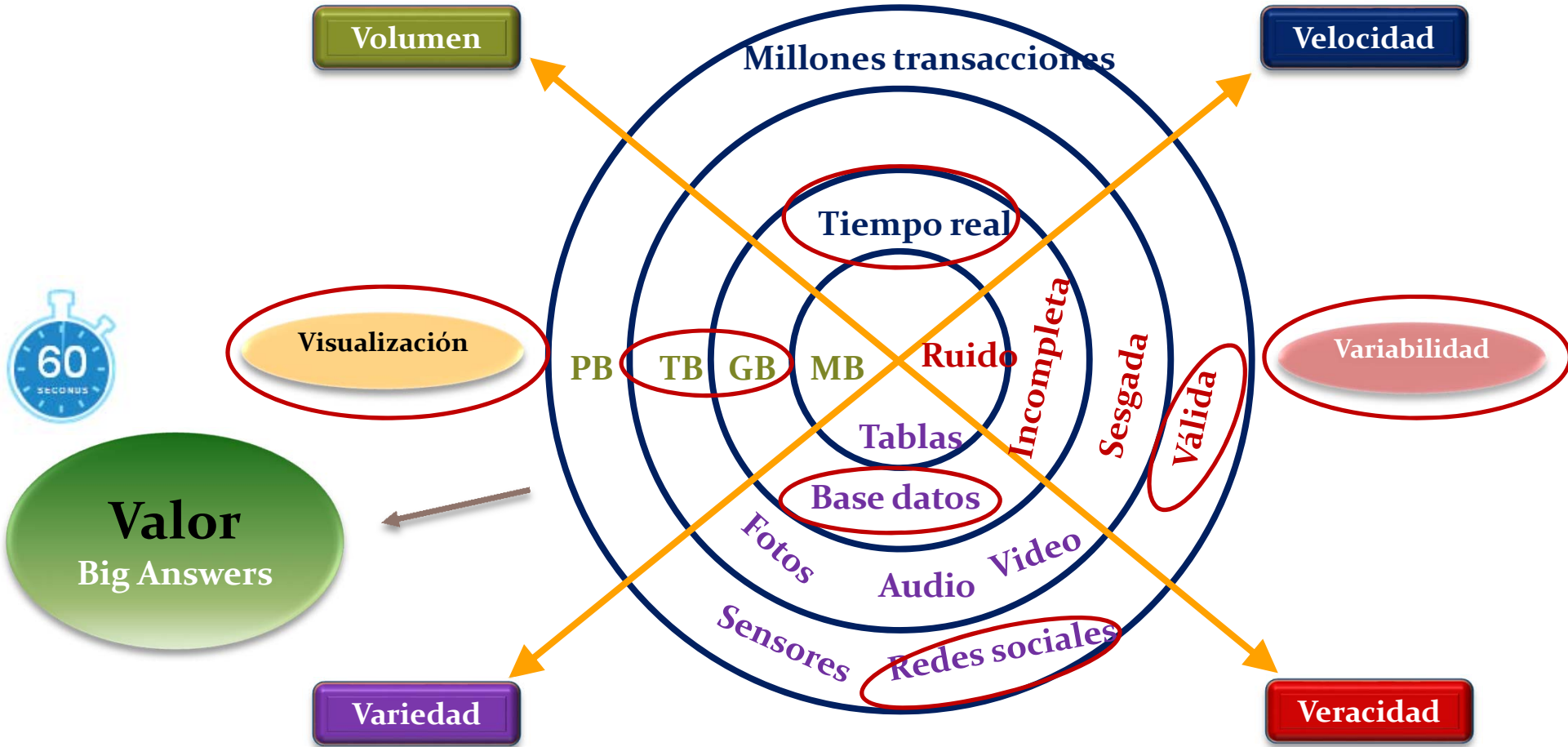
Internet es el motor del crecimiento

Fundamentos



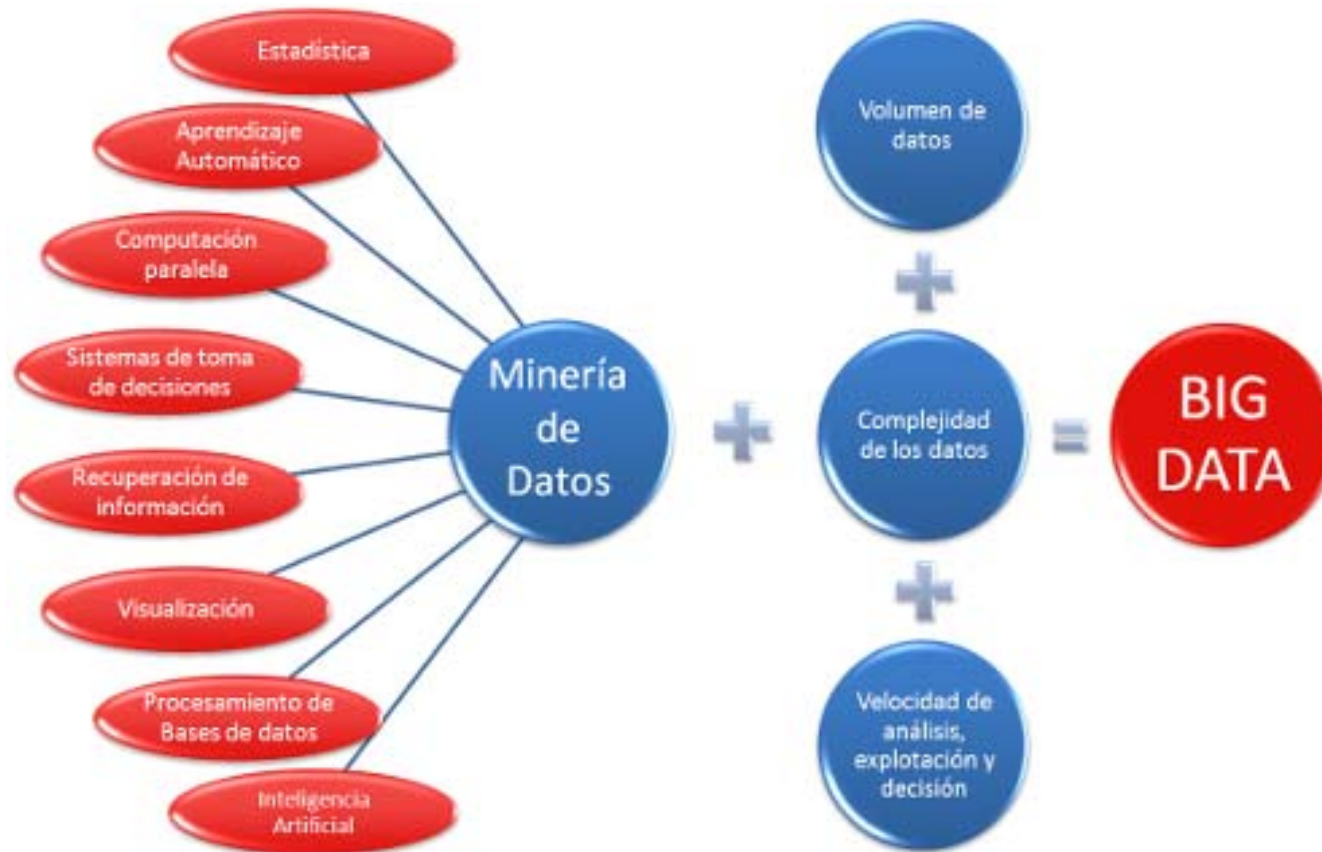
Sources: McKinsey Global Institute, Twitter, Cisco, Gartner, EMC, SAS, IBM, MEPTec, QAS

Fundamentos

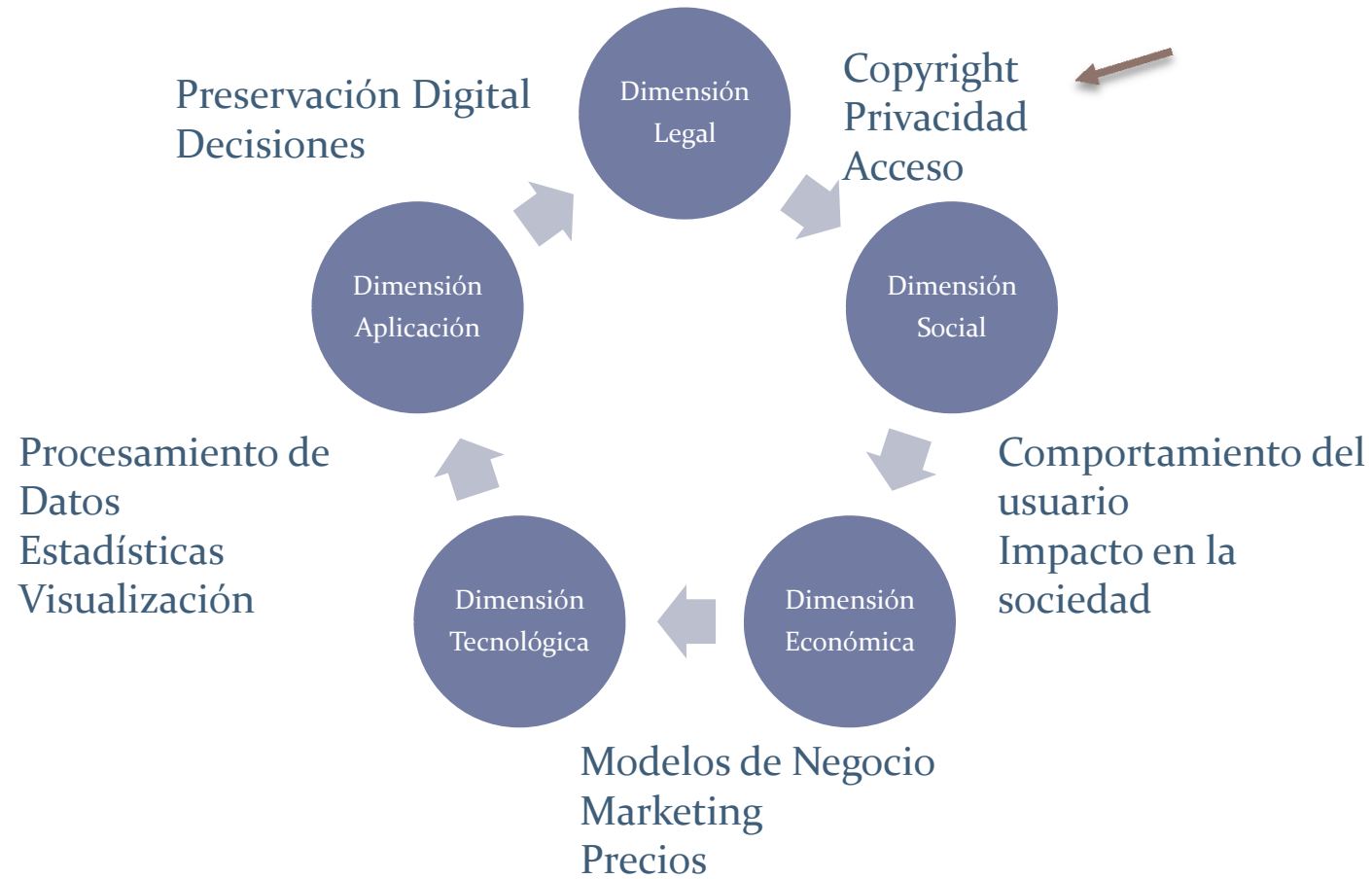


Visión del negocio en tiempo real

Fundamentos



Fundamentos



Fundamentos

Traditional Approach *Structured & Repeatable Analysis*

Business Users

Determine what question to ask



IT

Structures the data to answer that question

Monthly sales reports
Profitability analysis
Customer surveys



Big Data Approach *Iterative & Exploratory Analysis*



IT

Delivers a platform to enable creative discovery

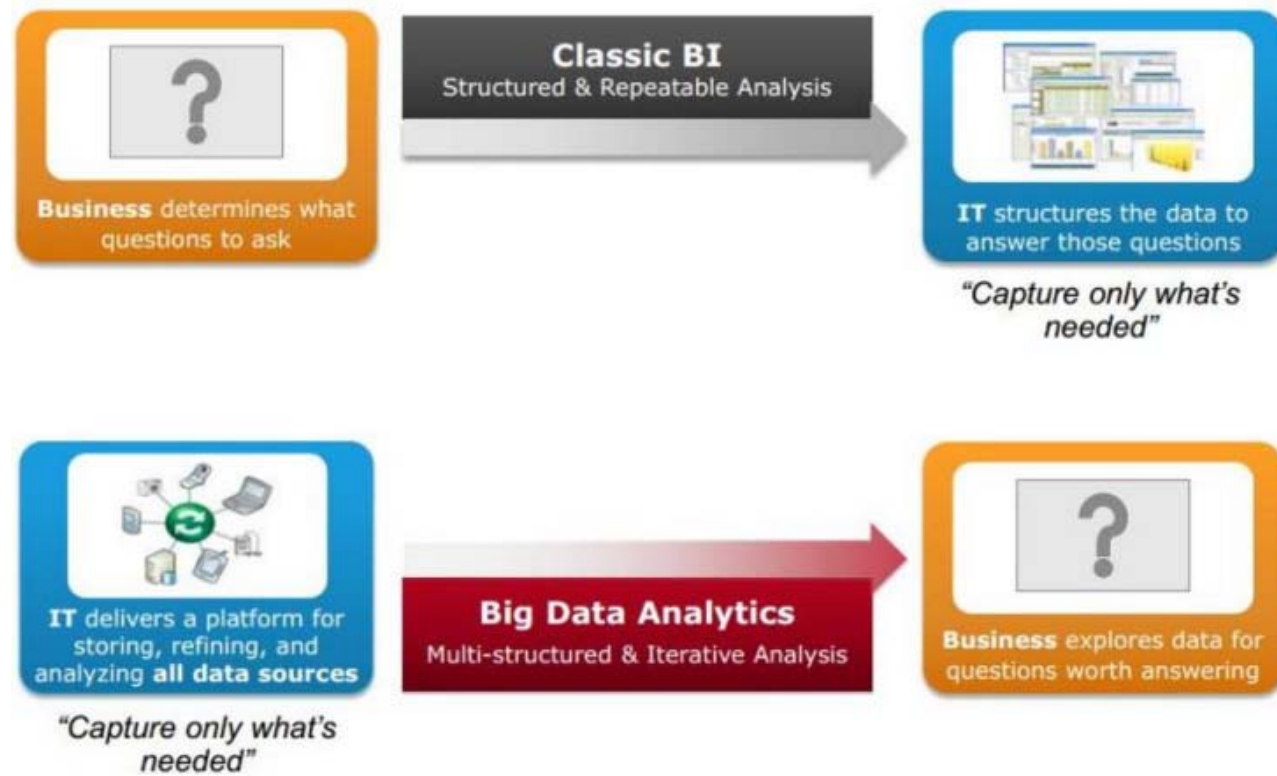
Business

Explores what questions could be asked

Brand sentiment
Product strategy
Maximum asset utilization
Preventative care



Fundamentos



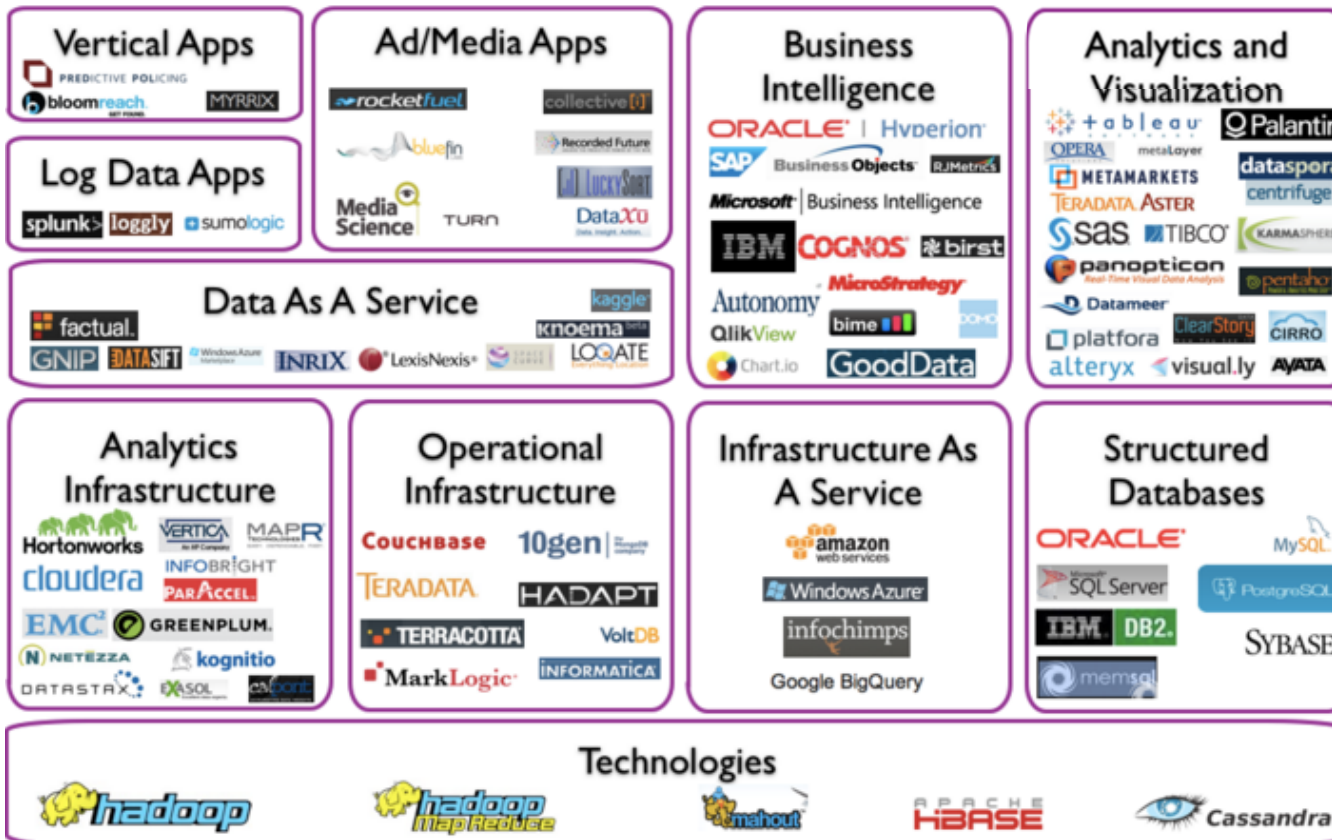
Proyectos



Cuadrante Mágico de Gartner sobre Plataformas Analíticas y de Business Intelligence 2016

Proyectos

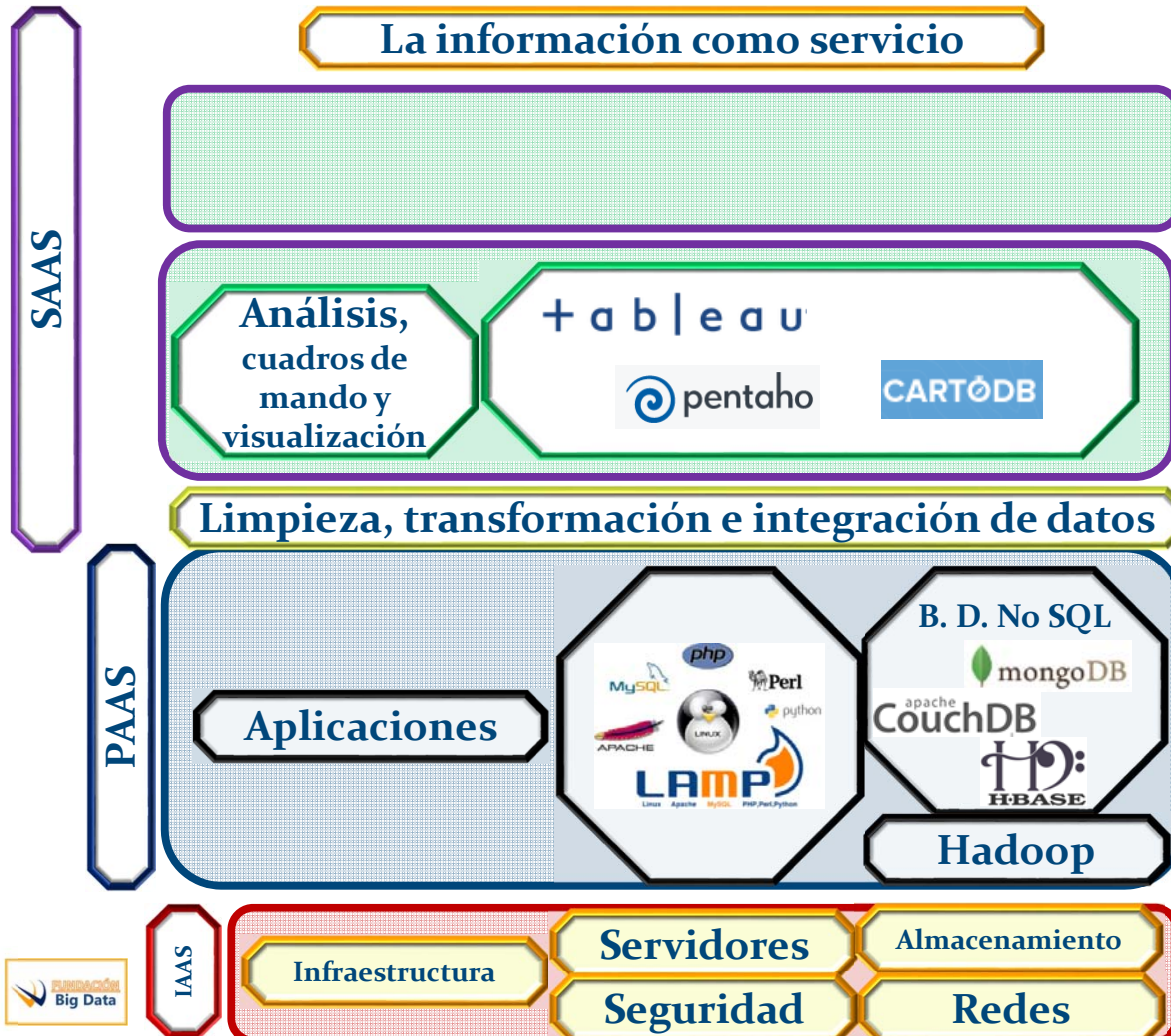
Big Data Landscape 2012



Proyectos

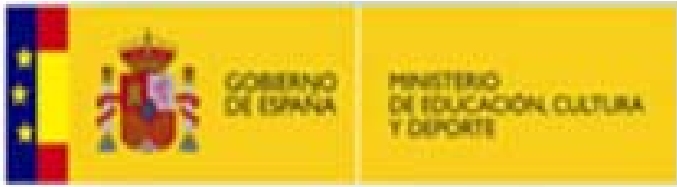


Proyectos



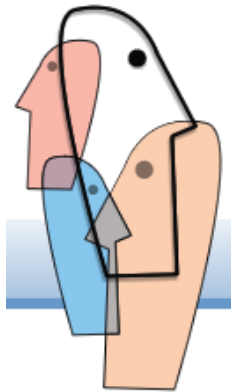
Fuente:



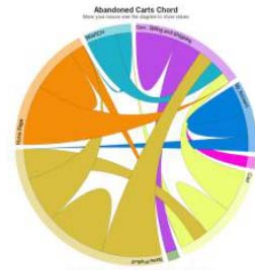


Proyectos

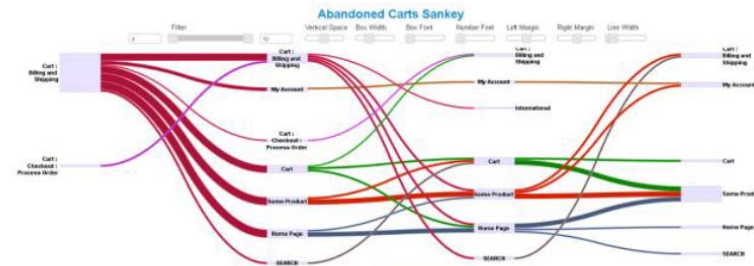
Visualización de datos



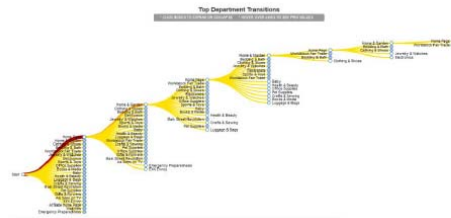
Many Eyes



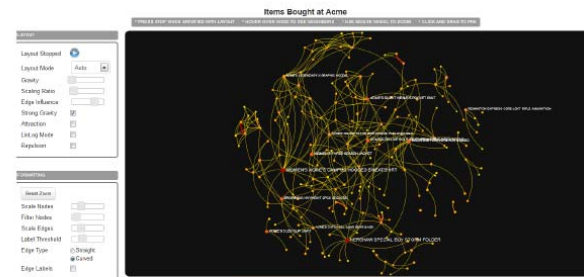
Chord



Flow

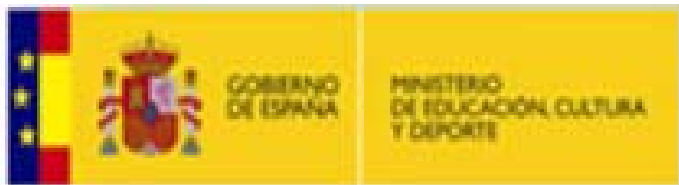


Hierarchy



Affinity





Proyectos Proyección

Empresas del mundo por capitalización bursátil Mayo 2016		
Ranking	Empresa	Cap. Bursátil (miles millones €)
1	Apple Inc.	488,92
2	Alphabet Inc-A (Google)	448,84
3	Microsoft Corp.	375,54
4	Exxon Mobil Corp	310,37
6	Facebook Inc-A	280,32
9	Amazon.Com Inc	238,21
32	Oracle Corp	142,77

66 Inditex - 89,79
Ibex 35 >100

Entorno Big Data crece al 12,5% (TIC 6%-8%)

Empleo en España

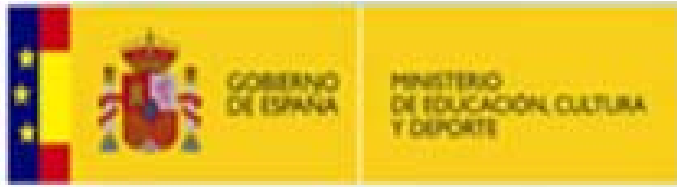
La demanda supera a la oferta

Perfil

2/3 son técnicos (java, base de datos, B.I., etc.)

1/3 son marketing, comercial, etc.

Ranking	Perfil
1	Analistas de datos
2	Programadores
3	Arquitectos tecnológicos
4	Data manager
5	Consultores B. D.
6	Product manager



Proyectos

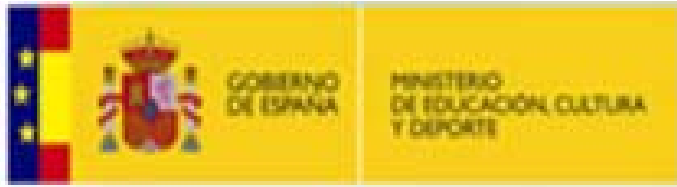
▫ Ejemplos: Juegos OnLine

▫ Antes del Big Data

- Recibir y procesar un elevado nivel de eventos era casi imposible
- Tecnologías principales:
 - Bds relacionales (i.e. MySql, Oracle)
 - Sistema de colas con consumidores

▫ Con Big Data

- Estadísticas y respuesta a eventos en tiempo real
- Tecnologías principales:
 - Storm, Erlang
 - NoSQL



Proyectos

Ejemplos: Anuncios Clasificados

Lo que se podía antes del Big Data

- Hacer portales de clasificados (inmobiliarios, etc) a nivel de un país.

- Tecnologías principales:

 - Bases de datos relacionales (i.e. MySQL, Oracle)

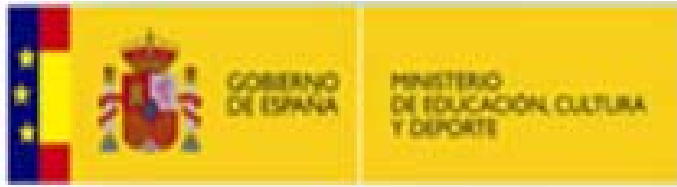
Lo que era casi imposible y ahora es bastante más sencillo

- Hacer un buscador por categorías a escala mundial

- Tecnologías principales:

 - Hadoop, Hbase, Storm

 - Solr, ElasticSearch

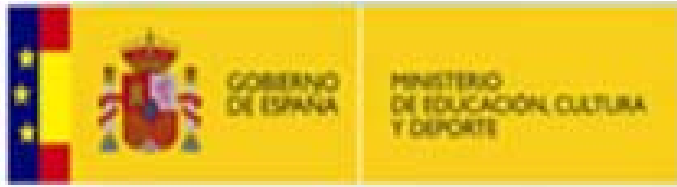


Proyectos

Análisis de influencia online y en redes sociales

- Lo que se podía antes del Big Data
 - Análisis en el servidor (análisis locales)
 - Tecnologías principales:
 - Bases de datos relacionales (i.e. MySQL)
 - Librerías en servidor.

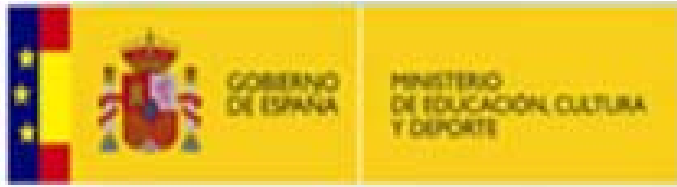
- Lo que era casi imposible y ahora muy razonable
 - Analizar la influencia de todos los individuos en una red social
 - Tecnologías principales:
 - Hadoop, Hbase, Storm
 - NoSQL: Voldemort, Cassandra
 - Hama, Giraph



Proyectos

Explotar datos agregados

- Lo que se podía antes del Big Data
 - Datos en silos
 - Presos de los DW (muy poco flexibles)
 - Tecnologías principales:
 - Datawarehouses
- Lo que era casi imposible y ahora es muy razonable
 - Explotar de nuevas maneras los datos y dar valor a terceros
 - Tecnologías principales:
 - Hadoop
 - NoSQL,
 - Gestores geográficos



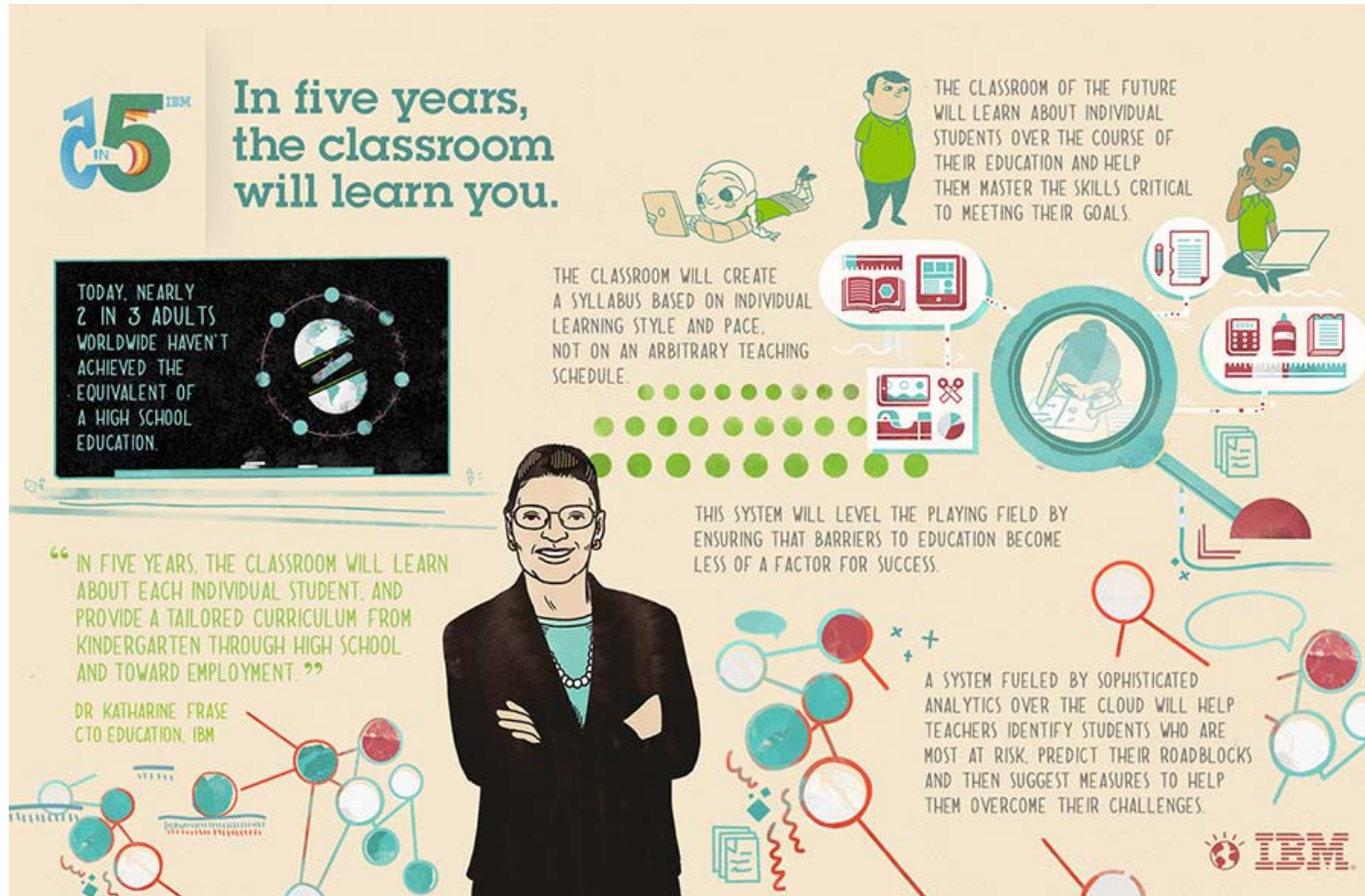
Proyectos

Redes de anuncios (Ad networks)

- Lo que se podía antes del Big Data
 - Estadísticas básicas en Bases de datos relacionales de clientes
 - Agregados de muy alto nivel (se pierde el grano fino)
 - Tecnologías principales:
 - BDs relacionales (i.e. MySql, Oracle)

- Lo que era casi imposible y ahora es muy razonable
 - No eliminar ni un dato
 - Estadísticas muy ricas y con mucho detalle para cada cliente
 - Tecnologías principales:
 - Hadoop
 - NoSQL, Splout SQL

Educación



15 IN **5** YEARS

In five years, the classroom will learn you.

THE CLASSROOM OF THE FUTURE WILL LEARN ABOUT INDIVIDUAL STUDENTS OVER THE COURSE OF THEIR EDUCATION AND HELP THEM MASTER THE SKILLS CRITICAL TO MEETING THEIR GOALS.

THE CLASSROOM WILL CREATE A SYLLABUS BASED ON INDIVIDUAL LEARNING STYLE AND PACE, NOT ON AN ARBITRARY TEACHING SCHEDULE.

THIS SYSTEM WILL LEVEL THE PLAYING FIELD BY ENSURING THAT BARRIERS TO EDUCATION BECOME LESS OF A FACTOR FOR SUCCESS.

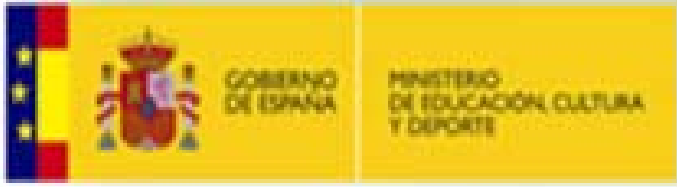
A SYSTEM FUELED BY SOPHISTICATED ANALYTICS OVER THE CLOUD WILL HELP TEACHERS IDENTIFY STUDENTS WHO ARE MOST AT RISK, PREDICT THEIR ROADBLOCKS AND THEN SUGGEST MEASURES TO HELP THEM OVERCOME THEIR CHALLENGES.

TODAY, NEARLY 2 IN 3 ADULTS WORLDWIDE HAVEN'T ACHIEVED THE EQUIVALENT OF A HIGH SCHOOL EDUCATION.

“IN FIVE YEARS, THE CLASSROOM WILL LEARN ABOUT EACH INDIVIDUAL STUDENT, AND PROVIDE A TAILORED CURRICULUM FROM KINDERGARTEN THROUGH HIGH SCHOOL AND TOWARD EMPLOYMENT.”

DR. KATHARINE FRASE
CTO EDUCATION, IBM

IBM



Educación-MECD

← → ↻

gooru Libraries ▾ My Content My Classes Log in

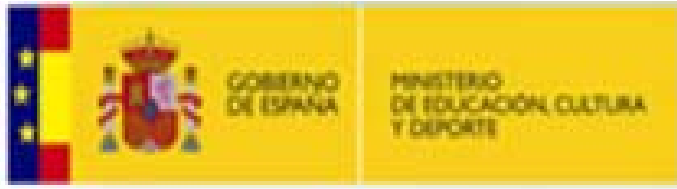
Education

Hi there! What are you looking for?

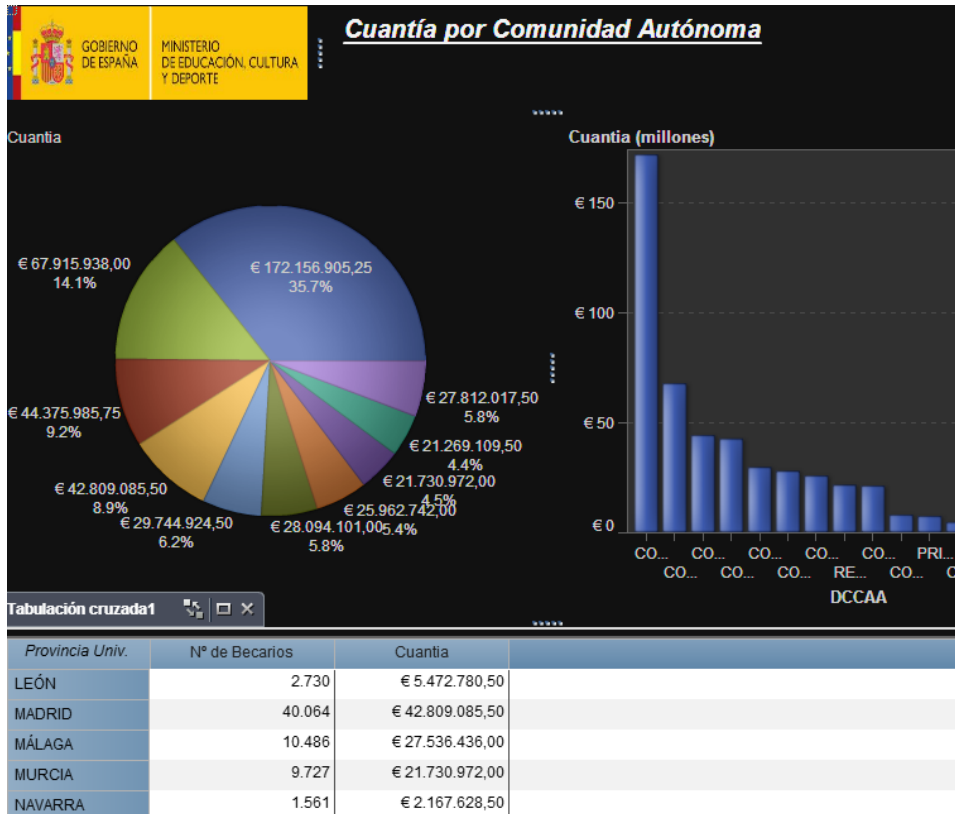
I am looking for learning materials in or

Already have an account? [Login here.](#)

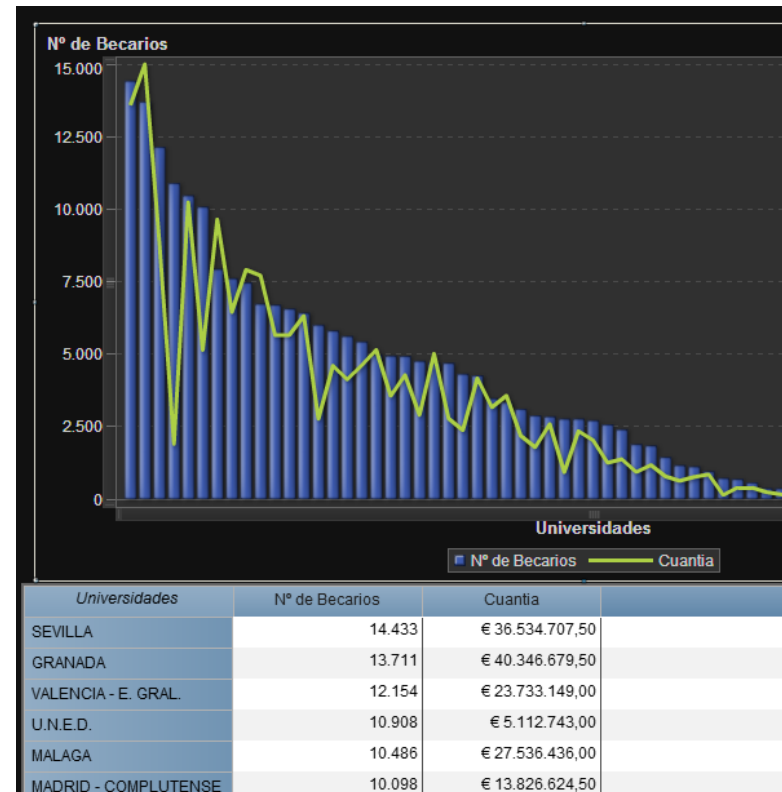
For Teachers	For Students	For Districts	For Partners
Discover standards-aligned content, customize content, and track student progress through data analytics.	Explore interests, build, and monitor progress through learning materials.	Collaborate with Gooru to unleash personalized learning and share a district-vetted curriculum.	Collaborate with mission-aligned partners to increase our collective impact on the education ecosystem.
<input type="button" value="See Stories"/>	<input type="text" value="Enter Class Code"/> <input type="button" value="Enter"/>	<input type="button" value="See Our Impact"/>	<input type="button" value="Learn More"/>

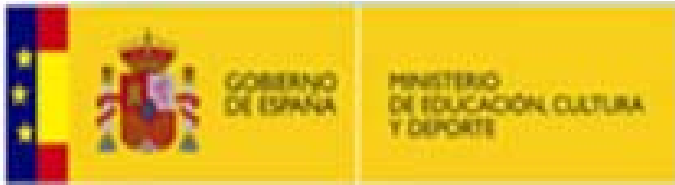


Educación-MECD



Becas y Ayudas





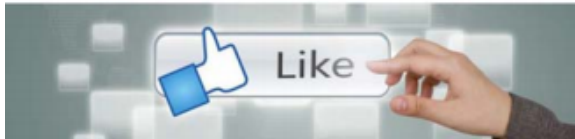
Educación-MECD

FP-Online

Plataforma e-learning



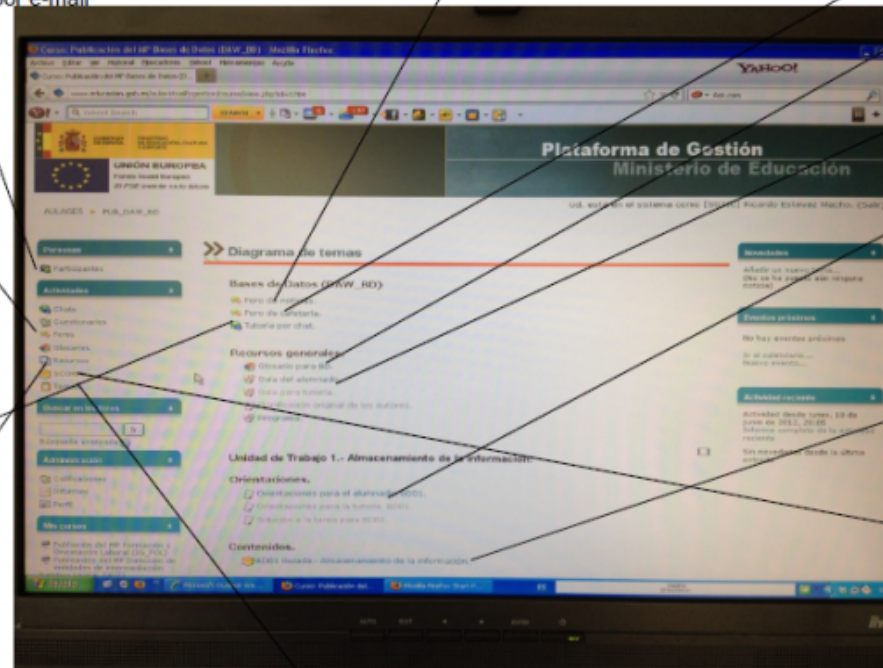
Miembros del curso
Puedo contactar por e-mail



Visualizar
Todos los foros

Preguntas
Directas al
tutor

Guías existentes
Orientaciones..



Mensajes generales a todos los participantes

Foro para hacer
vida social

Términos de la
Unidad

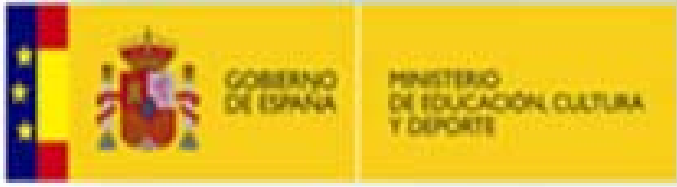
Explicación
Uso herramienta

Sobre la Unidad de
trabajo: Orientación
del tema, Índice,
Objetivo, tiempo
necesario, consejos,
recomendaciones..

La lección en sí

Acceso paquete
Scorm de todas
Las unidades

Tareas para todas las Unidades



Cultura

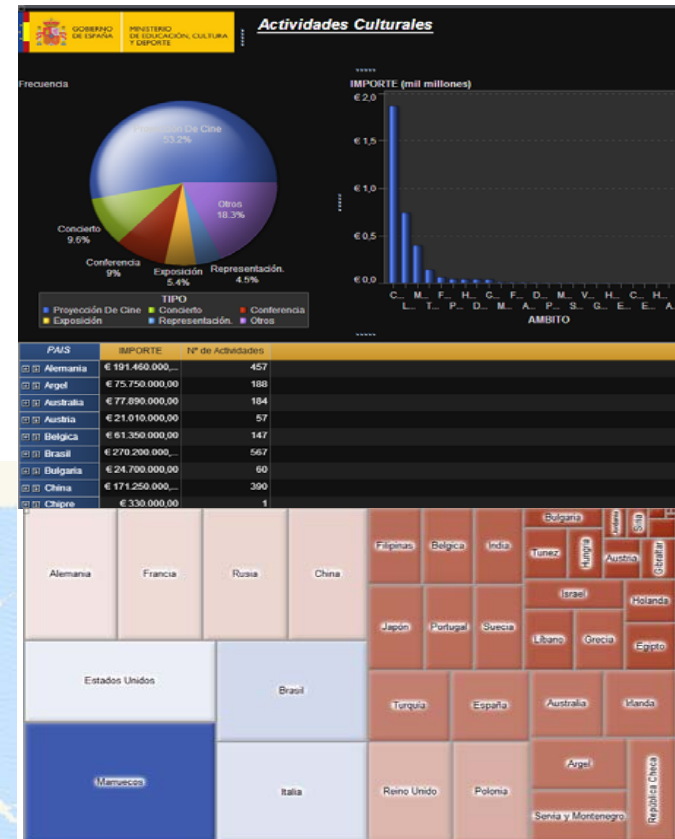
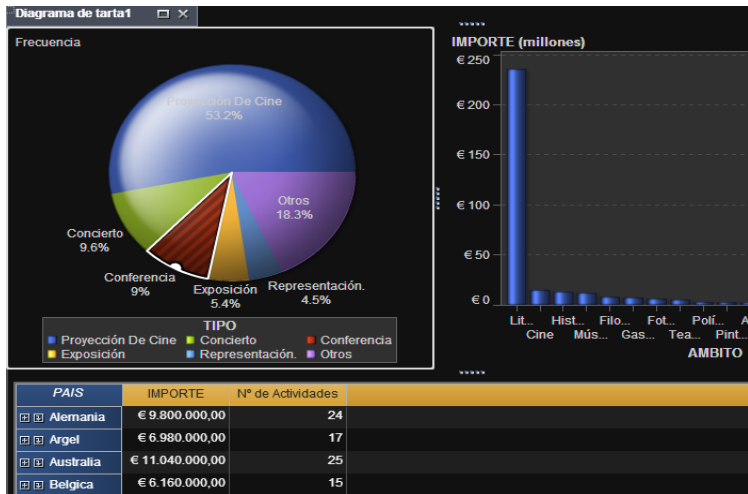


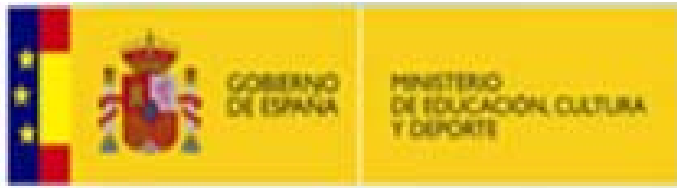
Graz (Austria)
Paisaje Móvil

Big Data



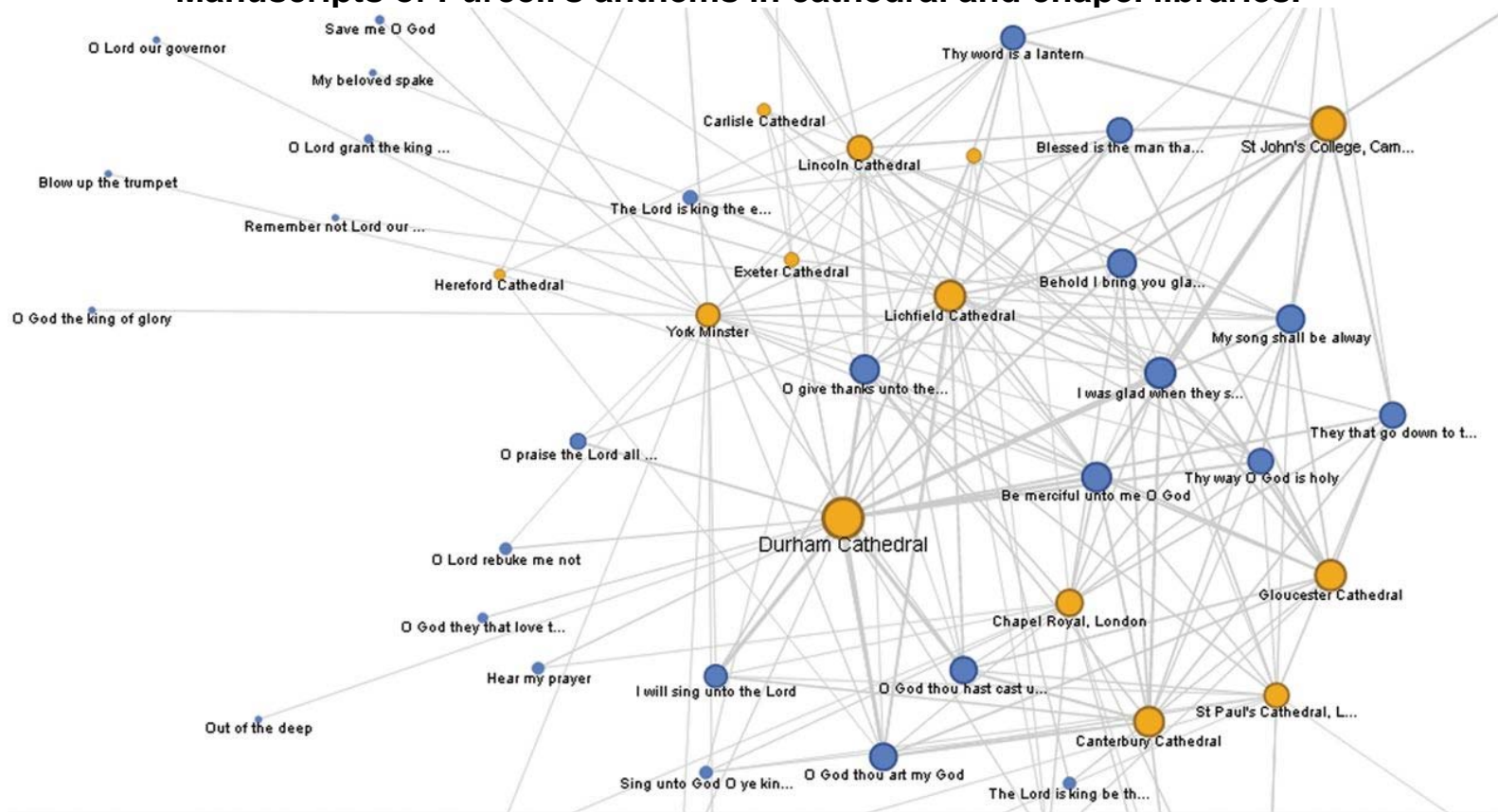
Cultura-MECD





Cultura

Manuscripts of Purcell's anthems in cathedral and chapel libraries.



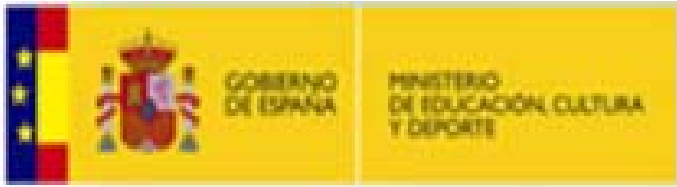
Writing a Big Data history of music

Stephen Rose et al. *Early Music* 2015;43:649-660

© The Author 2015. Published by Oxford University Press.

LONDON

EARLY MUSIC



Cultura-MECD

BIBLIOTECA DIGITAL HISPÁNICA Bienvenidos - Benvinguts - Ongi etorri - Benvidos - Welcome - Bienvenue
 BIBLIOTECA NACIONAL DE ESPAÑA Contacto | Ayuda | Alertas

Libros, manuscritos, partituras, fotografías... Todos los campos **BUSCAR**

Busque en el texto de los documentos [Búsqueda avanzada >](#)

Inicio Descubrir colecciones Acerca de la digitalización

Bienvenido al portal libre y gratuito de documentos digitalizados de la Biblioteca

Con el apoyo de *Telefónica*

Nuevas incorporaciones DESCRIPCIÓN DEL UNIVERSO (1613)

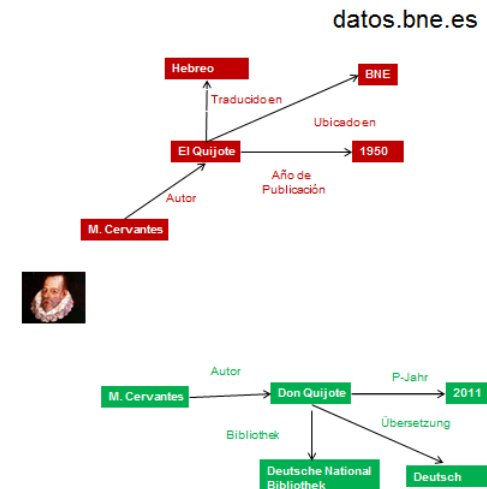
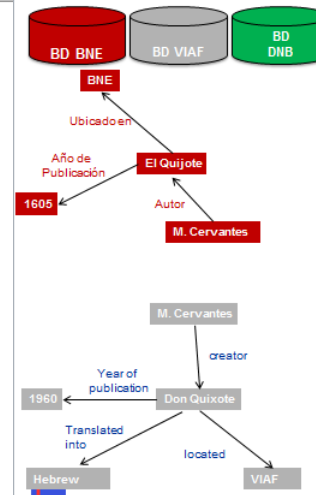


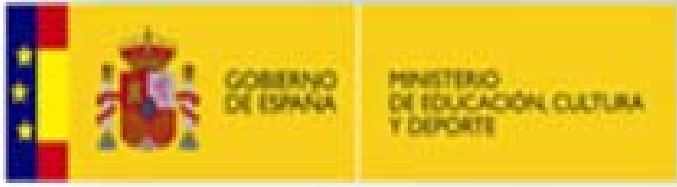
BNE:
 Recolección masiva del dominio .es y selectiva por acontecimientos

Linked
 Open
 Data

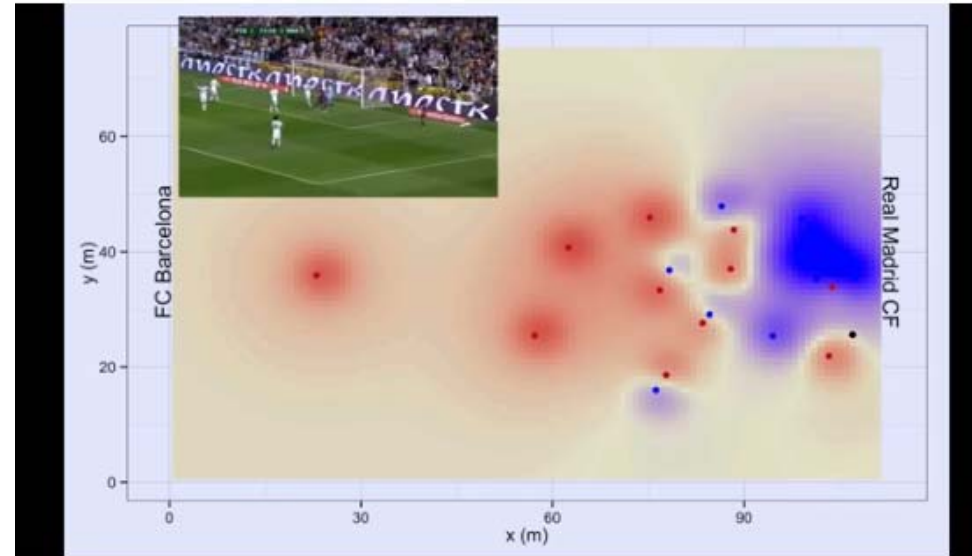
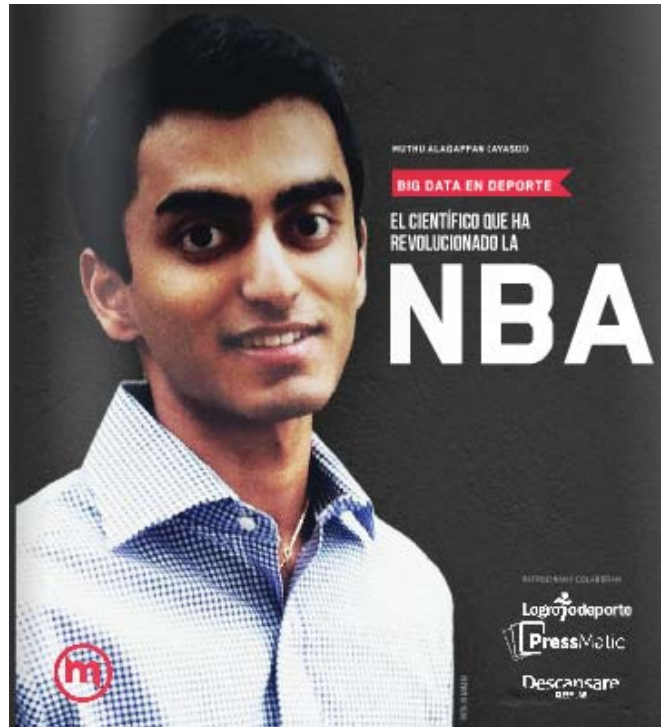
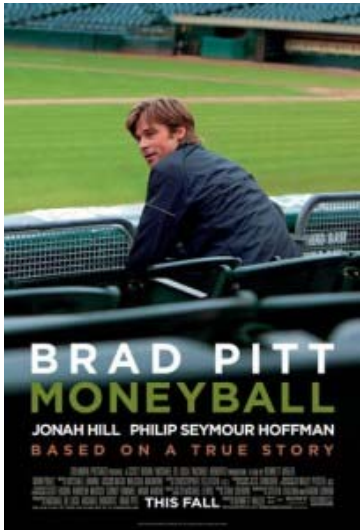
Modelos
 RDF(S)

RISP –BIG- MECD

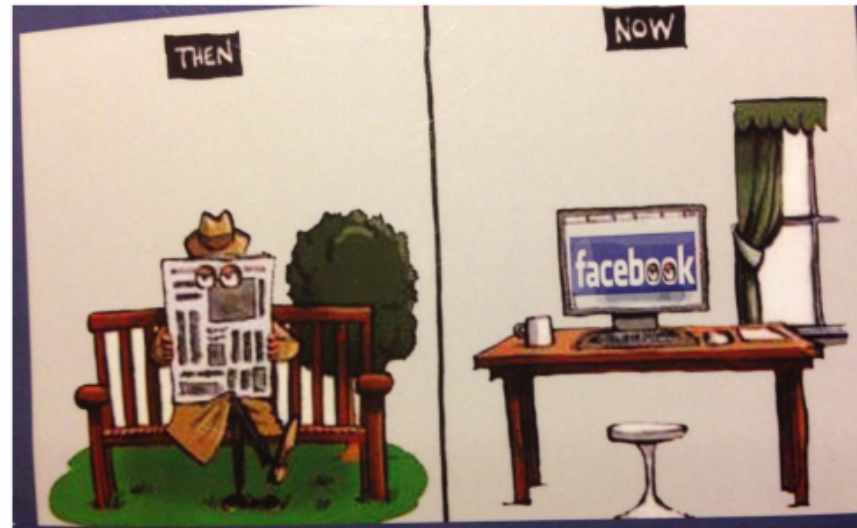


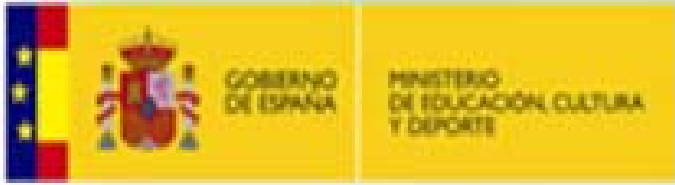


Deporte

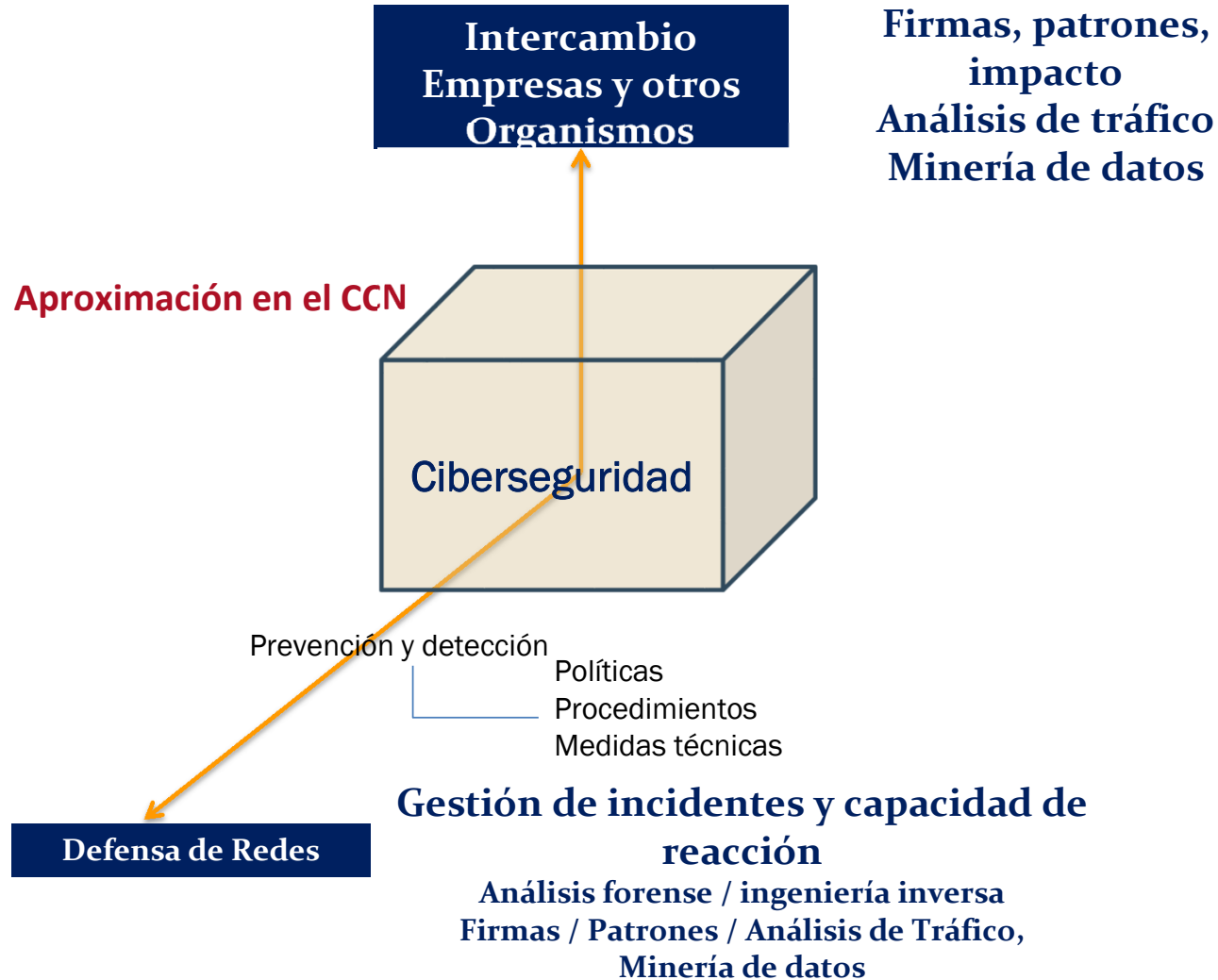


Ciberseguridad





MECD_CCN-AGE



Ciberseguridad Ejemplo de Uso

BLUE COAT

Inteligencia de la amenaza

DeepSee® Applications
Reports, GeoIP, Reconstruction,
Packet Analysis



Partner Integration



Reputation, Threat Feeds,
and File Analysis



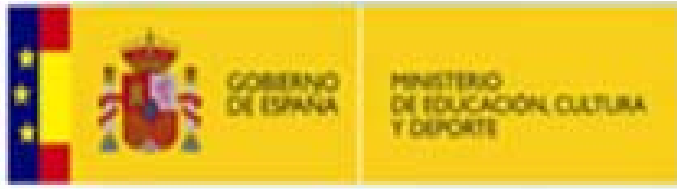
Big Data Analytics



DeepSee Big Data Security Analytics

Sensores





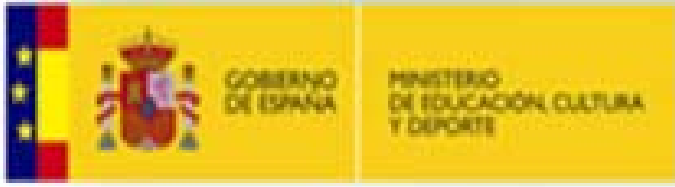
MECD

LEY DE TRANSPARENCIA

➤ Servidor de Informes

Aplicación	Informe	Usuarios	Programado	Publicado
Fondo Social Europeo	Centros Sin duplicados	carmen.cabanillas, juan.castador, mousana.sagura	No	SI
Fondo Social Europeo	Centros totales	carmen.cabanillas, juan.castador, mousana.sagura	No	SI
Consultas Información OAC	Consulta con parametros 2	alfonso.lazaro, jansorio.alonso	No	No
Consultas Información OAC	Consulta Test con parametro	alfonso.lazaro	No	No
Consultas Información OAC	Consulta test	alfonso.lazaro	No	No
SIDP	Corteo personal por provincia, cuerpo y relación administrativa	paloma.iglesias, rosa.herneros, taniocio.rodriguez	No	SI
Teles	Crucé Datos Teles Leídas con Becarios FDI (NSF ok y Fecha Lectura nula en TP)	carlos.hernero, luis.cholbi	No	SI
Teles	Crucé Datos Teles Leídas con Becarios FDI (NSF OK y Fecha Lectura distinta en ambos esquemas)	carlos.hernero, luis.cholbi	No	SI
Titulos	Envío a universidades	pricente.mas, raquel.gomez	No	No
SIDP	Estadística de Personal Docente por Cuerpos	paloma.iglesias, rosa.herneros, taniocio.rodriguez	No	SI
Titulos	Estadística desde 2000	pricente.mas, margarita.lezcano, maria.mojascano	No	SI
Titulos	Estadística desde 2000_tipo_titulo	argelia.lobato, pricente.mas, margarita.lezcano, maria.mojascano	No	SI
Consultas Información OAC	estadísticas en historico en 2014	alfonso.lazaro, carlos.hernero, guillermo.malonado, jeferson.alfonso, jesus.ibañez, pricente.mas,	No	No





MECD

LEY DE TRANSPARENCIA

➤ Servidor de Informes

Gestor de Informes

Inicio > Administración > Informes

Criterios de búsqueda X

Aplicación: Seleccione... Nombre: Usuario:

Informes

Aplicación	Informe	Usuarios	Programado	Publicado
		carmen.cabanillas	No	Si
			No	Si
		u.alfonso	No	No
		Garmen Ca	No	No
			No	Si
		ibi	No	Si
		omaza	No	No
			No	Si
		te.mas	No	Si
		henery	No	No
		mas		

Centros Sin duplicados

Informe | Permisos | Sentencia SQL | **Diseño**

Estado: Publicado

Publicado Programado

Diseño del Informe

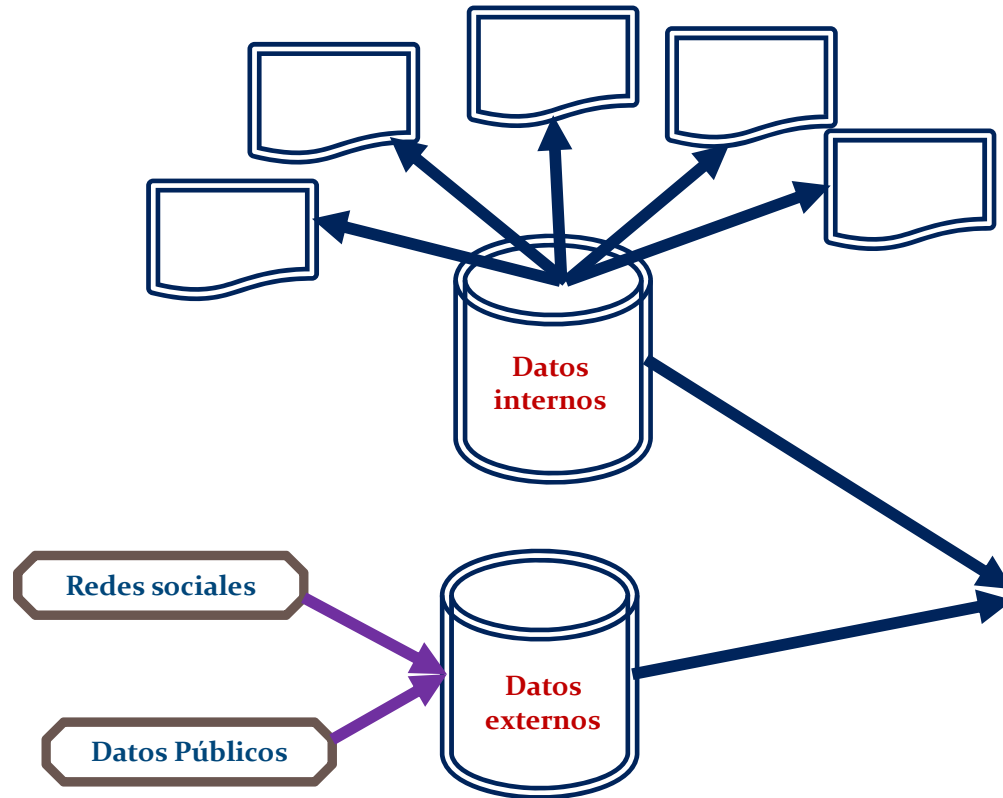
Tipo de gráfico: SOLO DATOS

Campo	Formato en Excel	Posición en Excel
CCAA	Alfanumérico	
Provincia	Alfanumérico	
Localidad	Alfanumérico	
Código Centro	Texto	
Nombre Centro	Alfanumérico	
Tipo de centro	Alfanumérico	
Zona rural	Alfanumérico	
Director	Alfanumérico	
Teléfono	Alfanumérico	
Email	Alfanumérico	
Domicilio	Alfanumérico	
Fecha Inicio	Alfanumérico	
Fecha Fin	Alfanumérico	

Posición de los campos sobre la hoja Excel

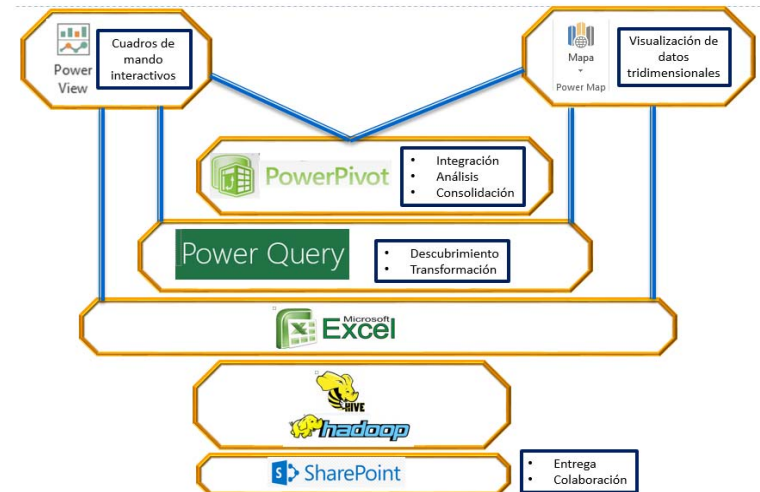
	A	B	C
1	FILTROS	(Todas)	
2			
3	COLUMNAS		
4			
5	FILAS		
6			
7			
8			DATOS
9			

Ministerio de Educación, Cultura y Deporte



FUNDACIÓN Big Data

Power BI



The screenshot shows the Microsoft Excel interface with the 'DAT' ribbon selected. The 'De origen' dropdown menu is open, displaying various data source options. The 'De Facebook' option is highlighted with a red circle and a red arrow. To the right of the menu, a list of fields is visible, including 'first_name', 'gender', 'last_name', 'link', 'location', 'locale', 'name', 'timezone', 'updated_time', 'username', 'verified', 'work', and 'connections'.

first_name
gender
last_name
link
location
locale
name
timezone
updated_time
username
verified
work
connections

PREGUNTAS



Muchas gracias

carmen.cabanillas@mecd.es