

CEiNE.

Centro de Investigación en
Inteligencia de Negocios



Big Data “Derribando Mitos”

Lautaro Cuadra - Octubre 2013



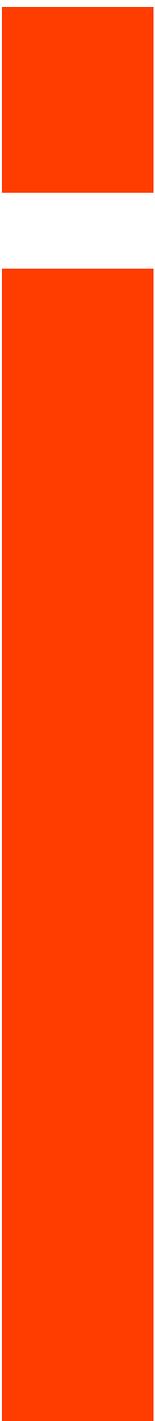
INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

Telefonica



fcfm

FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE



La era del “Big Data”

Prometían que sería grandioso...

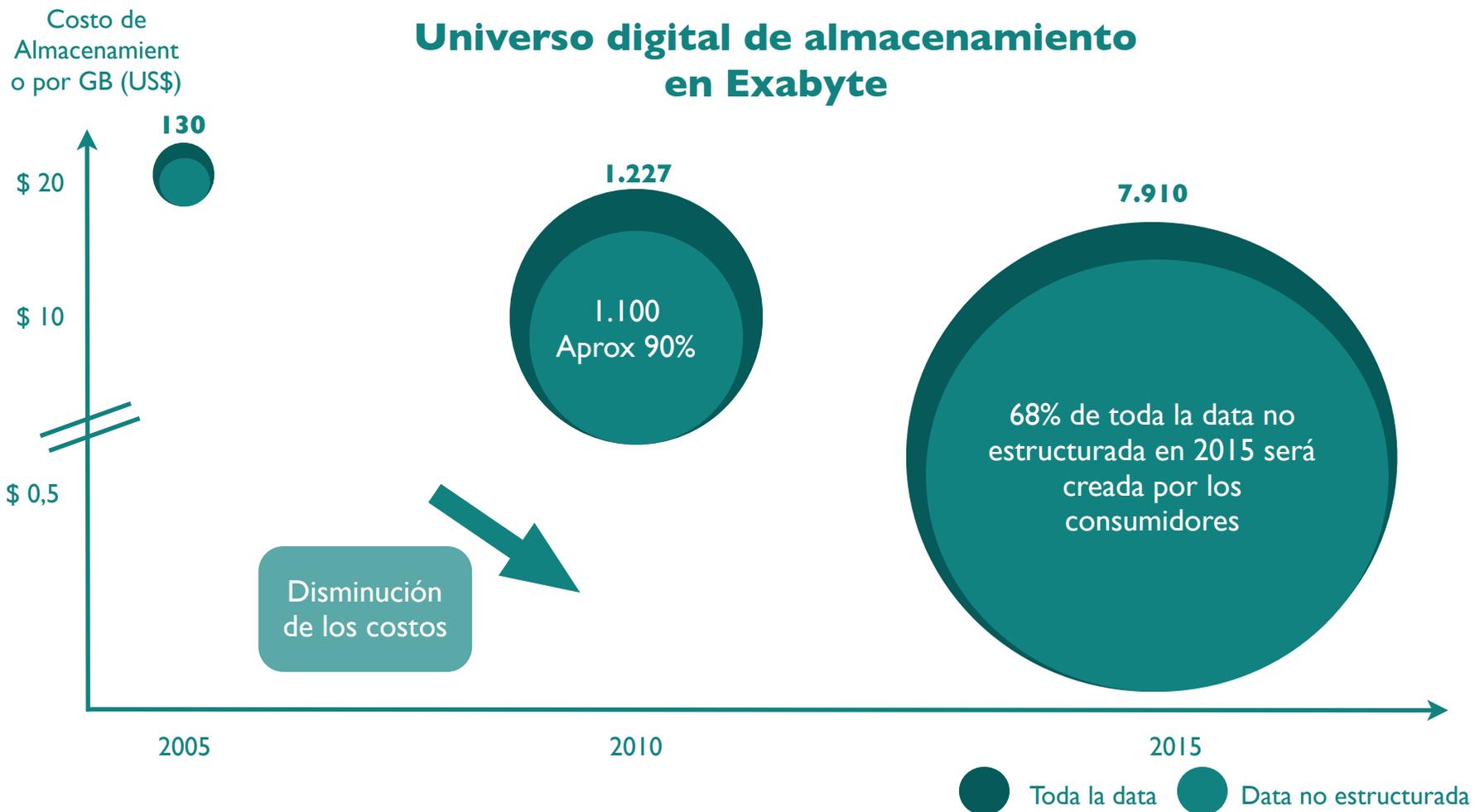
“La era del Big Data”

“...Una revolución en cómo la sociedad utiliza la **información** está irrumpiendo el mundo de los **negocios**, en la **política** y en la **vida cotidiana...**”

La era del big data “ Kenneth Cukier CNBC– 03/12

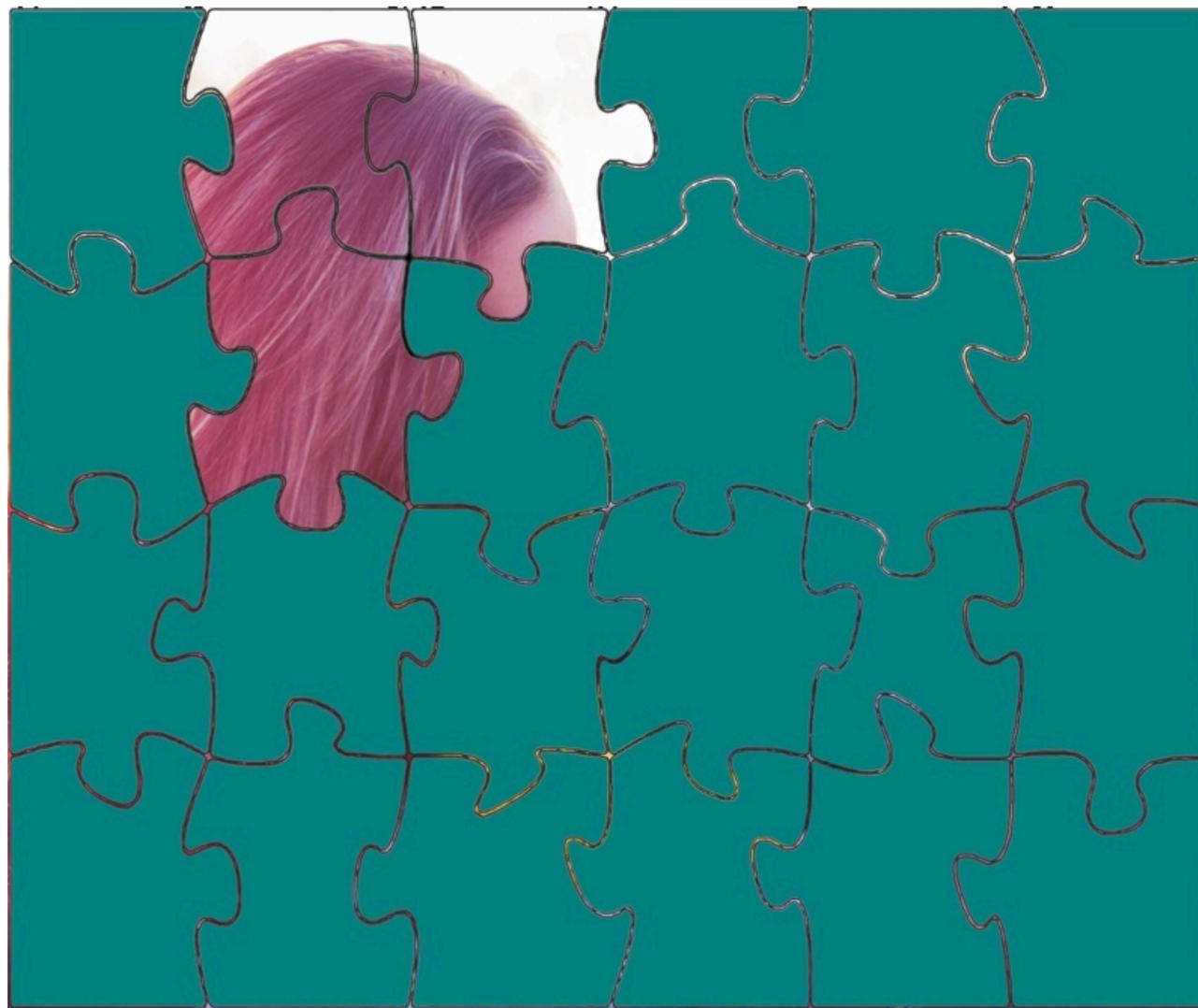


La avalancha de la información...



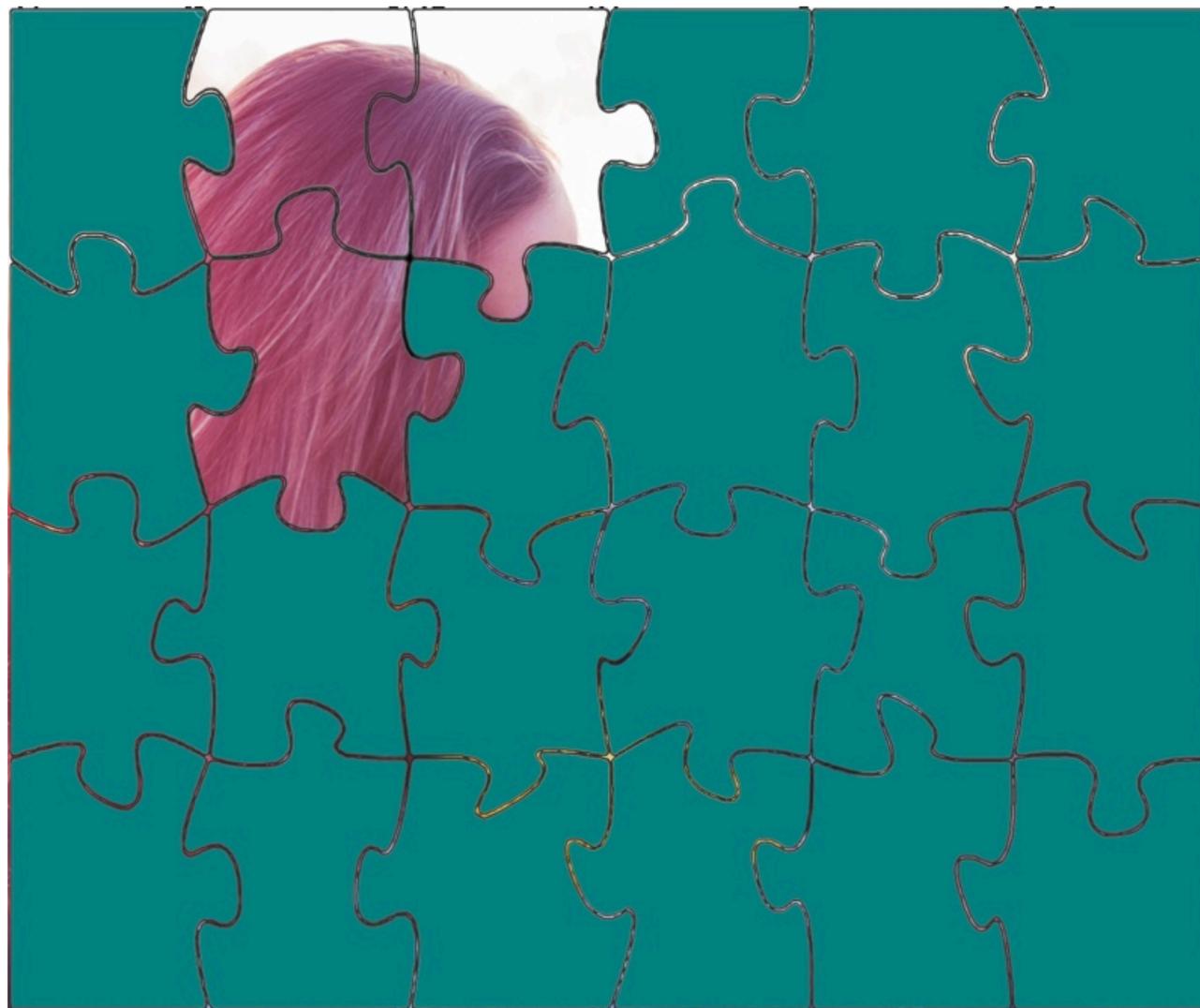
1 Exabyte = 1.000.000.000 Gigabyte > 5 Eb = Todas las palabras alguna vez dichas por toda la humanidad

Nuevas fuentes de información



Nuevas fuentes de información

Lo que hoy se
puede hacer...

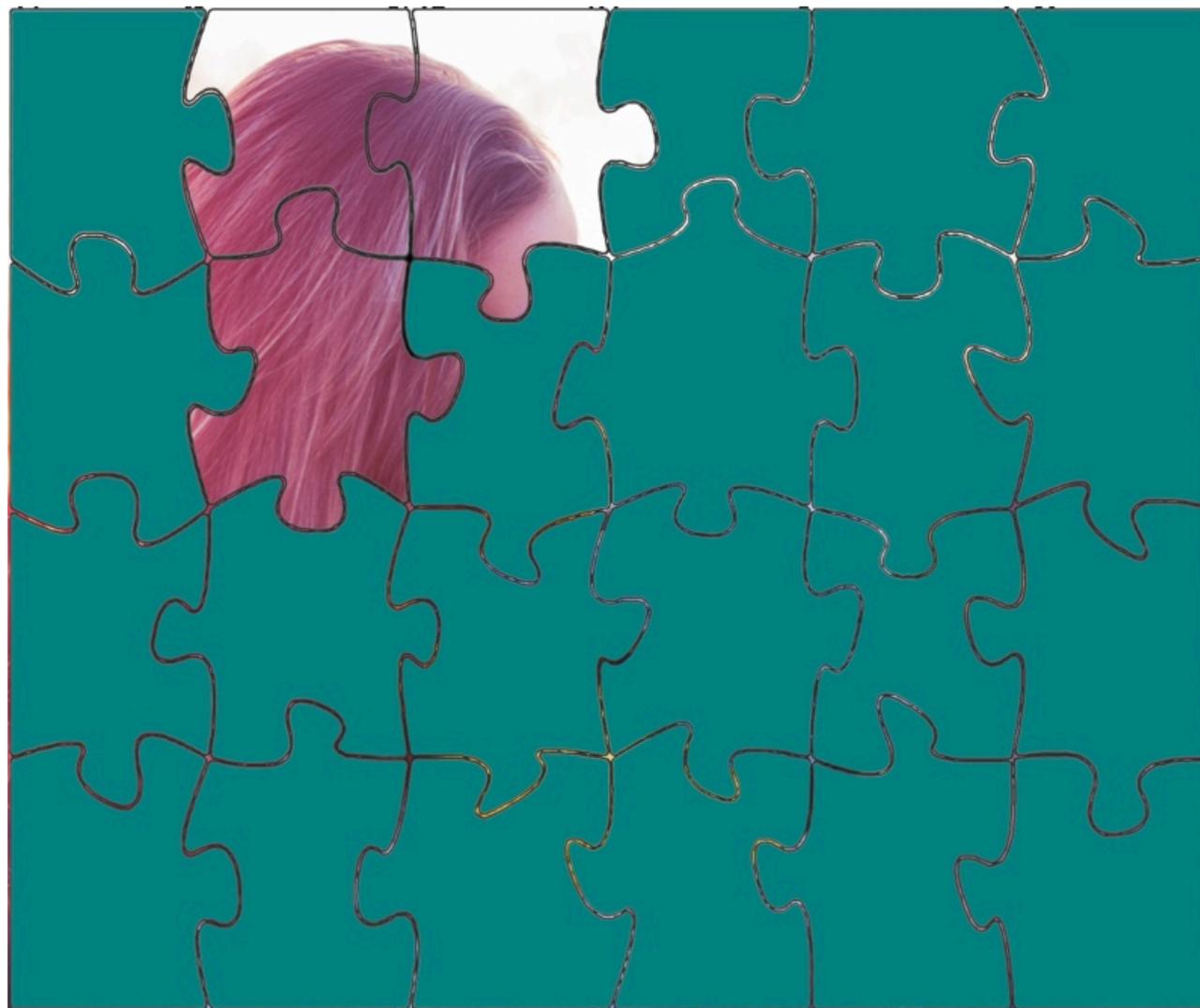


Nuevas fuentes de información

Lo que hoy se
puede hacer...



CRM
Data W.
Demográficos

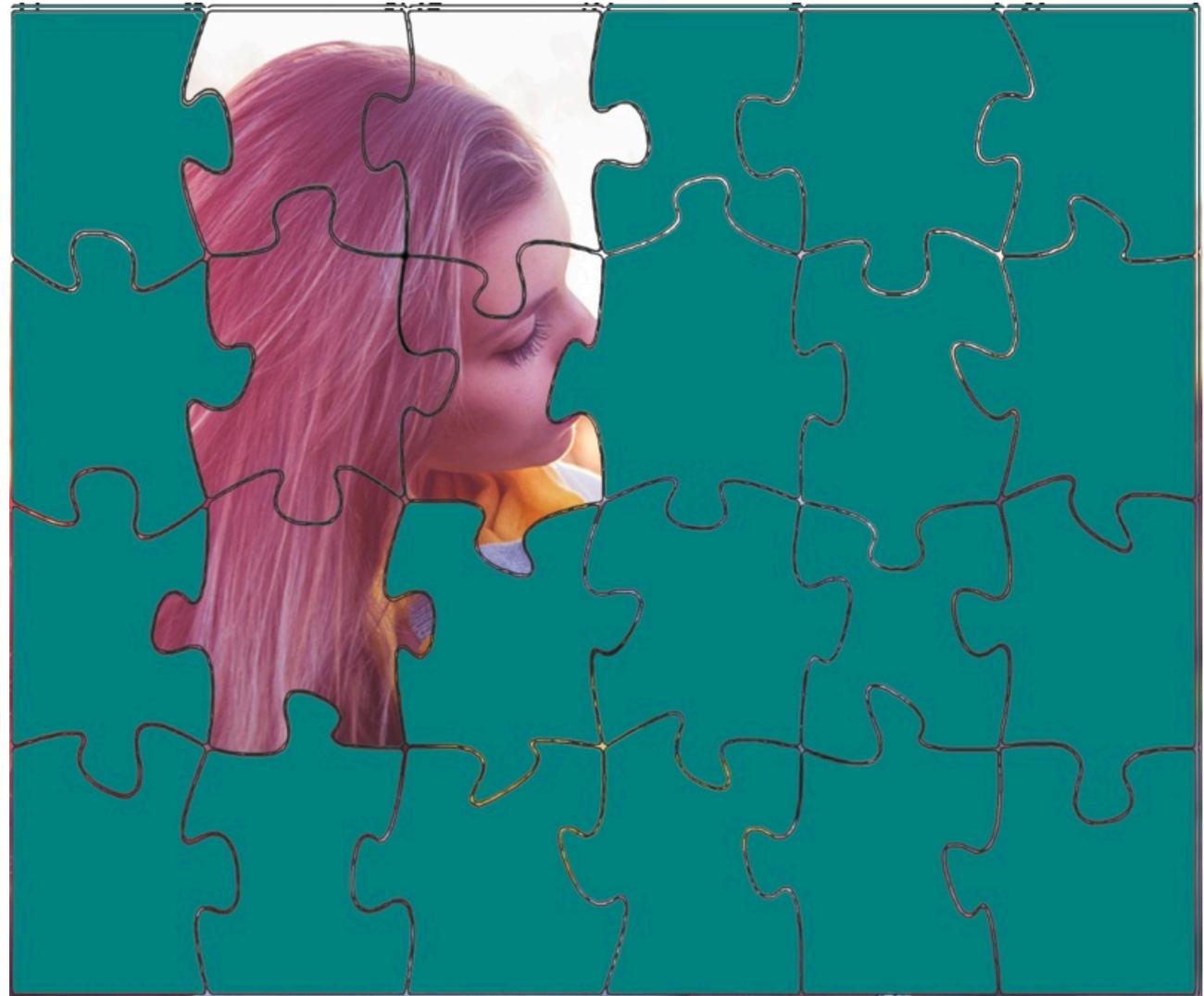


Nuevas fuentes de información

Lo que hoy se
puede hacer...



CRM
Data W.
Demográficos



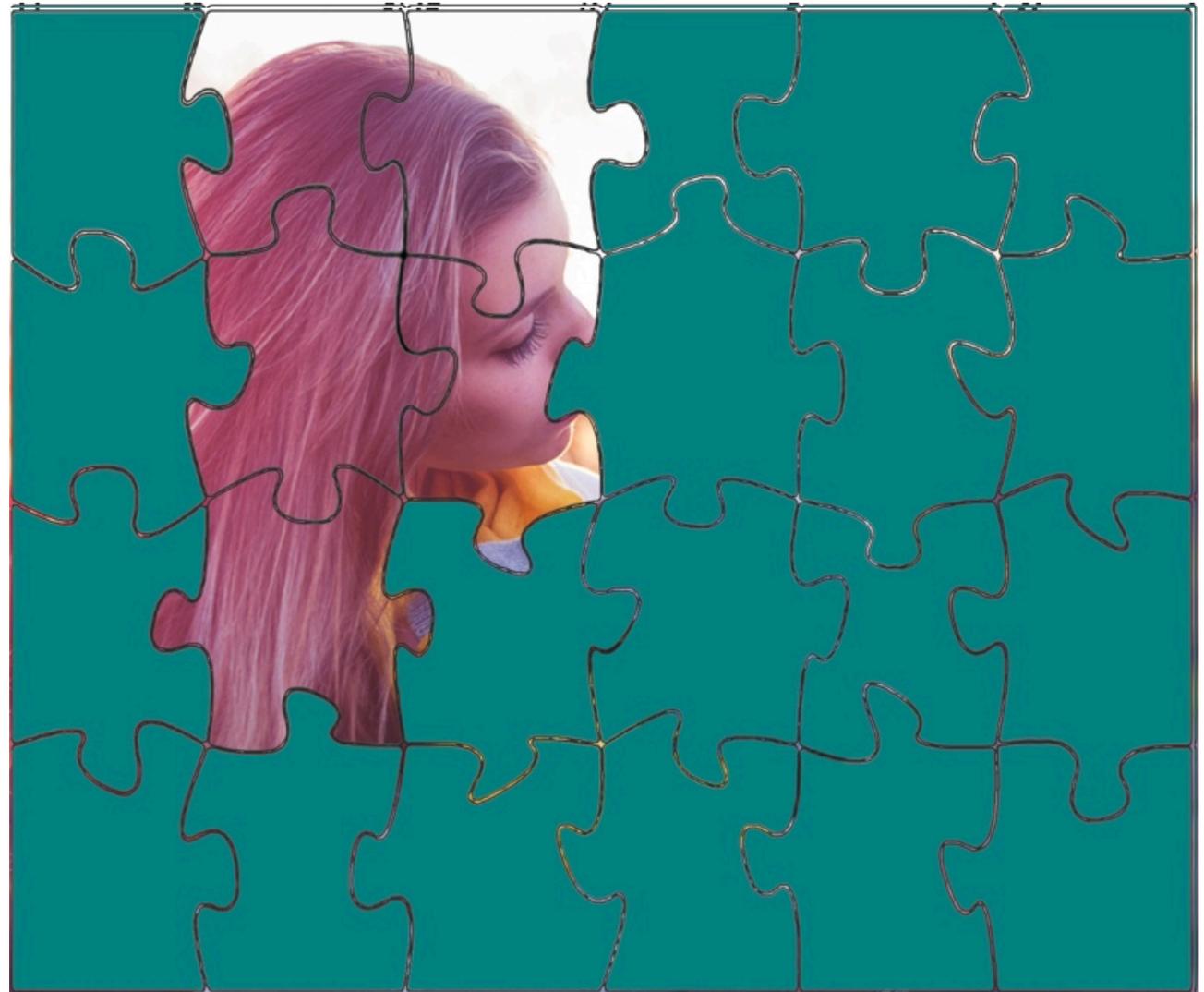
Nuevas fuentes de información

Lo que hoy se puede hacer...



CRM
Data W.
Demográficos

Lo que mañana se podrá hacer...



Nuevas fuentes de información

Lo que hoy se puede hacer...

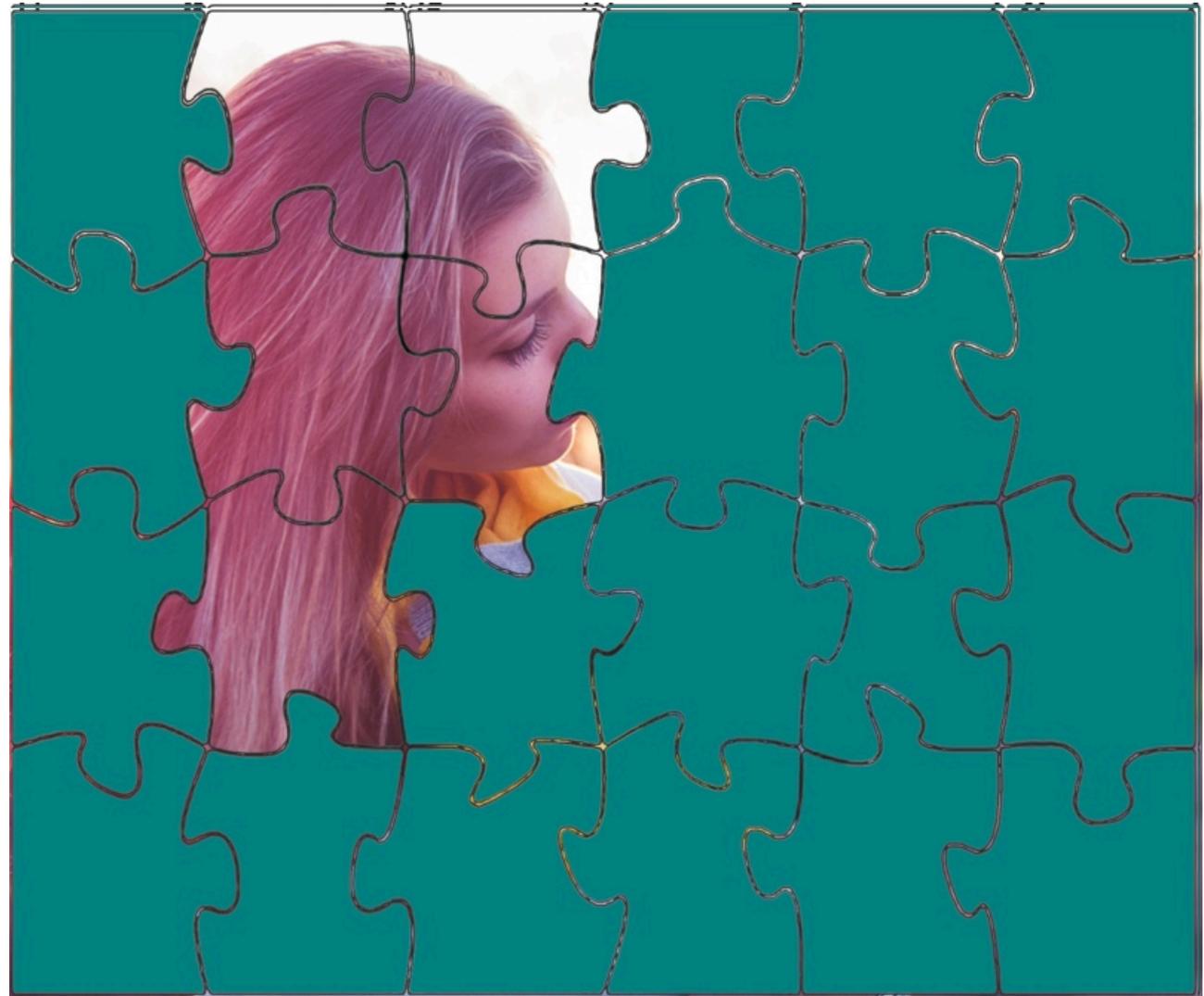


CRM
Data W.
Demográficos

Lo que mañana se podrá hacer...



Ubicación
Geográfica



Nuevas fuentes de información

Lo que hoy se puede hacer...

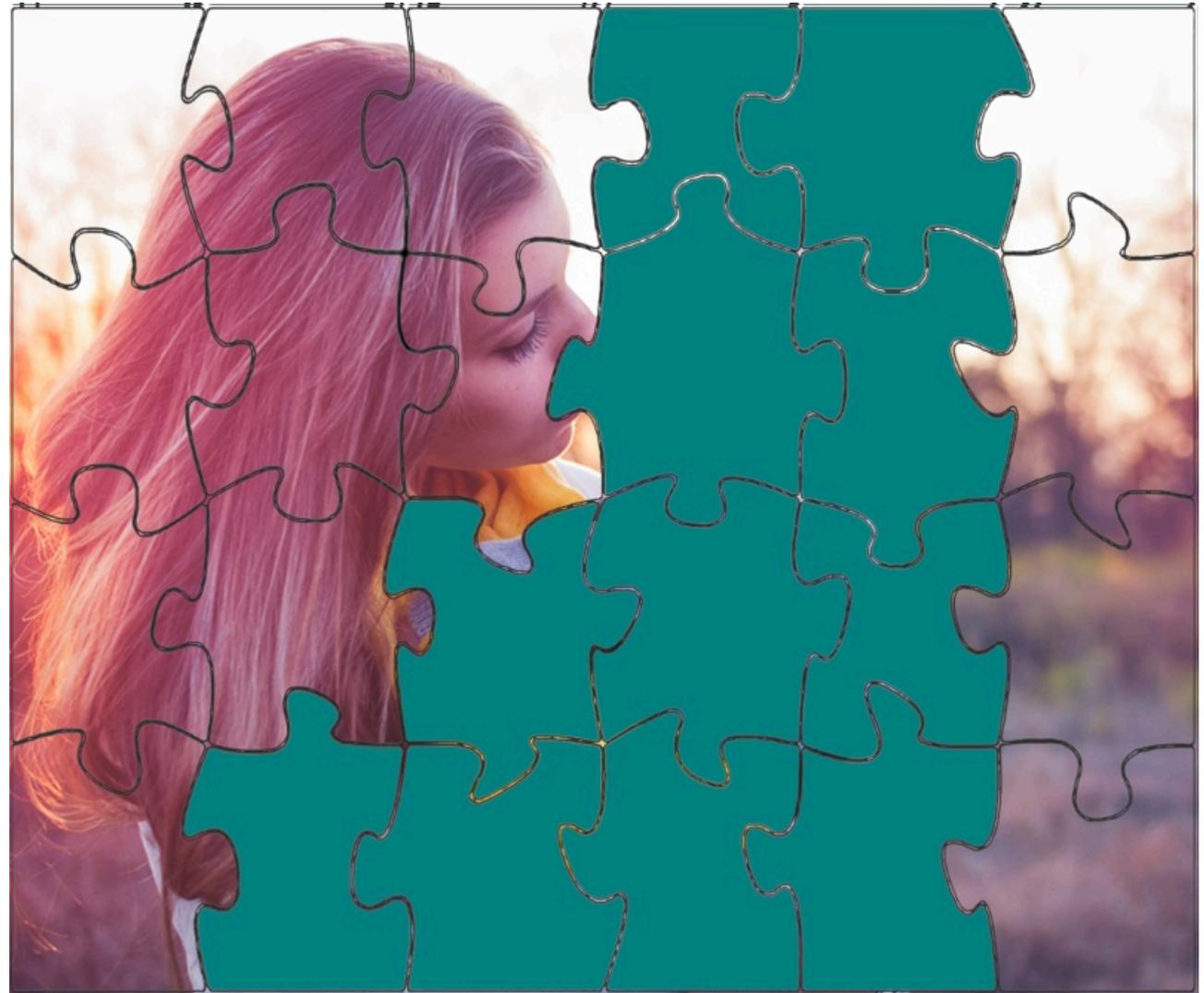


CRM
Data W.
Demográficos

Lo que mañana se podrá hacer...



Ubicación
Geográfica



Nuevas fuentes de información

Lo que hoy se puede hacer...



CRM
Data W.
Demográficos

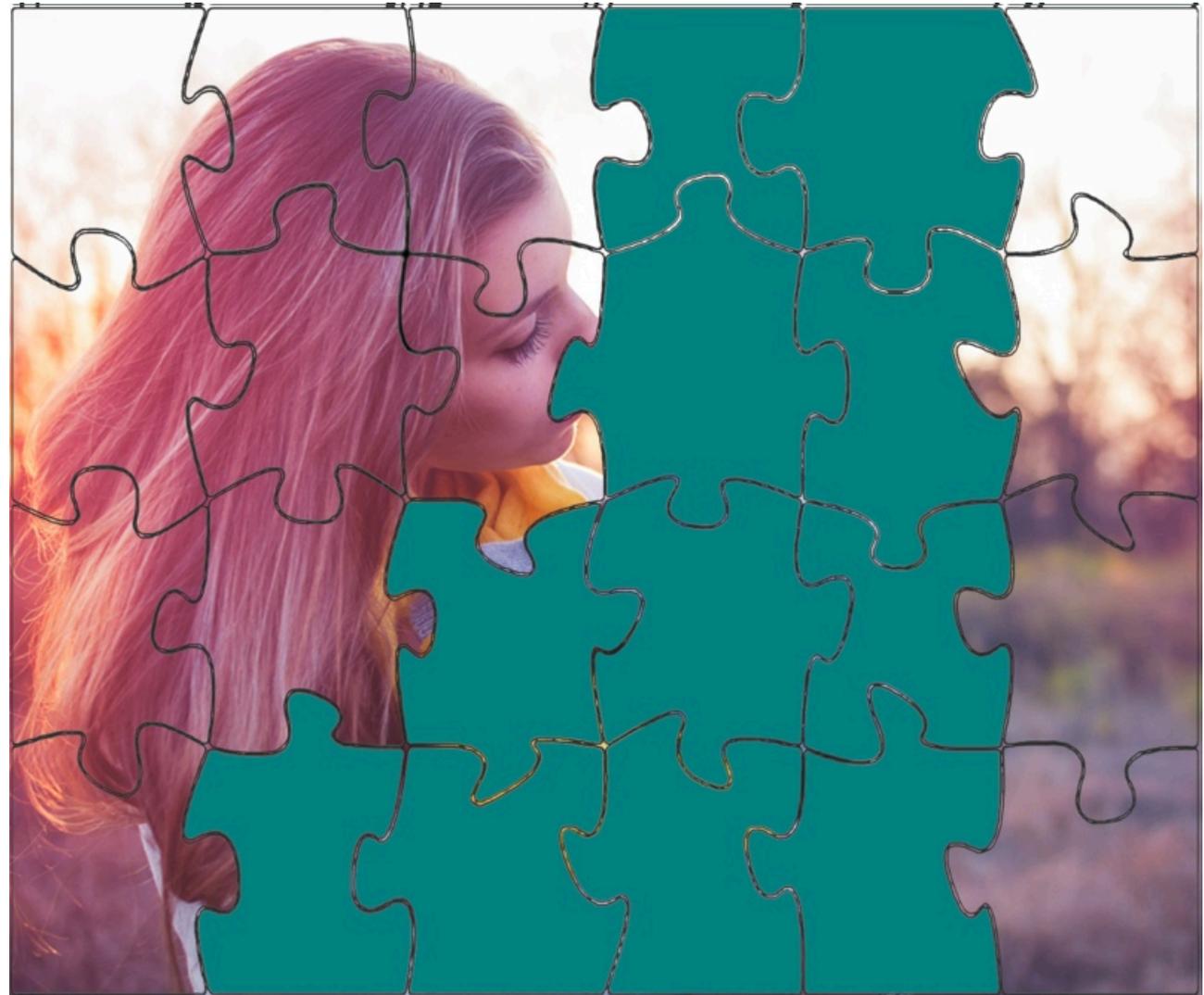
Lo que mañana se podrá hacer...



Ubicación Geográfica



Logs de Compra



Nuevas fuentes de información

Lo que hoy se puede hacer...



CRM
Data W.
Demográficos

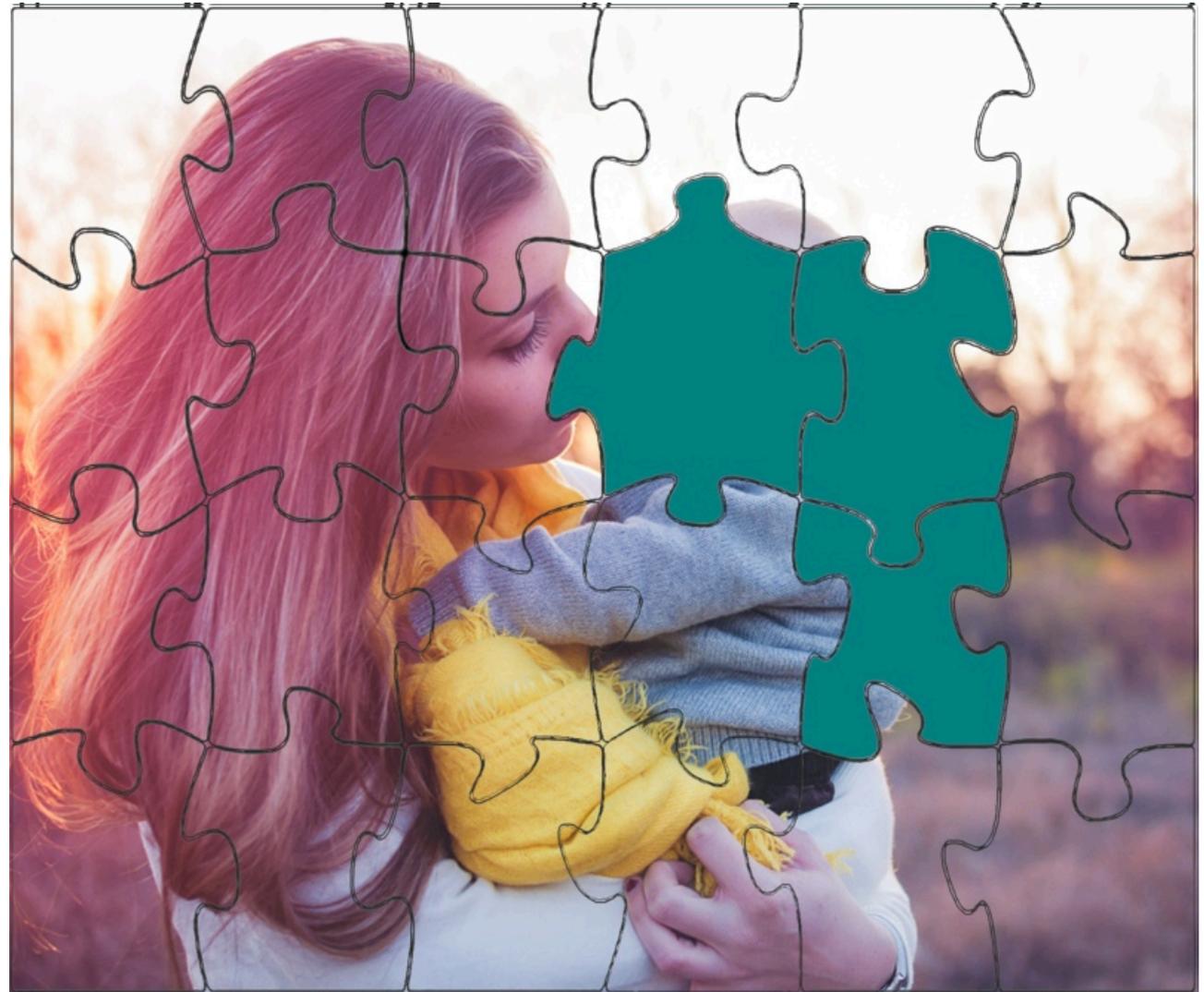
Lo que mañana se podrá hacer...



Ubicación Geográfica



Logs de Compra



Nuevas fuentes de información

Lo que hoy se puede hacer...



CRM
Data W.
Demográficos

Lo que mañana se podrá hacer...



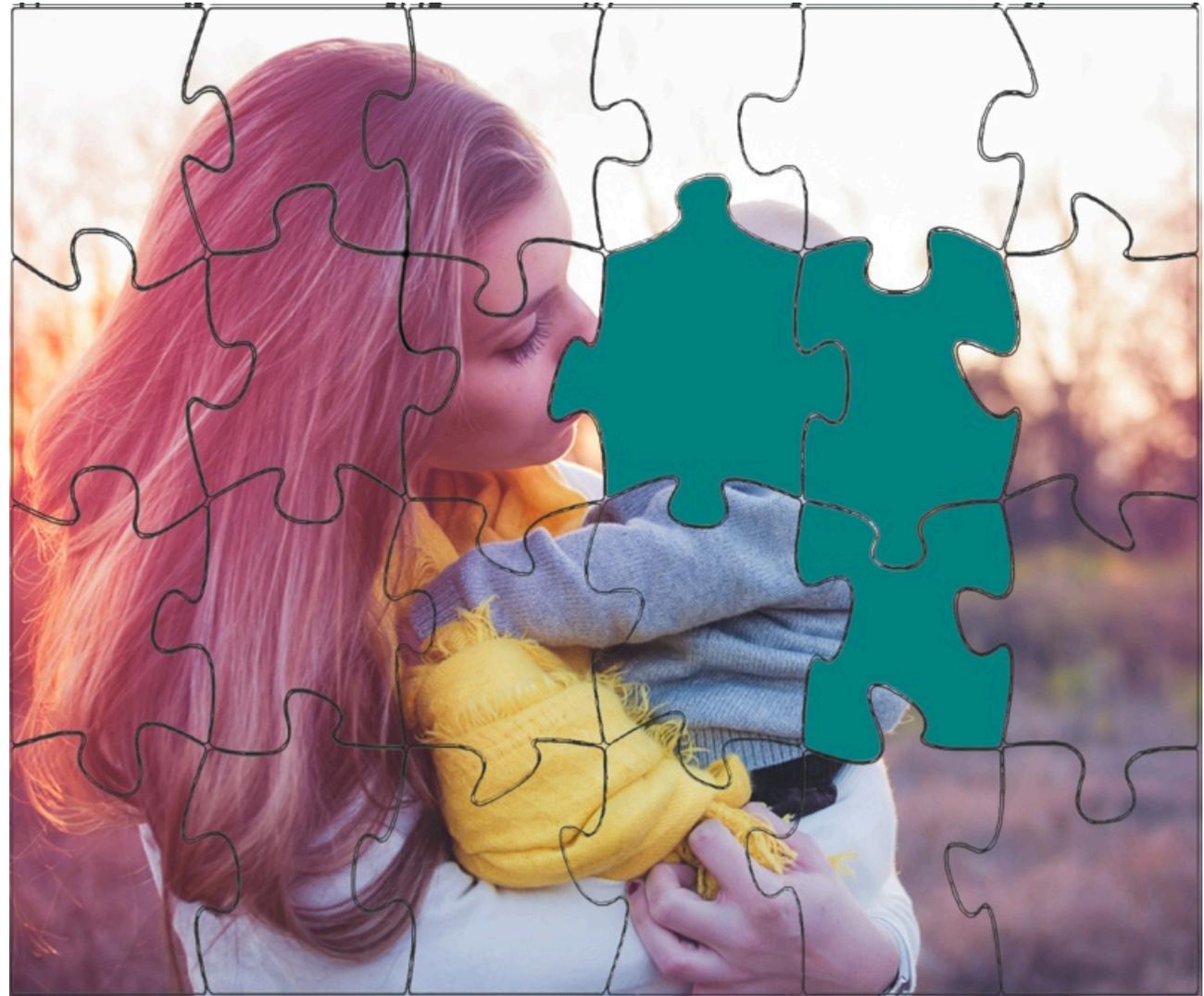
Ubicación Geográfica



Logs de Compra



Redes Sociales



Nuevas fuentes de información

Lo que hoy se puede hacer...



CRM
Data W.
Demográficos

Lo que mañana se podrá hacer...



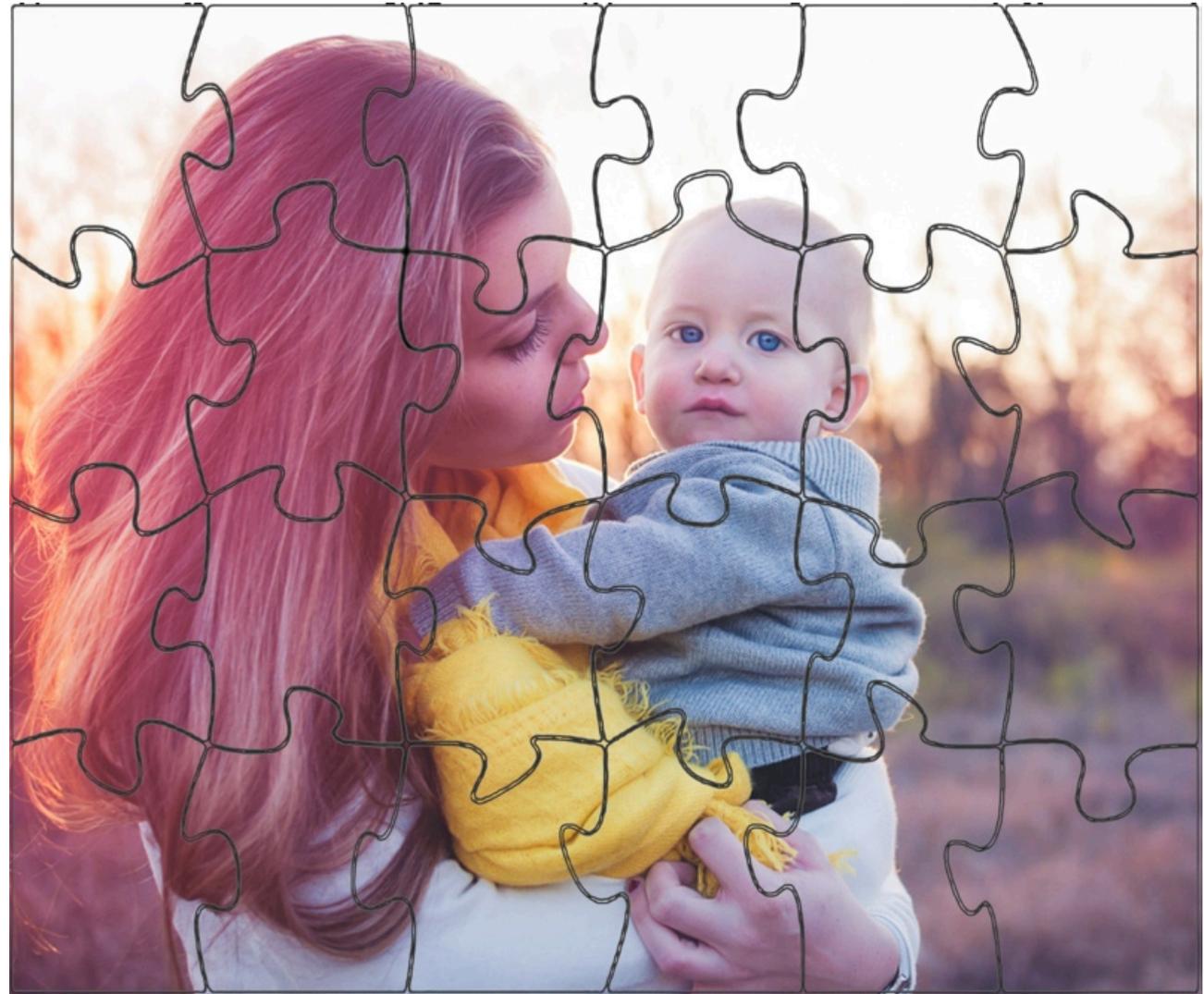
Ubicación Geográfica



Logs de Compra



Redes Sociales



Nuevas fuentes de información

Lo que hoy se puede hacer...



CRM
Data W.
Demográficos

Lo que mañana se podrá hacer...



Ubicación Geográfica



Logs de Compra



Redes Sociales



Una mejor comprensión del Cliente (360°)

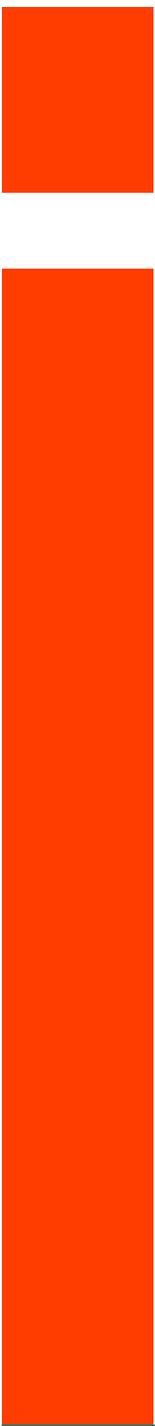
Surgen Miedos...

Para el 2015 se espera que el mercado del Big Data ascienda a US \$ 16,9 billones. McKinsey 2012

60% de aumento en márgenes operativos de las empresas de Retail con Big data McKinsey 2012

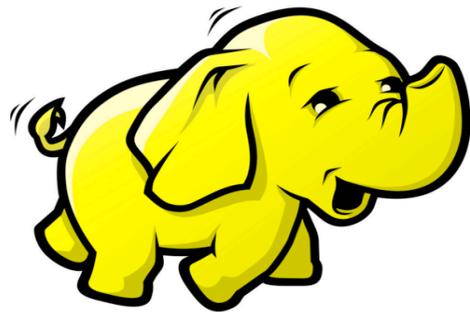
Para el 2015 habrán 4.4 millones de trabajos para Big data, pero solo un tercio podrán ser ocupados Gartner 2012

¿Debe nuestra empresa apostar por Big Data?



Hacia un concepto de “Big Data”

Definición de Big Data...



+

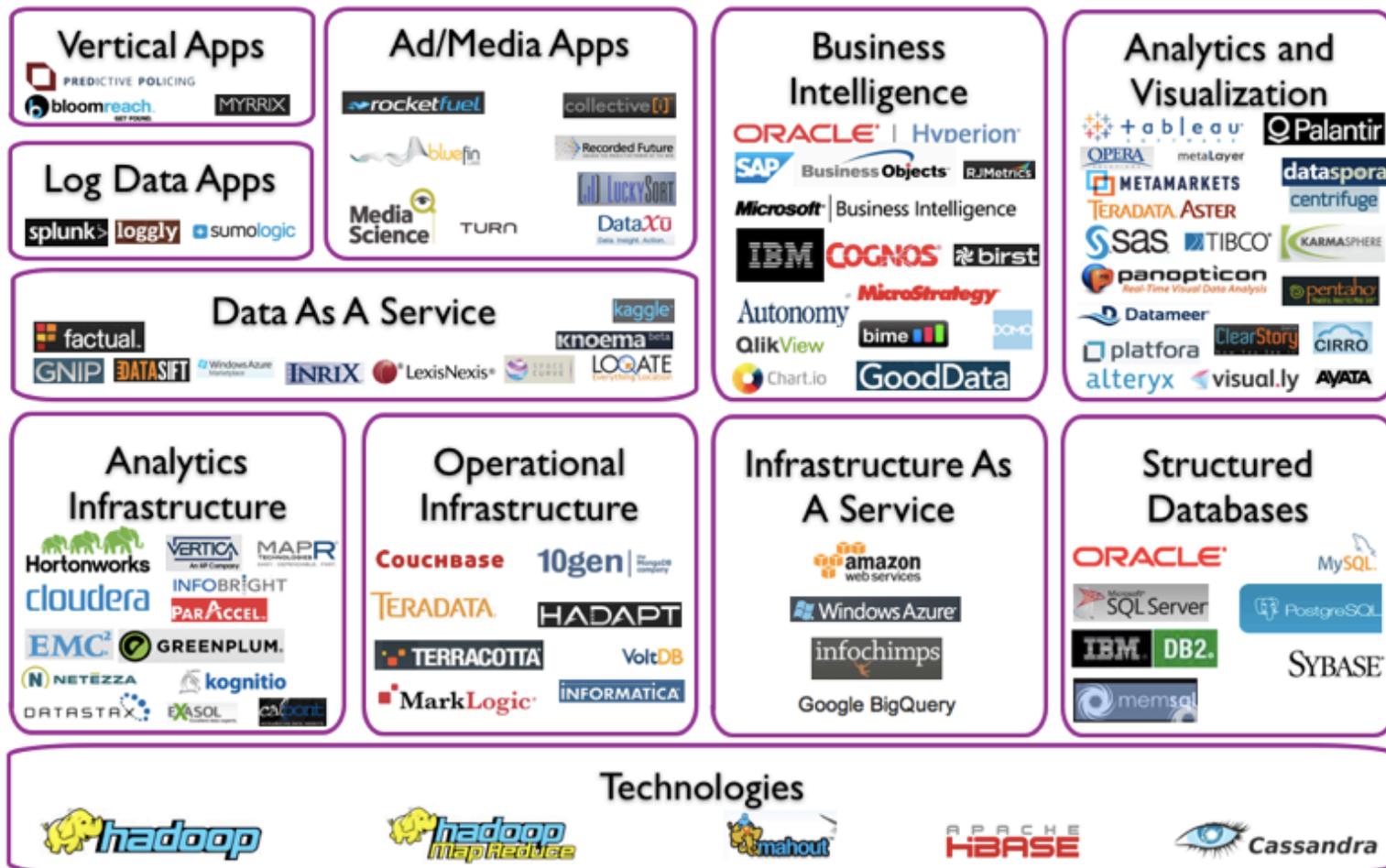


≠



Definición de Big Data...

Big Data Landscape



Copyright © 2012 Dave Feinleib

dave@vcddave.com

blogs.forbes.com/davefeinleib

Definición de Big Data...

Término aplicado a conjuntos de datos **cuyo tamaño supera las capacidades de administración de las herramientas tradicionales** para capturar, administrar y procesar dichos datos en un periodo de **tiempo tolerable**¹

¹ Presentación Plataforma IBM Big Data

Definición de Big Data...

“...Big Data son grandes **volúmenes** de datos, que crecen a gran **velocidad** y de variadas fuentes de datos, que requieren **nuevas formas** de procesamiento para mejorar la **toma de decisiones**, descubriendo información y optimizando procesos ...”

Gartner





El concepto de valor

El concepto de valor



“El Big Data **no puede reemplazar el juicio humano**. Más allá de lo que dice una predicción basada en datos, también las personas usan pistas **sociales, subjetivas e intuitivas** en el proceso de tomar una decisión”

- Caron Carlson

El fin último es generar **valor** a las compañías brindando **información relevante** a los que finalmente tomarán **las decisiones**, es decir, **generar valor a personas**.



Oportunidades con Big Data

¿Es Big Data algo nuevo?



En cada década ha existido un problema de **Big data**.

El concepto en el contexto que hoy conocemos se le atribuye a **John Mashey** de mediados de los 90's

Ciencias como la **astronomía, geología y meteorología** llevan mucho tiempo resolviendo problemáticas de tipo Big Data

Lo Nuevo...

1- Diferentes fuentes de datos y tipos de datos

2- Avances tecnológicos (Hardware - DB)

3- Desarrollos de aplicaciones para el descubrimiento información.

Nuevo paradigma de los datos...

La extensión lógica de la analítica actual en las empresas

Representa el mayor número de oportunidades para las empresas de comprender mejor a sus clientes

Fuentes de Datos

Externas

GPS/ Celulares

Historial de Crédito

Viajes/Pasajes

Censo/Servel

Google+

Facebook

Instagram

Twitter

Linkedin

Blogs

Sensores externos/M2M

Internas

CRM/ERP

Perfil Web

Registro de Compra

RRHH

Inventario

Financieros

Foros Online

Consultas Web

Log de Negocio

Plataformas sociales/ Sharepoint

Sensores Internos/ M2M

Documentos de Texto

Estructurados

No estructurados

Tipo de datos

Lo que las empresas están acostumbradas...

Es la data en la cual experimentar y aprender acerca del valor potencial de lo no estructurado

Nuevo paradigma de los datos...

Representa el mayor número de oportunidades para empresas de entender mejor a sus clientes

La extensión lógica de la analítica actual en las empresas

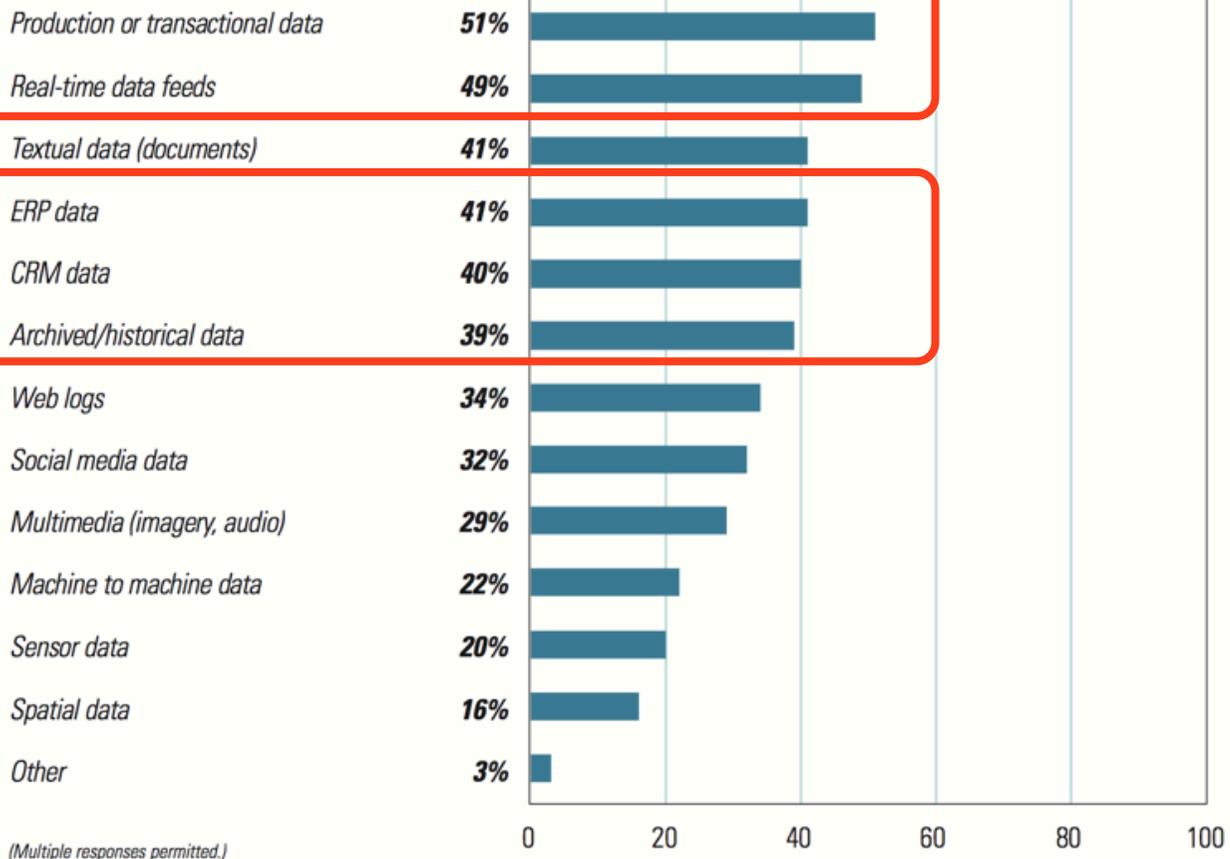
Exter

Fuentes
Datos

Inter

Lo que las empresas están acostumbradas

Figure 6: Types of Data Used in Big Data Initiatives



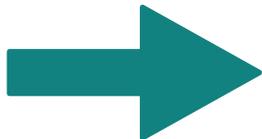
data en la cual experimentar y aprender acerca del valor potencial de lo no estructurado

Tipo de datos

Variables de Geolocalización...

“Location based services”

Servicios que permiten **generar valor** a sus usuarios utilizando como una de sus **variables** de entrada la **posición geográfica** obtenida a través de algún dispositivo.



Outdoor



GPS

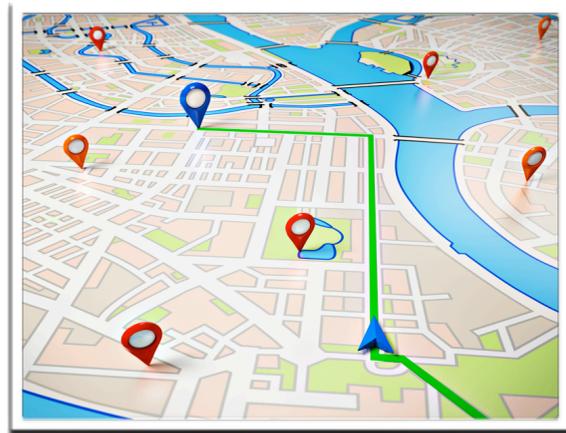
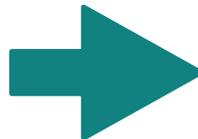


Celular



Wifi

Indoor

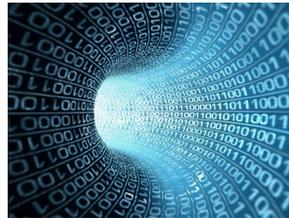


Oportunidades con Geolocalización...



LBS

+



Big Data

=

Llegar al **público adecuado**
En el **tiempo adecuado**
En el **lugar adecuado**

Oportunidades con Geolocalización...

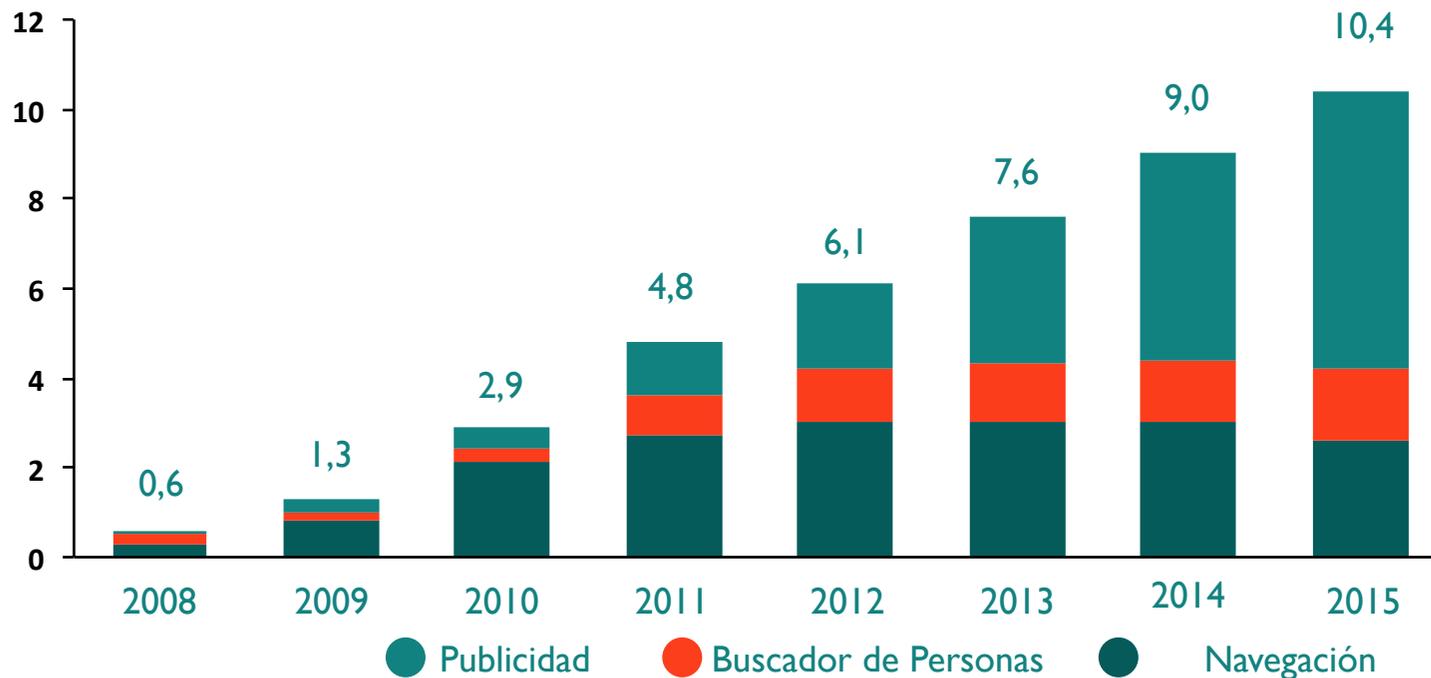


Oportunidades con Geolocalización...

Salud y Bienestar

\$US BN

LBS Forecast 2011-2015



amigos y familia

de bienes y personal

Redes sociales como fuente de datos ...



Redes Sociales

No solo como **canal online** (escuchar al cliente y generar comunicación efectiva)

Contienen información de perfiles, **grupo de amigos, familia y cercanos**, de gustos, **sentimientos**, percepciones, opiniones, entre otros.

Oportunidades en redes sociales ...



**Redes
Sociales**

+

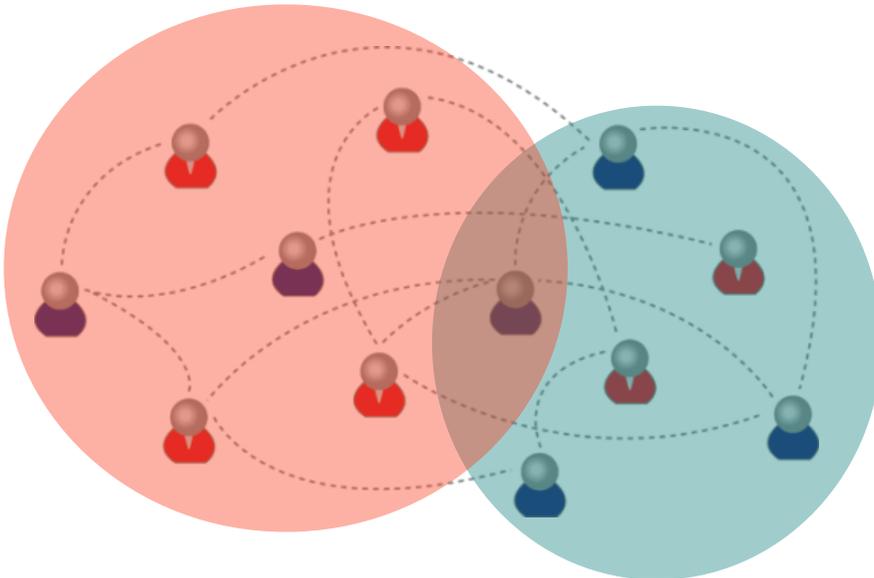


Big Data

=

Análisis de Redes
Sociales (SNA)
aplicada a nuevas
problemáticas

Análisis de Redes Sociales



◆ Buscar miembros clave

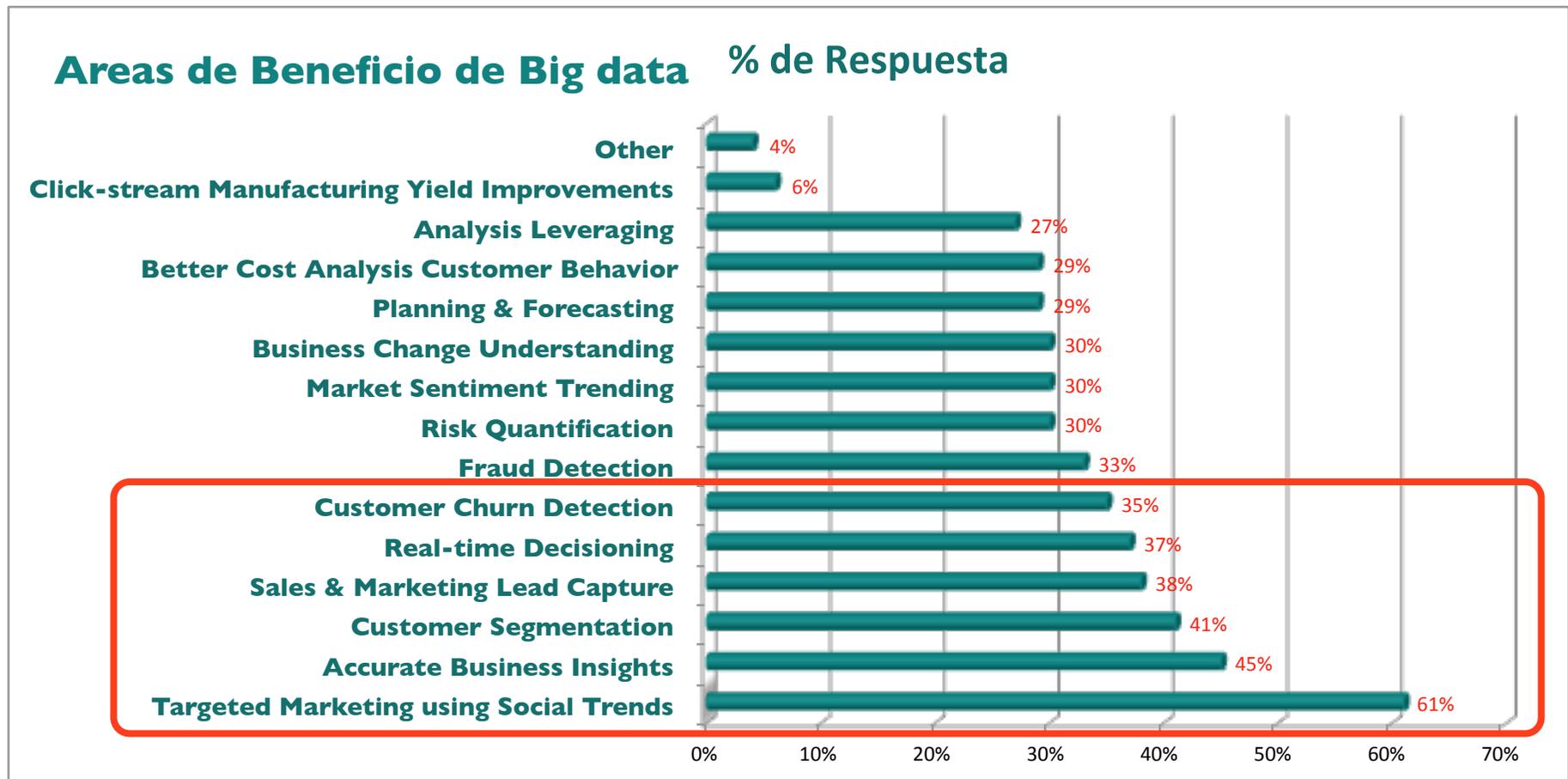
Miembros clave en redes de
llamados telefónicos

◆ Comunidades

Segmentación basada en
relaciones de amistad/Institución

Comunidades de productos
relacionados por boleta en retail

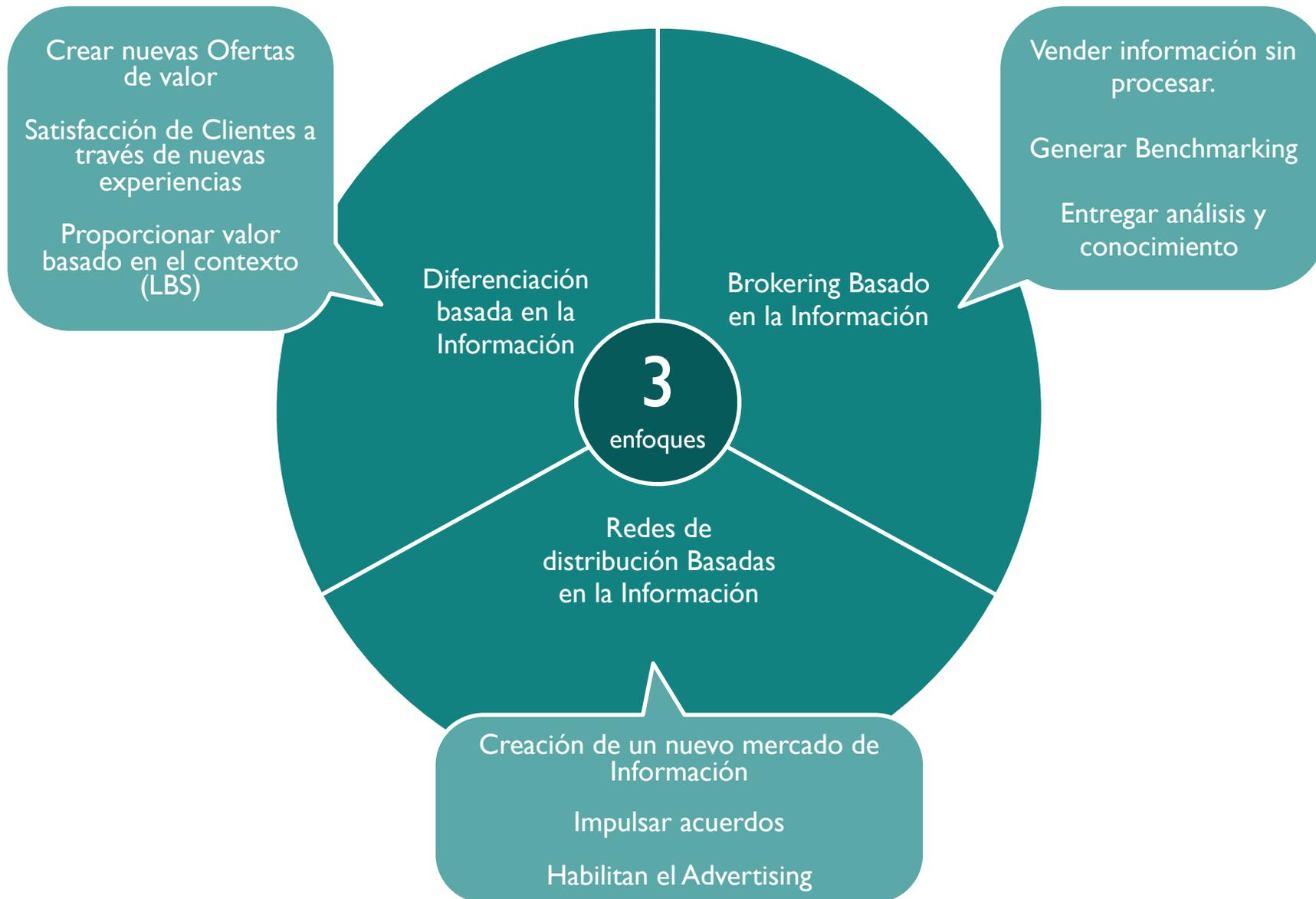
Otras oportunidades con Big Data

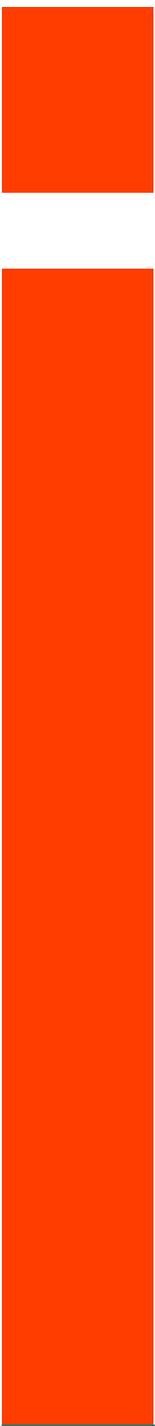


Oportunidades en Marketing, análisis de clientes, web análisis, Gestión de operaciones y análisis de riesgo y/o fraude.

Modelos de negocios...

Qué más hago con la información...





De Big data a Smart Data

El Smart Data ...

Carencia del Big Data

- El Big Data permite reaccionar, pero no innovar, **no reemplaza el juicio humano y hay que revisar los resultados.**
- El Big Data permite trabajar con todos los tipos de datos disponibles, pero **no optimiza en base a las experiencias** y no responde de manera inteligente posibles cambios de escenarios de forma ágil.



“Smart Data le da sentido al Big Data”



Amit Sheth

“...proporcionan **valor** a partir de los desafíos que plantea aprovechar el volumen, la velocidad, la variedad y veracidad de grandes volúmenes de datos, y a la vez que proporciona información **accionable** y mejorar la toma de decisión...”

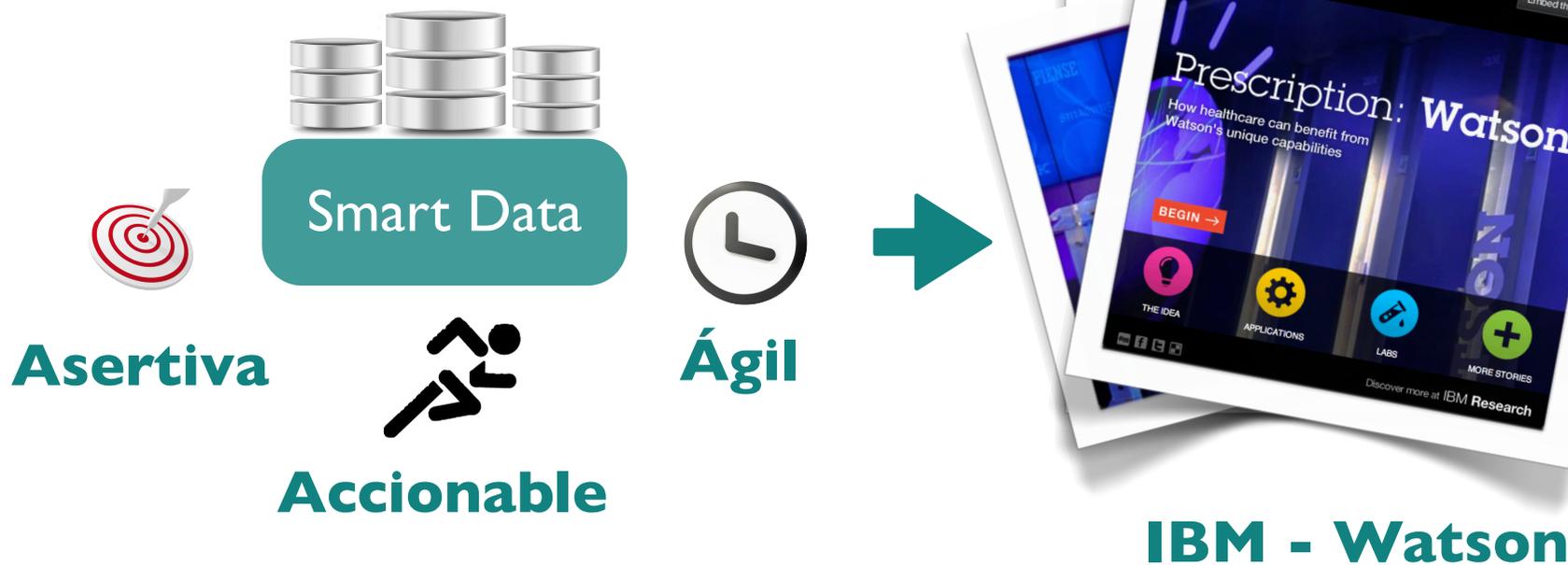
El Smart Data ...

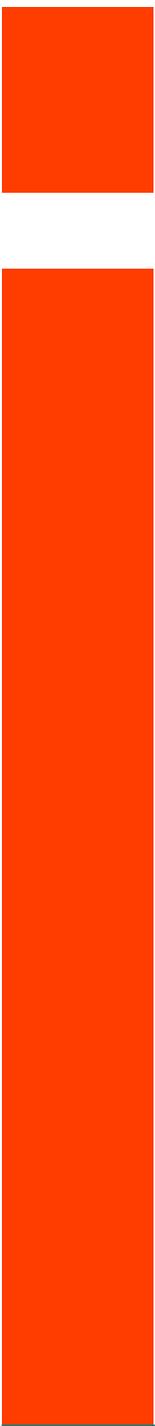


Amit Sheth

“...Smart Data se centra en el valor **accionable** logrado por la **participación de personas** en las fases de creación, procesamiento y consumo de datos para mejorar la **experiencia humana...**”

Las 3's A del Smart Data





La analítica como problema central

El huevo o la gallina



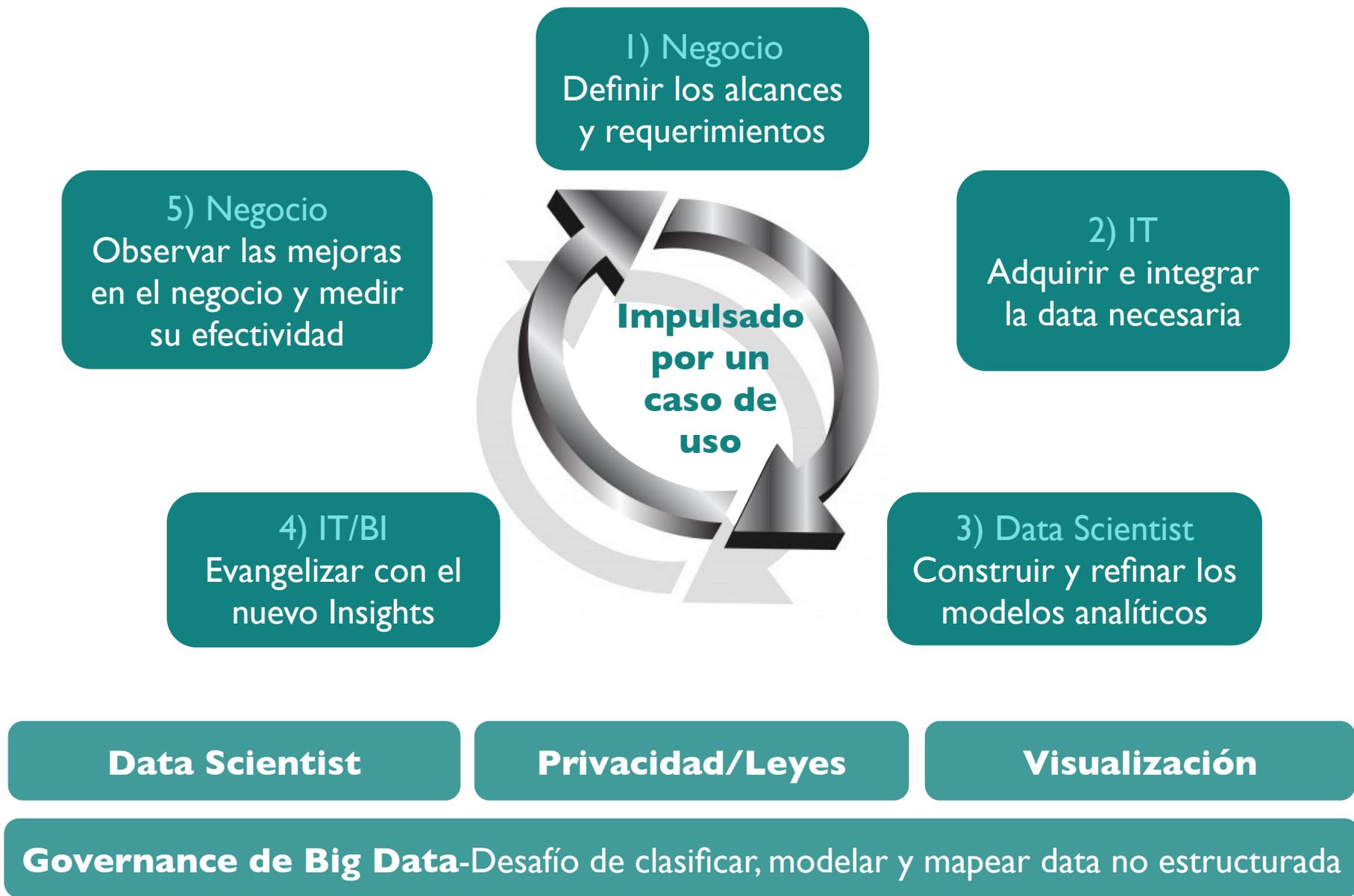
Un caso de uso de Big data



Una Herramienta de Big Data

¿Es la **herramienta** quien me permitirá explorar y detectar casos de uso relevantes a mi empresa (IBM) **o** debo tener una **necesidad** de resolver una problemática que con las **herramientas tradicionales** no puedo?

Enfoque de la problemática



Cuando las técnicas de siempre no funcionan...



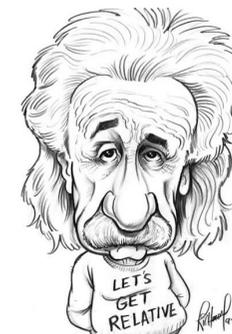
Small Data

Técnicas tradicionales
Métodos estáticos

Volumen y Variedad



Velocidad
C=Cte.



Big Data

Nuevas técnicas
Explorar

Al tener **variadas** fuentes de datos, que crean nueva data a gran **velocidad** generando un **volumen** estratosférico, el paradigma del análisis de datos cambia.

Para enfrentar este nuevo paradigma es necesario generar **nueva analítica, madura y orientada** a resolver las problemáticas que trae consigo el Big Data

El imput central es la analítica

“Big Data no es acerca de la data, es acerca de la analítica”

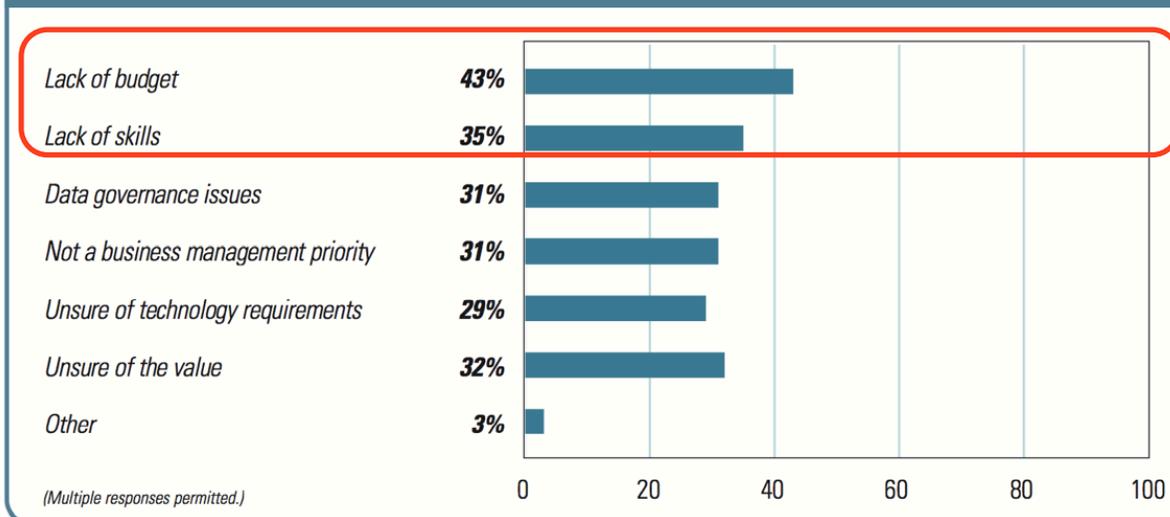
- Gary King Harvard University



- El valor está en la **calidad** de la analítica
- Provee una **ventaja** competitiva sustentable
- Requiere **capacidades y el conocimiento**

Data scientist...

Figure 27: Big Data Business Barriers



Data scientist...



Analista



Data Scientist

“El Rol de los Científicos de la data ha sido descrito como una **parte de análisis y una parte de Arte**”



Data scientist...



Analista



Data Scientist

- * Trabaja desarrollando nuevos algoritmos y modelos para aprovechar los datos, en general tienen grado de Magíster o Ph.D.
- * Tienen vasta experiencia en el uso de herramientas BI y habilidades para comunicar resultados.
- * “Pasión por los datos, pasión por descubrir la Verdad”

Desafío de la visualización de la data...



Minority report - Interacción

Mostrar la data
Visualmente

| | A | B | C | D | E |
|----|---------|------|----------|---------|-------|
| 1 | MES | REF | CANTIDAD | IMPORTE | TOTAL |
| 2 | Febrero | 1245 | 5 | 50 | 250 |
| 3 | Abril | 1265 | 6 | 12 | 72 |
| 4 | Enero | 1245 | 4 | 53 | 212 |
| 5 | Marzo | 1269 | 2 | 45 | 90 |
| 6 | Abril | 1267 | 4 | 25 | 100 |
| 7 | Marzo | 1265 | 6 | 35 | 210 |
| 8 | Junio | 1245 | 8 | 60 | 480 |
| 9 | Enero | 1235 | 12 | 25 | 300 |
| 10 | Febrero | 1236 | 5 | 30 | 150 |
| 11 | Junio | 1278 | 6 | 35 | 210 |
| 12 | Mayo | 1236 | 3 | 45 | 135 |
| 13 | Mayo | 1258 | 4 | 40 | 160 |
| 14 | Abril | 1236 | 5 | 42 | 210 |



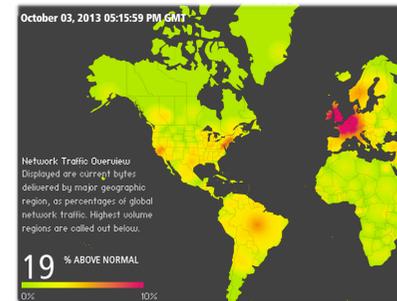
Desafío de la visualización de la data...



Minority report - Interacción

Mostrar la data
Visualmente

```
2012 [notice] apache/2.2.14 (Ubuntu) PHP/5.3.2-1ubuntu4.1
2012 [notice] caught SIGTERM, shutting down
2012 [notice] apache/2.2.14 (Ubuntu) PHP/5.3.2-1ubuntu4.1
2012 [notice] caught SIGTERM, shutting down
2012 [notice] apache/2.2.14 (Ubuntu) PHP/5.3.2-1ubuntu4.1
2012 [notice] caught SIGTERM, shutting down
2012 [notice] apache/2.2.14 (Ubuntu) PHP/5.3.2-1ubuntu4.1
2012 [notice] caught SIGTERM, shutting down
2012 [error] [client 10.1.200.48] File does not exist: /v
2012 [error] [client 10.1.200.48] File does not exist: /v
error_log "Press CTRL-C to stop" for help=
HTML, like Gecko) Chrome/18.0.1025.168 Safari/535.19"
May/2012:12:59:29 -0500 "GET /favicon.ico HTTP/1.1" 404 43
 Gecko) Chrome/18.0.1025.168 Safari/535.19"
May/2012:12:59:40 -0500 "-" 408 0 "-" "-"
May/2012:13:00:55 -0500 "GET /MS.txt HTTP/1.1" 200 1489 "
```





3 ideas a modo de resumen...

3 ideas centrales

1 - Las oportunidades con Big Data están en cómo puedo aprovechar las nuevas fuentes de información, aprendiendo con problemáticas pequeñas, para luego generar analítica de calidad para Big Data

2 - El huevo es primero, se debe pensar en una problemática, en un caso de uso. Hacer una lista y levantar las necesidades que tengo y las preguntas que deseo responder

3 - El nuevo paradigma de los datos requiere gente preparada que sea capaz de innovar y probar nuevas herramientas, no son solo analistas, son expertos capaces de adaptar la data disponible a las problemáticas específicas de negocio

CEiNE.

Centro de Investigación en
Inteligencia de Negocios



INGENIERIA INDUSTRIAL
UNIVERSIDAD DE CHILE

Telefonica



fcfm

FACULTAD DE CIENCIAS
FÍSICAS Y MATEMÁTICAS
UNIVERSIDAD DE CHILE