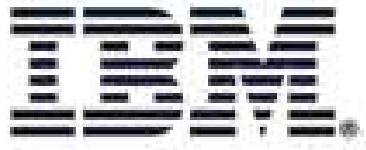


La oportunidad de Big Data para las organizaciones en Venezuela ¿Está Ud. preparado?

Leonardo González Barceló
IBM Latin America Big Data
Sales Leader

Caracas, Junio 11, 2013





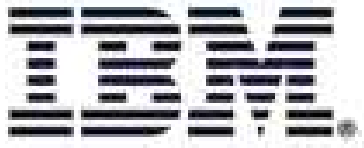
¿Es Big Data Un Tópico de Tendencia Actual?

Big Data: 2,100,000,000 resultados

BI: 600,000,000 resultados

DW: 90,200,000 resultados





IBM Institute for Business Value:

3 de 4 organizaciones tienen actividades de Big Data en marcha

- **La era de Big Data**

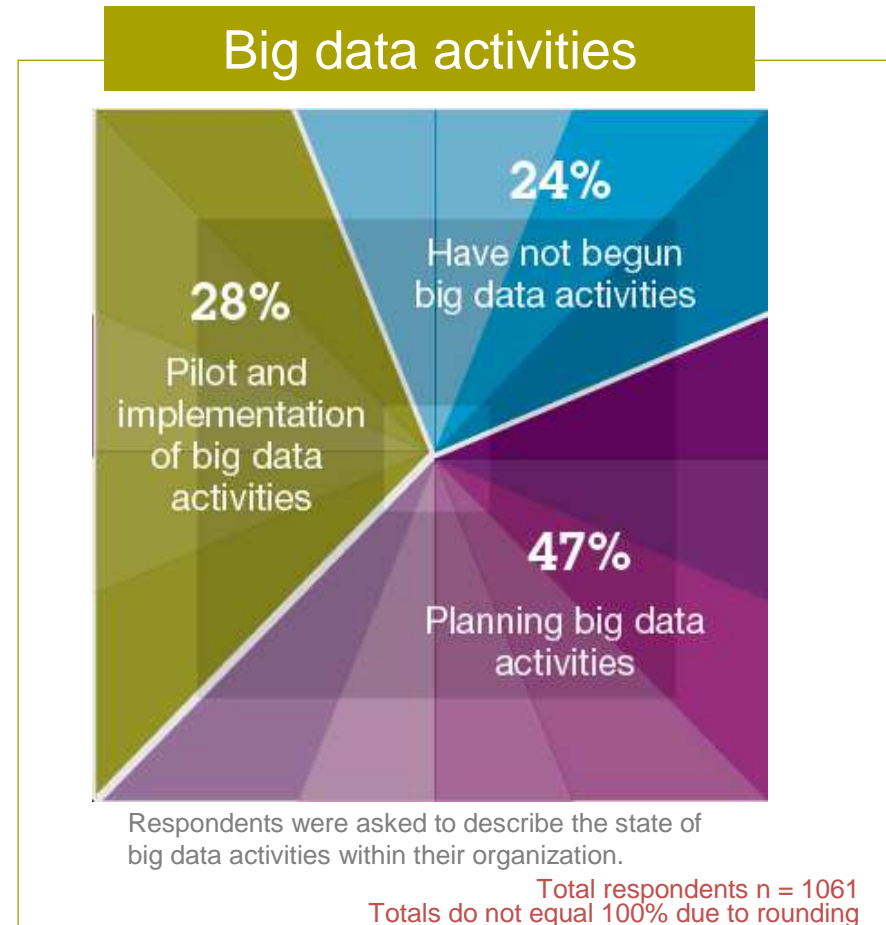
- Casi la mitad de las organizaciones encuestadas, reportaron análisis internos acerca de planes de Big Data
- Big data se ha movido fuera de IT hacia el área de negocios

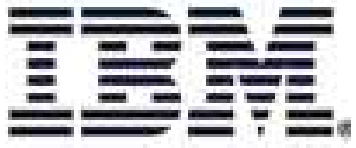
- **Poniendo en Marcha**

- Más de un 25% de las organizaciones tienen pilotos o implementaciones de Big Data activas
- El uso de Big Data está siendo real

- **Acceleración Prevista**

- La cantidad de pilotos ejecutándose, sugiere que las implementaciones de Big Data van a crecer de forma exponencial en los próximos años
- Una vez que las tecnologías base estén instaladas, el uso se propagará rápidamente en toda la organización





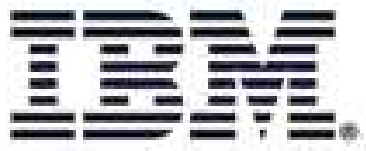
Los Patrones de Comportamiento Organizacional son consistentes a través de las 4 etapas de adopción de Big Data

Adopción de Big Data



When segmented into four groups based on current levels of big data activity, respondents showed significant consistency in organizational behaviors

Total respondents = 1061
Totals do not equal 100% due to rounding

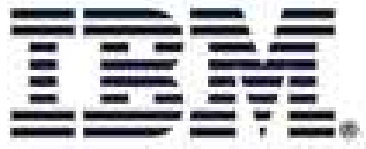


Datos Temporales Automáticos y Enriquecidos Especialmente

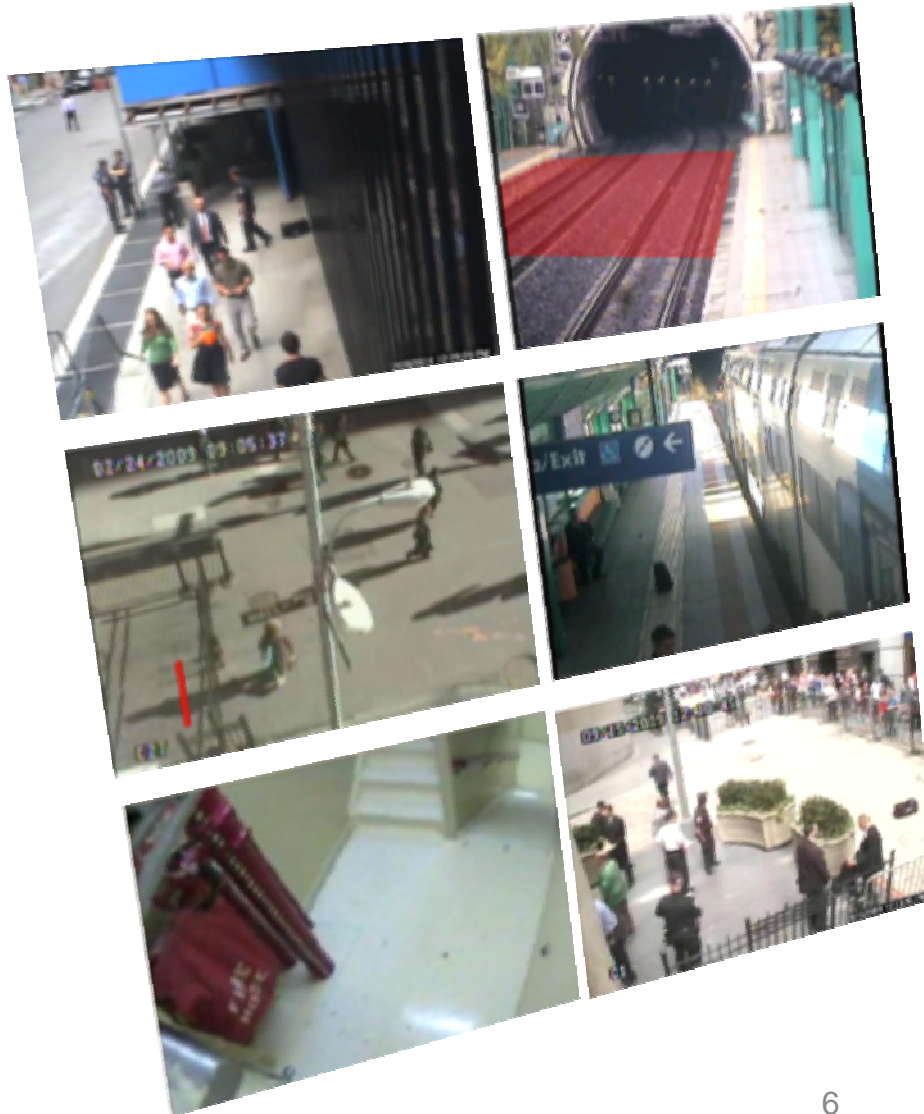


- En el 2010, Adam Savage, de “Cazadores de mitos”, sacó una foto de su vehículo con su smartphone. Luego publicó la foto en su cuenta de Twitter con la frase “Camino al trabajo”.
- Como sacó la foto con el smartphone, la imagen contenía metadatos que revelaban la ubicación geográfica exacta de dónde se había tomado la foto.
- Con sólo tomar y publicar una foto, Savage reveló la ubicación exacta de su casa, el vehículo que conduce y la hora a la que sale a trabajar.
- Ahora, más de 650.000 personas saben donde queda su casa.....





Datos Temporales Automáticos y Enriquecidos Espacialmente

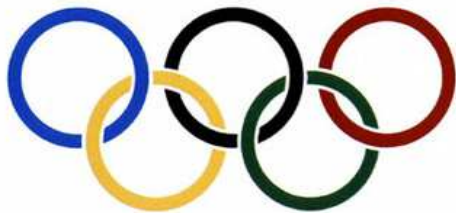


- En Londres, mientras se viaja desde el trabajo de regreso a casa, una persona es captada en video al menos 150 veces.
- Estos datos son utilizados para analizar prevención de crimen y seguridad nacional, (pero qué más se podría hacer?)



IBM

Beijing 2008



83x

6,000,000 usuarios on
Twitter

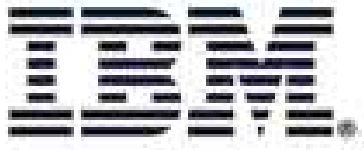
enviando 300,000
tweets por día

500,000,000 usuarios en
Twitter

enviando 400,000,000
tweets por día

1.333x

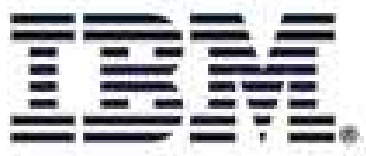




El tablero Olímpico de Expulsiones

Al menos 4 años de Entrenamiento
desperdiciados en 140 caracteres

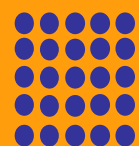




Nos estamos moviendo a una Nueva Era de Tecnologías de la Información y Desafío de las 4 V

12+ terabytes

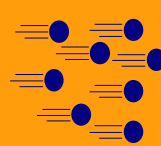
of Tweets
creados
diariamente



Volume

5+ million

Eventos de trading
por segundo.



Velocity

100's

De diferentes tipos de datos.



Variety

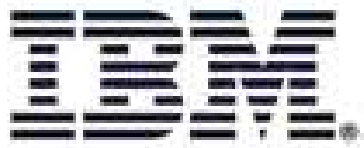


Veracity

Solo **1 in 3**

Tomadores de decisiones
Creen en la información que
tienen.





2011

2009

1.8 ZB

1 ZB

1 ZB=1T GB

2010

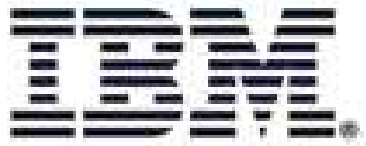
2020

35 ZB

(1 ZB = 1B TB)

4 Billones
de iPods
de 8GB





¿Se Puede Monetizar a una Persona de los Medios Sociales?



THE WALL STREET JOURNAL | **ARTS & ENTERTAINMENT**

DECEMBER 28, 2011, 7:00 PM ET

Man Leaves Job, Takes Twitter Followers, Gets Sued

"This will establish precedent in the online world, as it relates to ownership of social media accounts. We've actually been waiting to see such a case as many of our clients are concerned about the ownership of social media accounts vis-a-vis their branding."

25,826

Tweets

211

Following

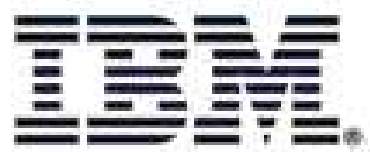
24,265

Followers

952

Listed

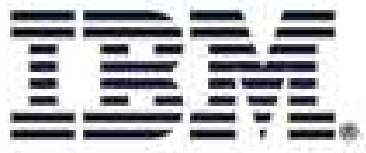




¡Y ya no es sólo para personas!



Fuente: http://www.mediabistro.com/alltwitter/how-to-get-fido-to-tweet-infographic_b17521



Los 5 Patrones de Big Data



Exploración de Big Data

Buscar, visualizar, entender big data para mejorar el conocimiento del negocio.



Visión de 360° del Cliente Mejorada

Lograr una verdadera vista unificada que cuente además con fuentes externas a la empresa.



Seguridad/Inteligencia Extensión

Menor riesgo, detectar fraude y monitorear la seguridad en la web en tiempo real.



Análisis de Operaciones

Anlizar una variedad de datos generados por instrumentos para mejorar los resultados del negocio.

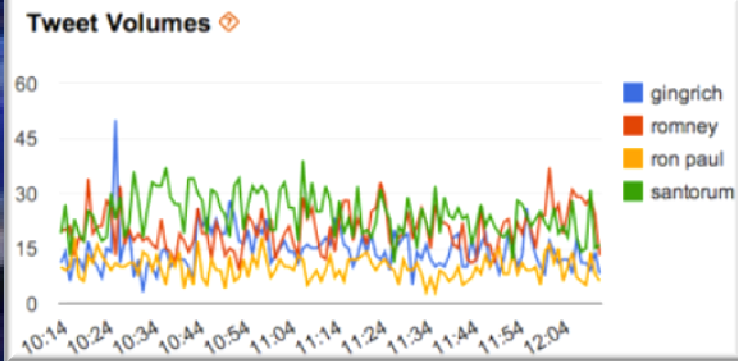


Enriquecimiento del Data Warehouse

Integrar las capacidades de big data y data warehouse para mejorar la eficiencia operacional y enriquecimiento de los datos

University of Southern Cal Monitoreo Debate Político

- Solución desarrollada para medir el sentimiento público durante las Primarias Republicanas y el Debate Presidencial
- Examina tendencias, volumen y contenido de millones de Twitters en tiempo-real
- Aceleradores Analíticos para entender el sentimiento (positivo, negativo, neutro)
 - Stream Computing y visualización
- Beneficios
 - Despliegue en Tiempo-Real del sentimiento público a medida que los candidatos responden preguntas
 - Predicción del ganador del debate basado en la opinión pública, versus solo la opinión de los analistas políticos





“Ayudar a detectar condiciones de riesgo vital hasta con 24 horas de anticipación”

University of Ontario Institute of Technology (UOIT)

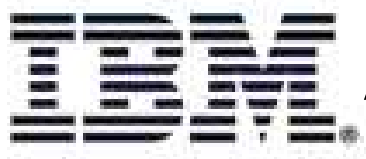
Ayuda a Detectar Riesgo de Bebés Prematuros con Anticipación

Capacidades Utilizadas:

Stream Computing

- Información que se utiliza para agregar y analizar cada 30 a 60 minutos, generando tableros de control
- Analizar 1.000 piezas de información de diagnóstico médico p/seg y se almacena en un modelo dinámico
- Resultados: Baja de 20% en la mortalidad de bebés prematuros del grupo de control en los ensayos
 - 120 bebés observados: 120K mensajes/seg... miles de millones/por día





Aplicaciones para Analítica de Big Data

Servicios de Salud Más Inteligentes



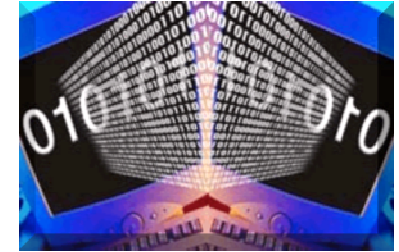
Ventas Multicanales



Finanzas



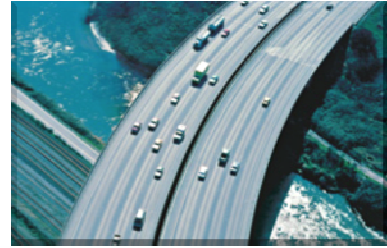
Análisis de Log



Seguridad del Territorio Nacional



Control de Tráfico



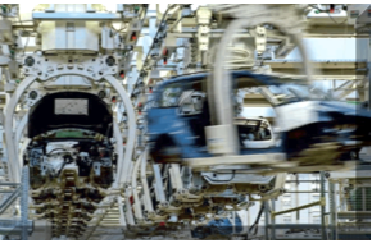
Telecomunicaciones



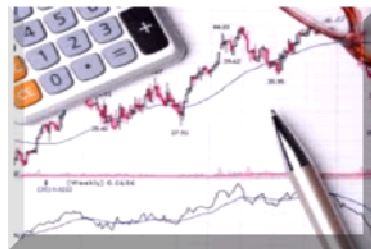
Calidad de Búsqueda



Fabricación



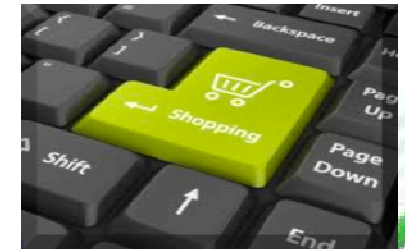
Analítica de Comercio

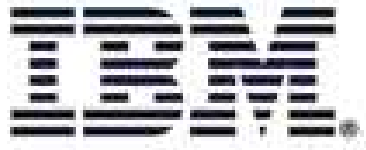


Fraude y Riesgo



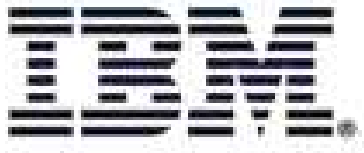
Minorista: Producir en serie, NBO





Plataforma IBM para BIG DATA



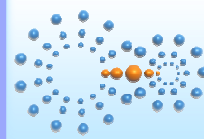


¿Qué hace una Plataforma de Big Data?



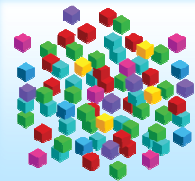
Analiza una Amplia Variedad de Información

Analítica novedosa en un amplio conjunto de información mixta que antes no se podía analizar



Analiza Información en Movimiento

Análisis de flujo de datos
Saturación de grandes volúmenes de datos y análisis ad-hoc



Analiza Volúmenes de Información Excesivos

Procesa de manera rentable y analiza PBs de información
Administra & analiza grandes volúmenes de datos estructurados relacionales



Descubre y Experimenta

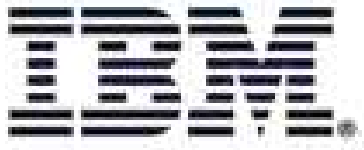
Analítica Ad-hoc, descubrimiento de datos y experimentación



Gestiona y Planifica

Aplica estructura de datos, integridad y control para garantizar la coherencia de las consultas repetibles





The IBM Big Data Platform

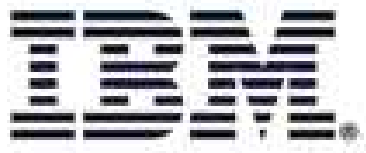
Solutions

Analytics and Decision Management

IBM Big Data Platform

Big Data Infrastructure





The Big Data Platform

Solutions

Analytics and Decision Management

IBM Big Data Platform

Data
Warehouse



***Entrega
Conocimiento
Detallado mediante
in-database
analytics avanzado
& operational
analytics***

**PureData for
Analytics &
Operational**

- Appliance Diseñado y Integrado para Analíticos
- Appliance para análisis operacional configurable

Big Data Infrastructure





NYSE: Market Risk/ Surveillance

Situación

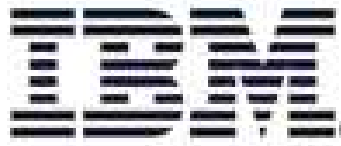
- Desafío de almacenar 7 años de historia sin incurrir en costos significativos
- Procesos de carga muy lentos, performance de queries de transacciones: 26 horas, 7 minutos para análisis del piso completo
- Volumen de datos creciendo 100-200% por año

Solución

- PureData for Analytics > 1PB
- Primera App desarrollada en <2 meses
- ROI dentro del primer año, solo por beneficios de storage

Resultados

- Análisis de +7 años de datos
- Alto rendimiento en consultas, sin necesidad de hacer tuning;
- Análisis de transacciones : 2 min, Análisis de piso completo: 5 seg.



The Big Data Platform

Solutions

Analytics and Decision Management

IBM Big Data Platform

Stream
Computing



Data
Warehouse

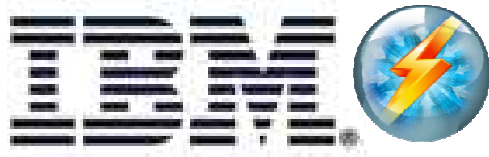


***Analiza Data en
Movimiento para
obtener
resultados en
“Tiempo Real”***

InfoSphere Streams –
Solución de software que
permite analizar
volúmenes masivos de
datos en movimiento, con
tiempos de respuesta de
milisegundos.

Big Data Infrastructure



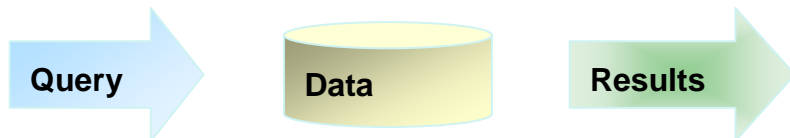


Stream Computing – Analizar datos en movimiento

Forma Tradicional



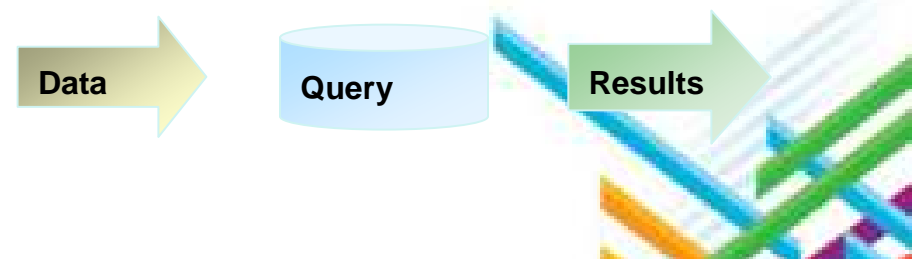
- Encontrar hecho histórico
- Buscar y analizar la información almacenada en el disco
- Paradigma Batch, modelo pull
- Query-driven: envía consultas a datos estáticos

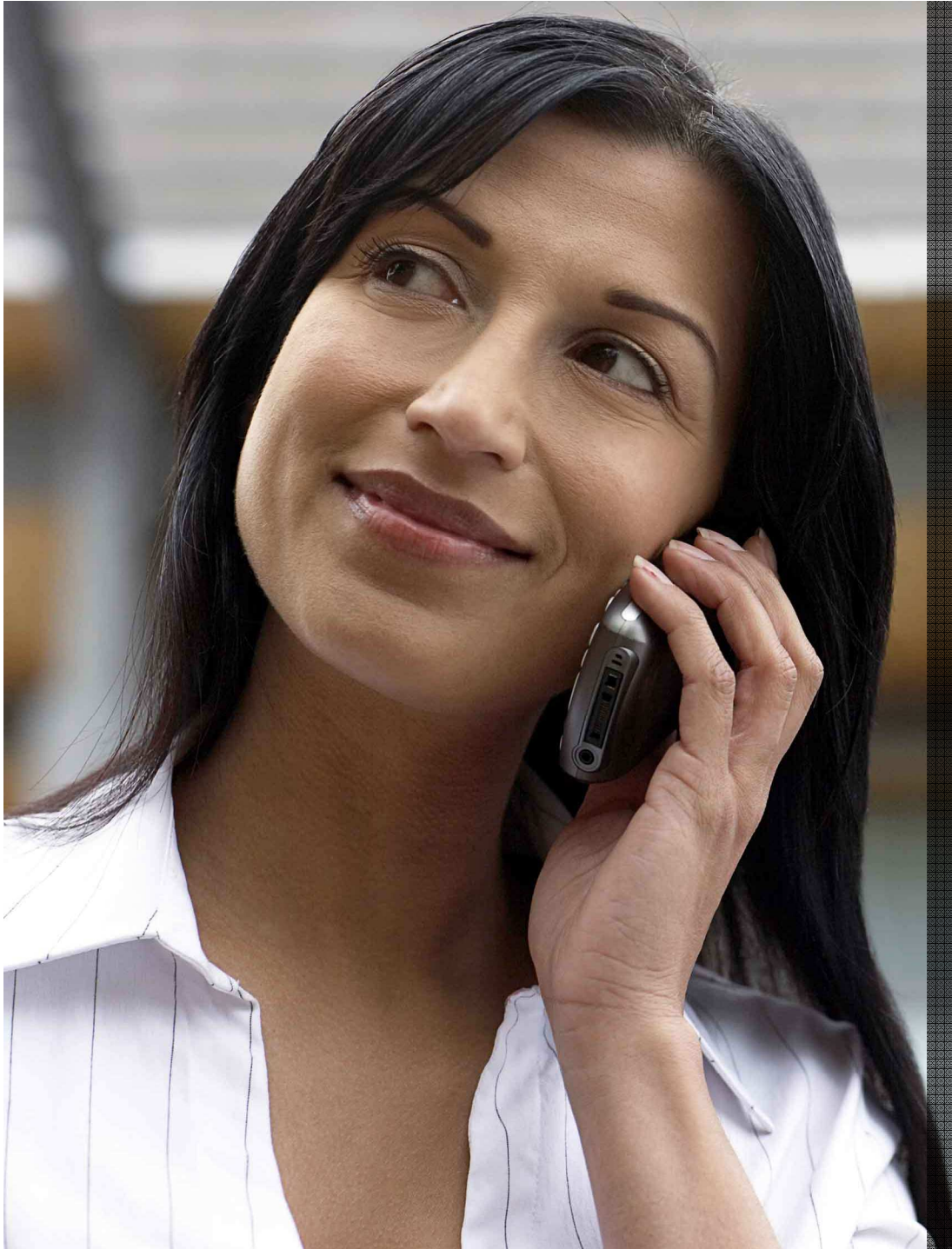


Stream Computing



- Encontrar hecho actual
- Analizar datos en movimiento - antes de que se almacena
- Paradigma baja latencia, modelo push
- Data driven – traer los datos a la consulta





Una Telco de Asia reduce los costos de facturación y mejora la satisfacción del cliente

Capacidades:

Stream Computing
Analytic Accelerators

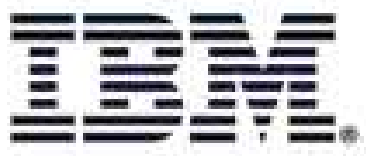
Mediación en tiempo real y el análisis de
6B CDRs por día

tiempo de procesamiento de datos
reducido de

12 hrs a 1 seg.

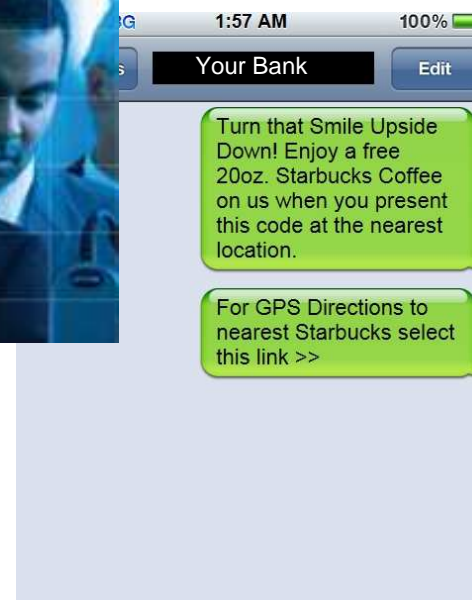
Costo de Hardware reducido a 1/8th

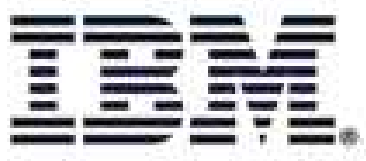
Solución proactiva de problemas (ej..
llamadas perdidas) mejoro la
satisfacción del cliente.



Caso de Uso – Analytics de Video

- Detección de Dispositivos
– x ej. Skimmers
- Reconocimiento Facial - Cliente

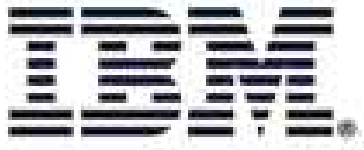




Caso de Uso - Multi-Feed Video Analytics

- Detecta Eventos Sospechosos – Largas Filas, robos en cajero, graffiti
- Alertas – Cambios en iluminación, basura, etc.
- Registro la historia de los eventos
- Mantiene el ATM en buen estado





La Plataforma de Big Data de IBM

Plataforma basada en Hadoop para analizar Petabytes de datos, ya sea estructurados y No-estructurados

■ **InfoSphere BigInsights**

■ **Open source Hadoop con capacidades empresariales**

- Performance Optimization
- Development tooling
- Enterprise integration
- Analytic Accelerators
- Application and industry accelerators
- Visualization
- Security

Solutions

Analytics and Decision Management

IBM Big Data Platform

Hadoop System



Stream Computing

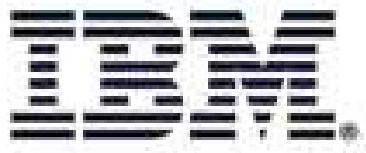


Data Warehouse



Big Data Infrastructure





Origen y Evolución de Hadoop



Publishes
MapReduce,
GFS Paper



Apache OpenSource
MapReduce & HDFS
projects created



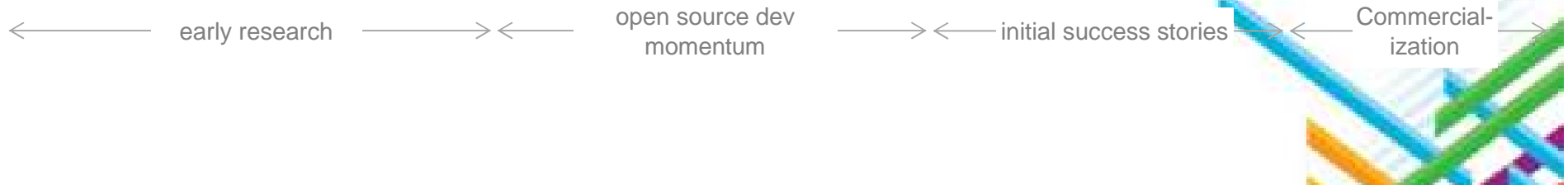
Runs 4,000
node Hadoop
Cluster

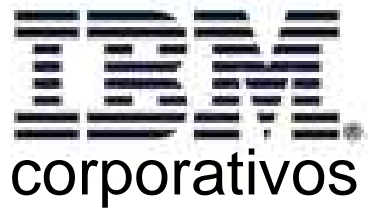


Launches SQL
Support for
Hadoop

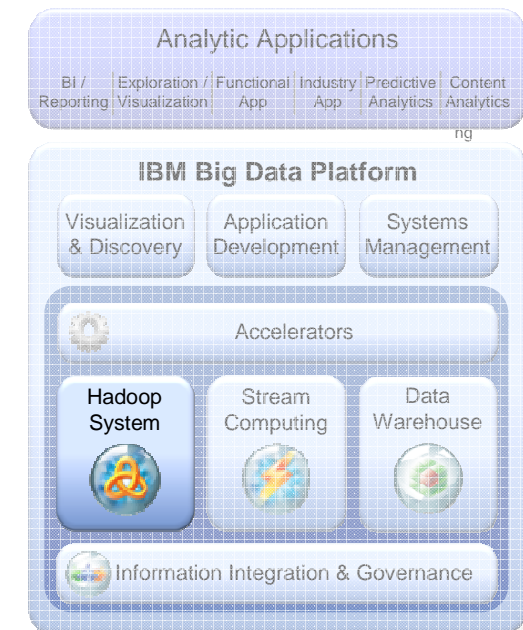
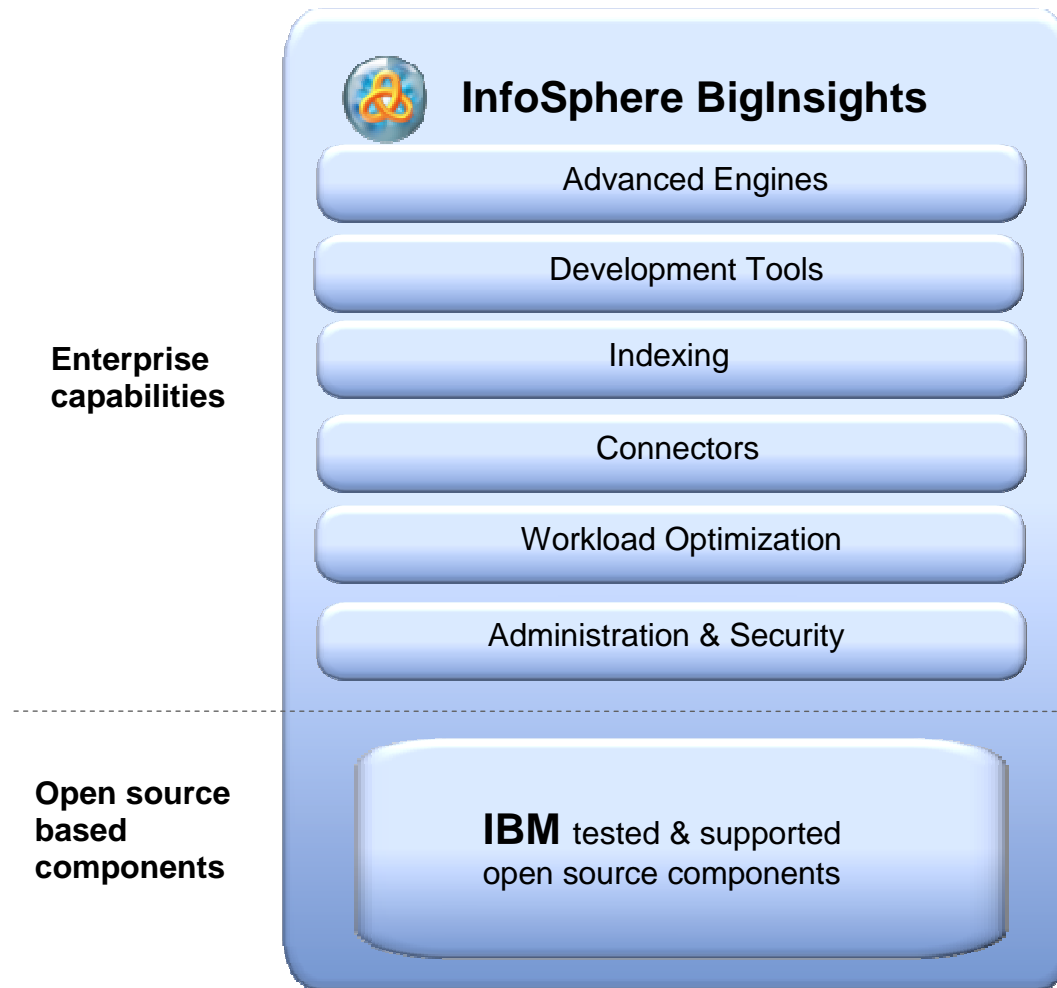


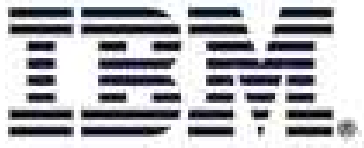
Big Insights
announced





BigInsights está desarrollado sobre Hadoop con capacidades para desarrollos de tipo

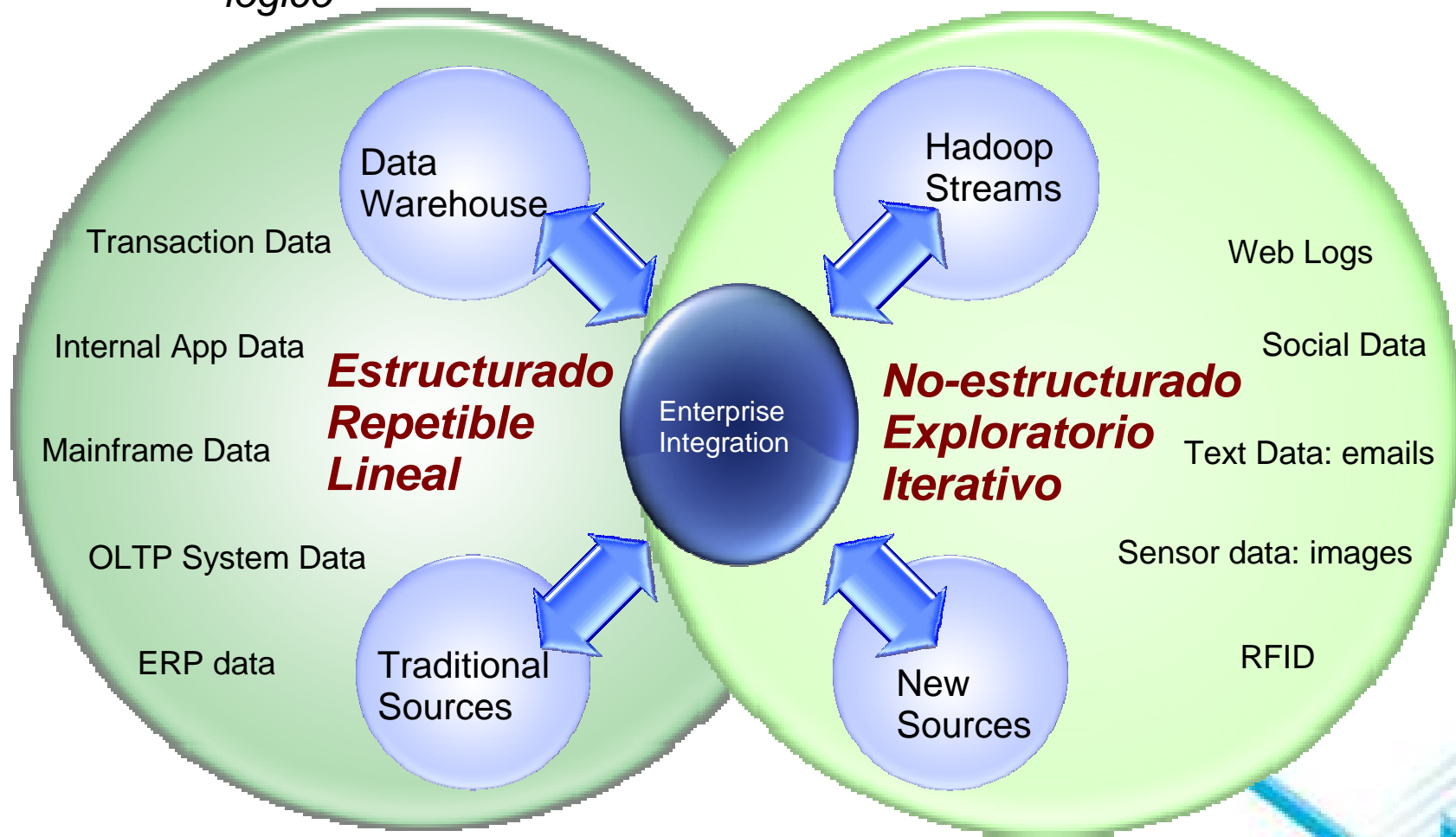


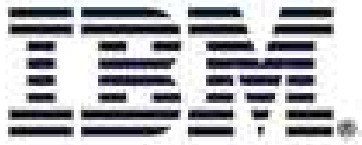


Data Warehouse vs Big Data

Enfoque Tradicional
Estructurado, analítico,
lógico

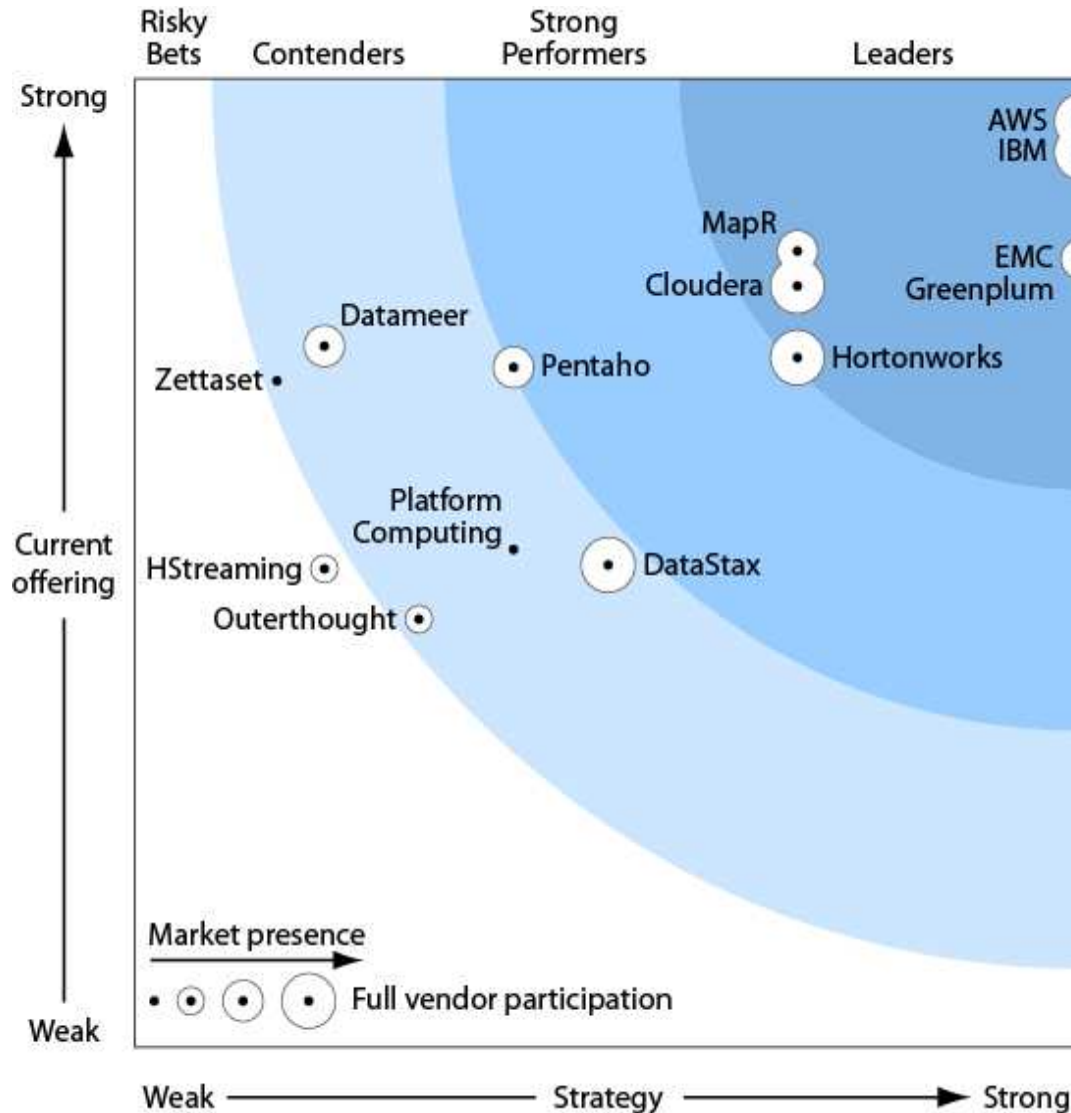
Nuevo Enfoque
Creativo, intuitivo





“IBM has the deepest Hadoop platform and application portfolio.”


First Ever Forrester Wave on Hadoop



FORRESTER®

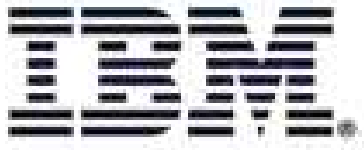
“IBM has the deepest Hadoop platform and application portfolio. IBM, an established EDW vendor, has its own Hadoop distribution; an extensive professional services force working on Hadoop projects; extensive R&D programs developing Hadoop technologies; connections to Hadoop from its EDW.”

—The Forrester Wave™: Enterprise Hadoop Solutions, 1Q12



Vestas Ejecuta Modelos Climáticos para determinar la ubicación más óptima de cada turbina

- Se dispone de datos públicos sobre vientos en redes de 284km x 284km (2.5o LAT/LONG)
- Tener más datos implica tener modelos más precisos y abundantes (sumando cientos de variables)
 - Biblioteca de vientos Vestas de 3.5 PB: crecerá más de 6 PB en el período próximo
 - Redes de granularidad de 27km x 27km: conducirán a simulaciones de 9x9, 3x3 a 10m x 10m
- Identificación de una reducción de semanas a horas en la colocación de turbinas
- Perspectiva: Ver la biblioteca de Vientos Vestas como TV HD tomaría 70 años



La Plataforma de Big Data de IBM

Gobernar la Calidad de los Datos y gestionar el ciclo de vida de la información

- **InfoSphere Information Server** – Limpia los datos, monitorea la calidad e integra big data con los sistemas actuales
- **InfoSphere Optim** – Administra el ciclo de vida de los datos.
- **InfoSphere Master Data Management** – Administra los datos maestros
- **InfoSphere Guardium** – Monitorea la seguridad de las bases de datos en tiempo real

Solutions

Analytics and Decision Management

IBM Big Data Platform

Hadoop System



Stream Computing



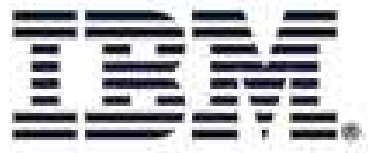
Data Warehouse



Information Integration & Governance

Big Data Infrastructure





La Plataforma de Big Data de IBM

Solutions

Analytics and Decision Management

IBM Big Data Platform



Accelerators

Hadoop
System



Stream
Computing



Data
Warehouse

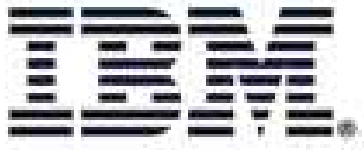


Information Integration & Governance

Big Data Infrastructure

***Aumenta la
velocidad de
generar valor
mediante
aceleradores pre-
definidos***

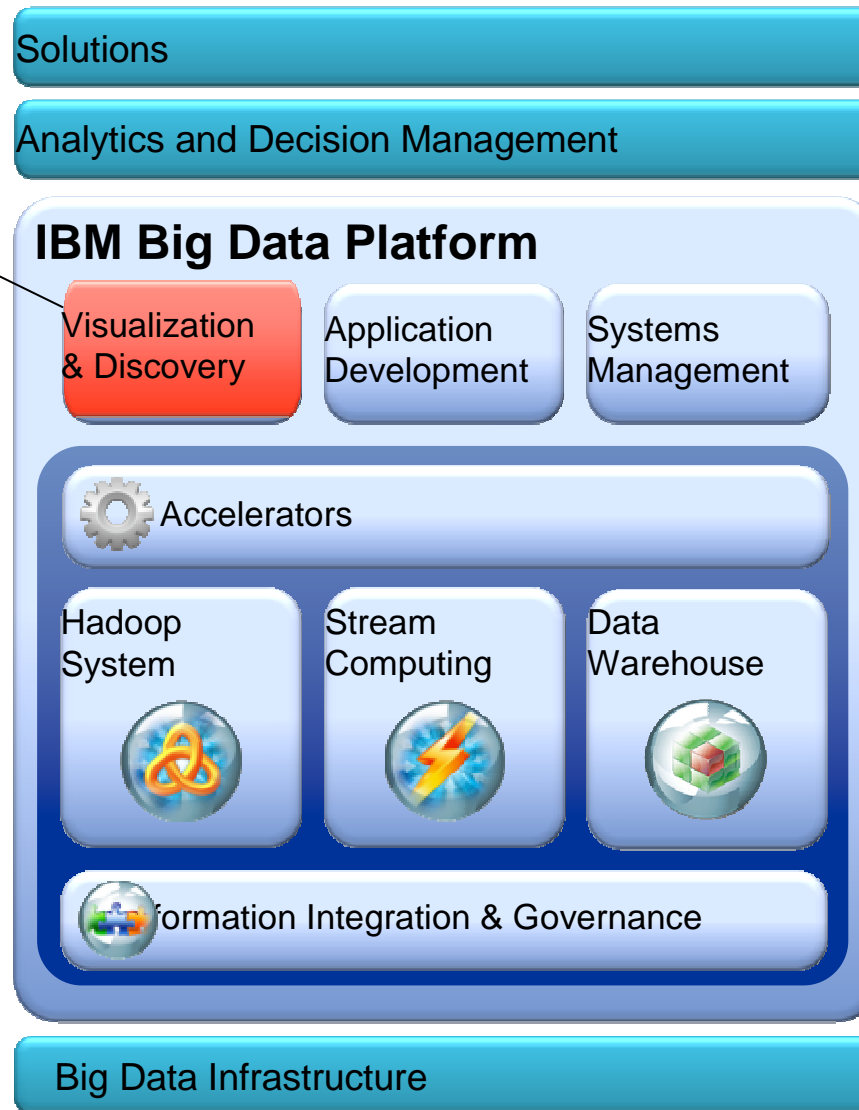
- **Accleradores Analíticos** – text analytics, geospatial, time-series, data mining
- **Accleradores de Aplicaciones** – financial services, machine data, social data, Telco event data
- **Modelos de Industria**
 - Modelos de datos y procesos basados en las mejores prácticas de cada Industria.

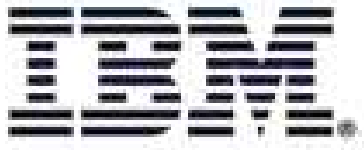


La Plataforma de Big Data de IBM

*Descubra,
entienda, busque
y navegue fuentes
de big data
federadas*

- **InfoSphere Data Explorer** – Software para Descubrimiento y navegación que provee acceso en tiempo real donde se fusiona Big Data junto con datos desde cualquier aplicación dentro de la organización, para obtener mayor conocimiento.





La Plataforma de Big Data de IBM

Las nuevas aplicaciones analíticas impulsar los requisitos para una plataforma big data

- Integrar y gestionar toda la variedad, velocidad, volumen y veracidad de datos (V4)
- Aplicar análisis avanzados a la información en su forma nativa
- Visualizar todos los datos disponibles para análisis ad-hoc
- Entorno de desarrollo para la construcción de nuevas aplicaciones analíticas
- Optimizado para Workload y scheduling
- Seguridad y Gobernabilidad

Solutions

Analytics and Decision Management

IBM Big Data Platform

Visualization
& Discovery

Application
Development

Systems
Management



Accelerators

Hadoop
System



Stream
Computing

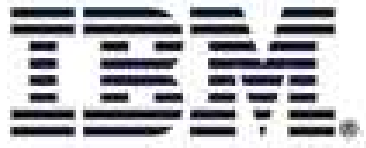


Data
Warehouse



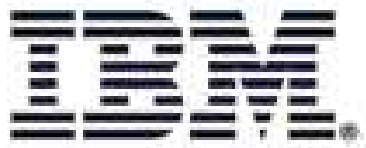
Information Integration & Governance

Big Data Infrastructure

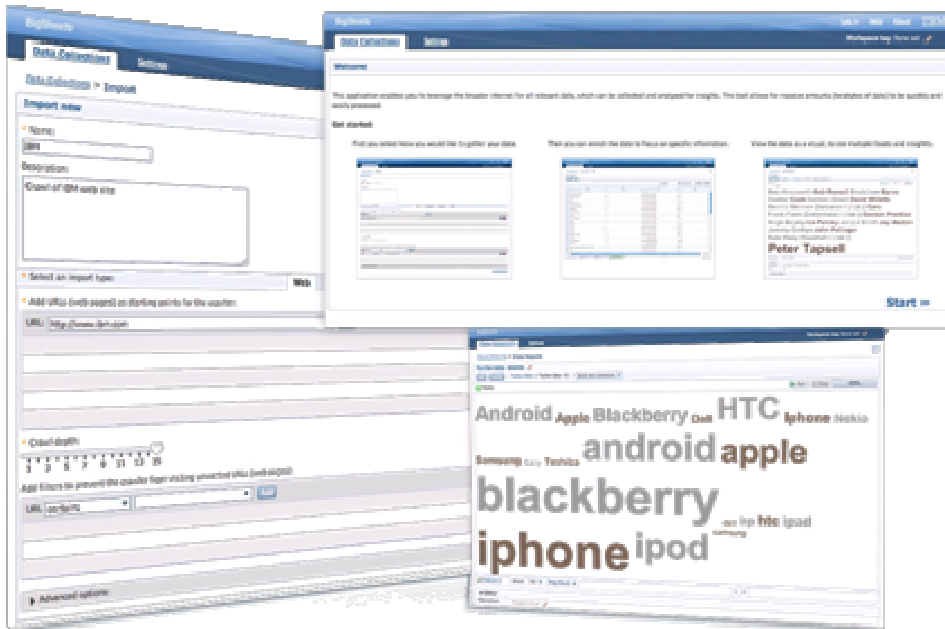


¿Por qué elegir IBM
para Big Data?
Del Lado de la Solución



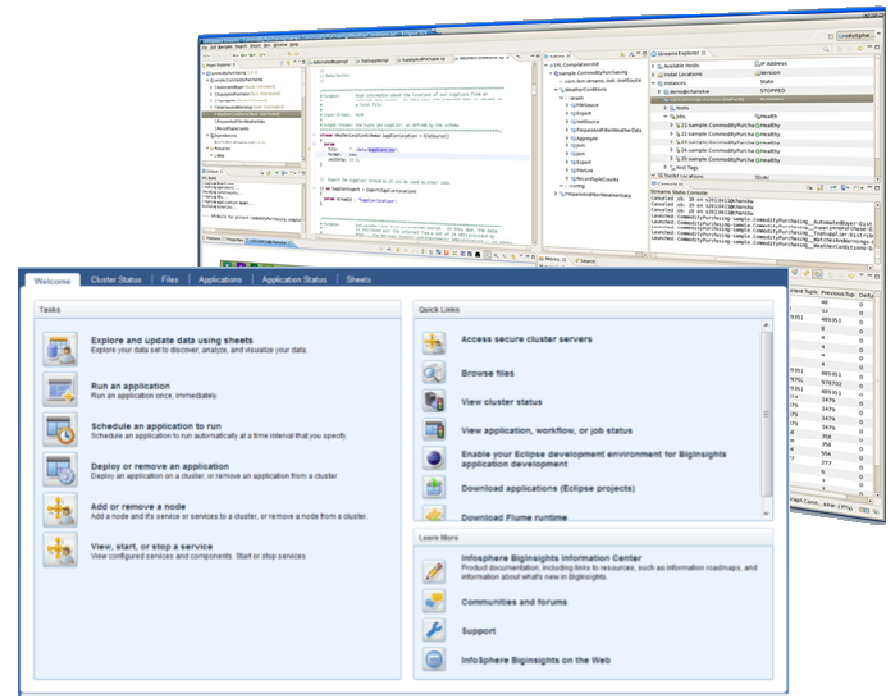


Facilidad de Uso para los Desarrolladores y los Usuarios de Negocios



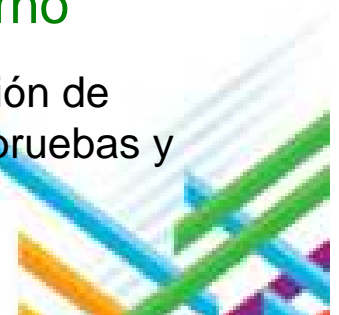
Visualización del Usuario Final

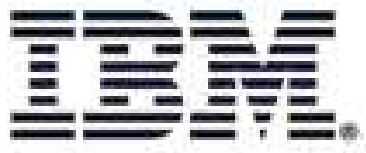
Exploración de datos, arrastre y analítica



Desarrollo de Entorno

Entorno conocido de gestión de herramientas y codificación, pruebas y optimización





¿Qué es BigSheets?

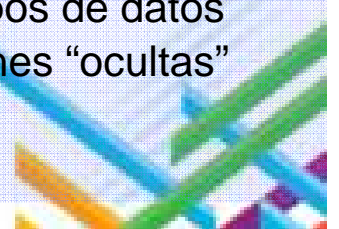
Herramienta analítica de Big Data basada en navegador para usuarios empresariales

Desafíos Big Data...

- Los usuarios empresariales necesitan un enfoque sin programación para analizar Big Data
- Es extremadamente difícil encontrar percepciones empresariales en datos de origen múltiple con formatos distintos
- Traducir datos sin explotar en percepciones empresariales factibles es un requisito común que requiere una visualización

¿En qué puede ayudar BigSheets?

- La interfaz de descubrimiento con apariencia de planilla de cálculo les permite a los usuarios analizar fácilmente Big Data **SIN PROGRAMACIÓN**
- Los "lectores" **INTEGRADOS** pueden funcionar con los datos en varios formatos comunes
 - Series JSON, CSV, TSV, producto de ruta web. . .
- Los usuarios pueden combinar y explorar **VISUALMENTE** varios tipos de datos para identificar las percepciones “ocultas”



[Data Collections](#) > [View Results](#) > [Create](#)


Unnamed Collection(1)


Save Exit


fx ☐ ☐ Fit column(s)


	A	B	C	D	E	F	G	H
	EMPNO	FIRSTNAME	LASTNAME	WORKDEPT	PHONENO	HIREDATE	JOB	EDLEVEL
1	10	Jennifer	Noonan	A00	3978	19950101	PRES	18
2	20	Pablo	Reinoso	B01	3476	20031010	MANAGER	18
3	30	Patricia	Schiapelli	C01	4738	20050405	MANAGER	20
4	50	Sanderson	Broudy	E01	6789	19790817	MANAGER	16
5	60	Franco	Bruno	D11	6423	20030914	MANAGER	16
6	70	Hedi	Simane	D21	7831	20050930	MANAGER	16
7	90	Coleen	Rieder	E11	5498	20000815	MANAGER	16
8	100	Ramesh	Khanna	E21	972	20000619	MANAGER	14
9	110	Andrew	King	A00	3490	19880516	SALESREP	19
10	120	Robert	O'Wager	A00	2167	19931205	CLERK	14
11	130	Heidi	Slimane	C01	4578	20010728	ANALYST	16
12	140	Peggy	Bonifacino	C01	1793	20061215	ANALYST	18
13	150	Jay	Longley	D11	4510	20020212	DESIGNER	16
14	160	Jun	Ashida	D11	3782	20061011	DESIGNER	17
				D11	2890	19990915	DESIGNER	16
				D11	1682	20030707	DESIGNER	17
				D11	2986	20040726	DESIGNER	16
				D11	4501	20020303	DESIGNER	16
				D11	942	19980411	DESIGNER	17
				D11	672	19980829	DESIGNER	18
				D21	2094	19961121	CLERK	14
				D21	3780	20041205	CLERK	17


Select a type of sheet:


 Filter


 Macro


 Load


 Pivot


 Combine

 Union

 Limit

 Distinct

 Copy

 Formula

Add Sheet using by entering a formula

Add sheets

Unnamed Collection

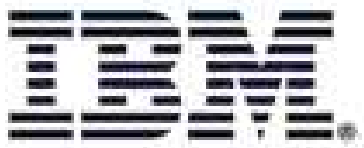
< >

Run

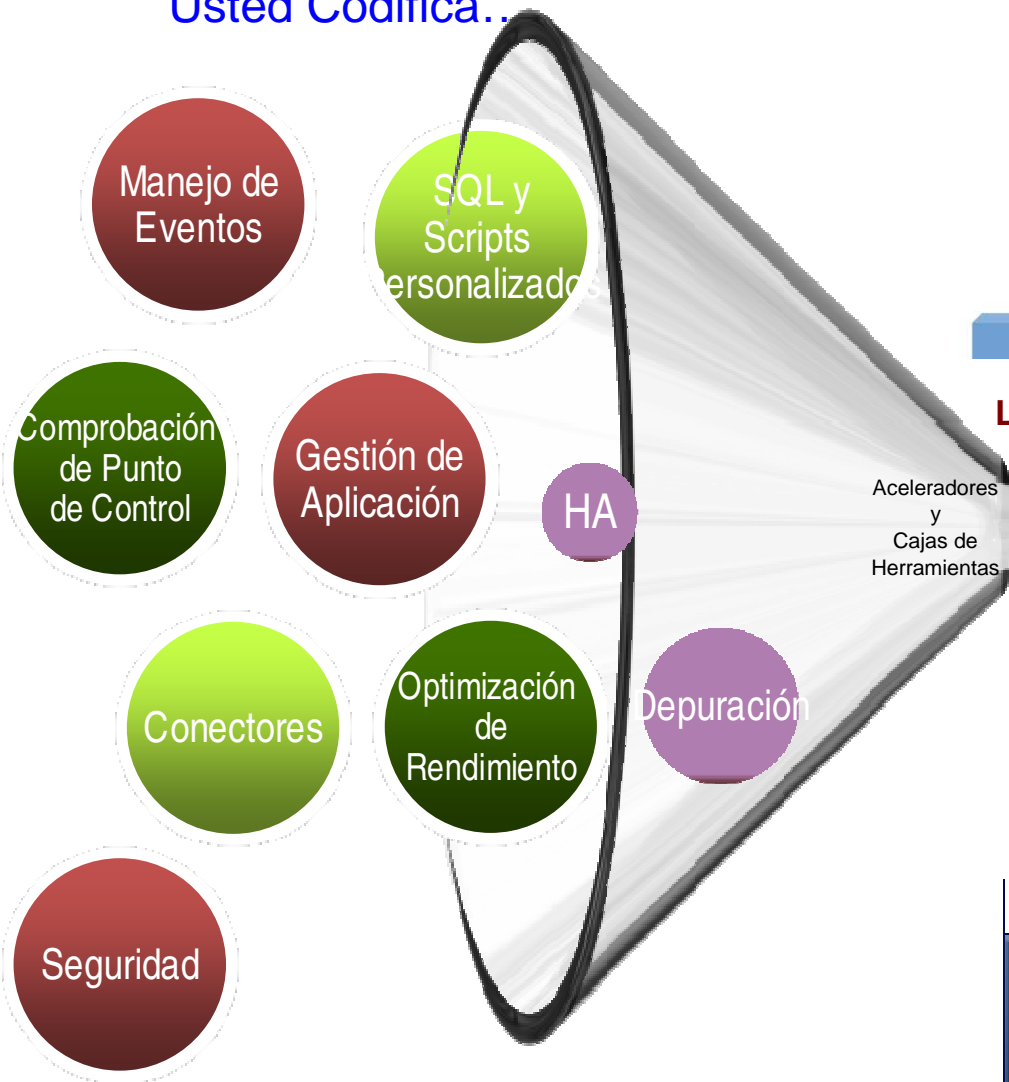
Stop

50%

Ready

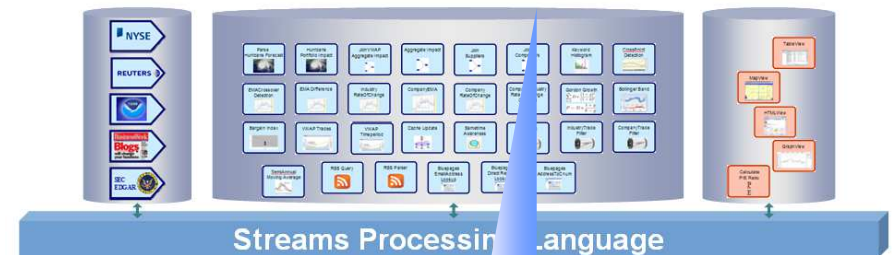


Sin una Plataforma Big Data
Usted Codifica...

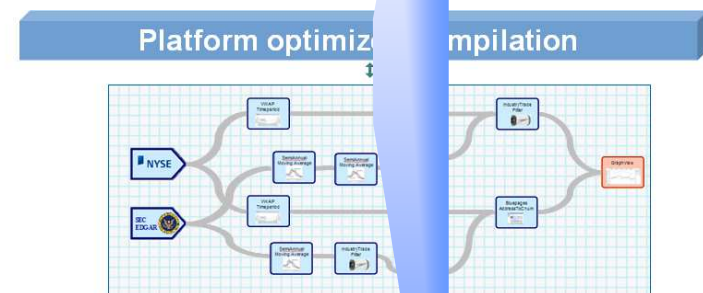


Plataforma IBM Big Data

Más de 100 aplicaciones de muestra y cajas de herramientas con cajas de herramientas diseñadas para la industria ¡con más de 300 funciones y operadores!

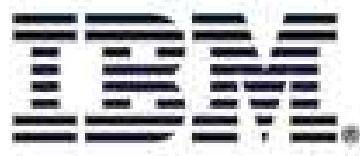


Los flujos de datos proveen servicios de desarrollo, despliegue, tiempo de ejecución e infraestructura



“Los desarrolladores de TerraEchos pueden crear aplicaciones 45% más rápido por la agilidad del Lenguaje de Procesamiento de Flujos...”

– Alex Philip, CEO y Presidente



Aceleradores para una relación de tiempo y valor más veloz!



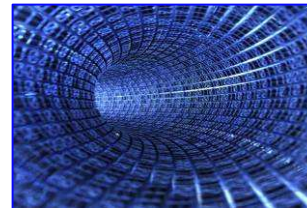
Telco



Finanzas



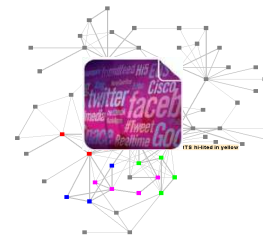
Energía más inteligentes



Minería de datos



Transporte Público



Analítica de Opinión

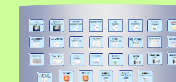
. . . hay más en camino.



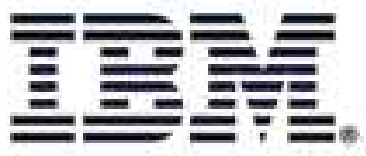
Más de 100
aplicaciones
de muestra



Cajas de
Herramientas
Definidas por el
Usuario



Cajas de
Herramientas
Estándar



IBM's Big Data Business Partner Ecosystem



boardreader™
Connecting Communities through Search



EasyAsk.
Search That Means Business

lucid
IMAGINATION

infogroup®



tresata
make • data • work

TerraEchos

syncsort

ALOGIX
SYSTEMS GROUP

FUZZY LOGIX

Platform Computing

Atigeo™
Compassionate Technology

CADOW
SOFTWARE

GOGRID

TOSHIBA
Leading Innovation >>>

Sharpe Engineering Inc.

BROCADE

eZly

SEMA
sistemas

ARISTA

Zettaset



CONCORD

SYSTECH
Systech Solutions, Inc.

Cognizant

WIPRO
Applying Thought

MEPS
REAL TIME, INC.

NP NewVantage Partners

Capgemini
CONSULTING. TECHNOLOGY. OUTSOURCING

Infosys®

lunexa
illuminate your data

HCL

LIGHTHOUSE
COMPUTER SERVICES

Perficient®

ESTUATE

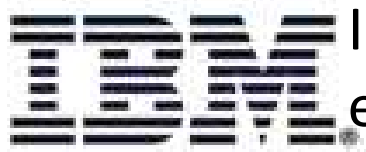
btrg

ONE POINT
SOLUTIONS

TAMGROUP

WaveStrong
Information Security Professionals

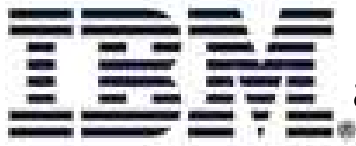




Inicie el viaje Big Data: Elementos comunes para convertirse en una empresa inteligente explotando información



Uso de principios de smarter planet



El Big Data BVA permite a su organización derivar beneficios de analizar grandes volúmenes de datos y aplicarlo en la toma de decisiones para mejorar los resultados de negocio

CÓMO EMPEZAR

Nuestro Big Data BVA ofrece:

- Explorar los flujos de datos que se generan pero que casi nunca se analizan
- Comprender qué soluciones existentes pueden mejorarse y nuevas perspectivas y oportunidades que se crean combinando datos estructurados y no estructurados
- Enfocar un punto / área específico de mejora
- Mejorar el modelado analítico utilizando sus propios datos
- Recomendaciones y próximos pasos

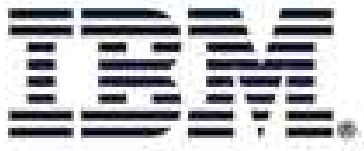


VENTAJAS

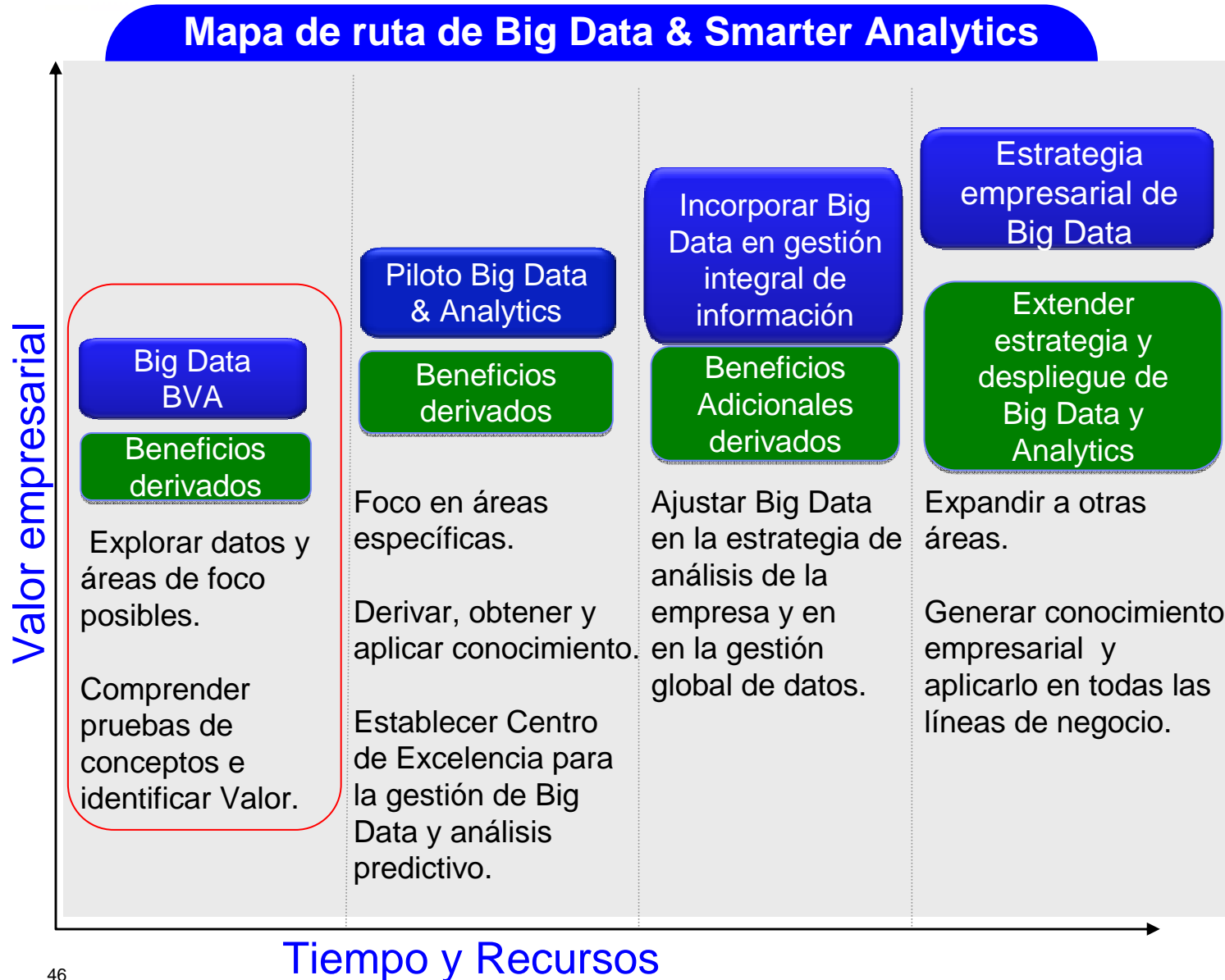
- **Ganar conocimiento con datos mejorados** - entrega rápida con una inversión reducida al inicio
- **Obtener una perspectiva fresca** - oportunidad para evaluar la gestión de datos de IBM, capacidad analítica avanzada y conocimiento de líder de mercado
- **Explorar oportunidades potenciales** - abrir puertas a oportunidades para entregar un valor significativo y ventajas competitivas al negocio

PRÓXIMOS PASOS

- Piloto / Rollout
- Proyectos específicos de Big Data y de análisis predictivo identificados durante el BVA
- Proyectos Analíticos más efectivos
- Centro de excelencia Big Data
- Fundaciones de Gestión de Información

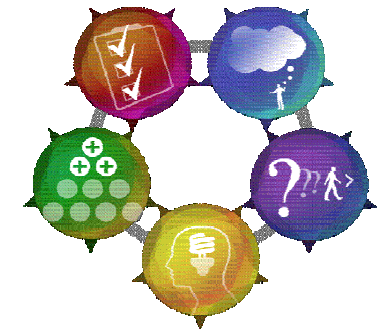


Mapa de ruta de Big Data - Explorar con un BVA

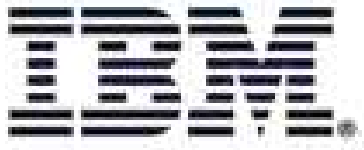


Acciones sugeridas:

1. Acordar compromiso
2. Asignar recursos para participar con IBM
3. Acordar alcance y recopilar los datos relevantes



Metodología Big Data



BigInsights en la Nube

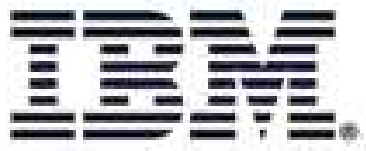
Haciendo el Aprendizaje de Hadoop Fácil y Entretenido

- El dictado **flexible** online permite aprender @su casa y @su paso
- Cursos **gratuitos**, materiales de estudio **gratuitos**
- Entorno de pruebas para ejercicios basado en la nube – **sin configuración**
- Más de 43.000 estudiantes registrados

<http://www.bigdatauniversity.com>

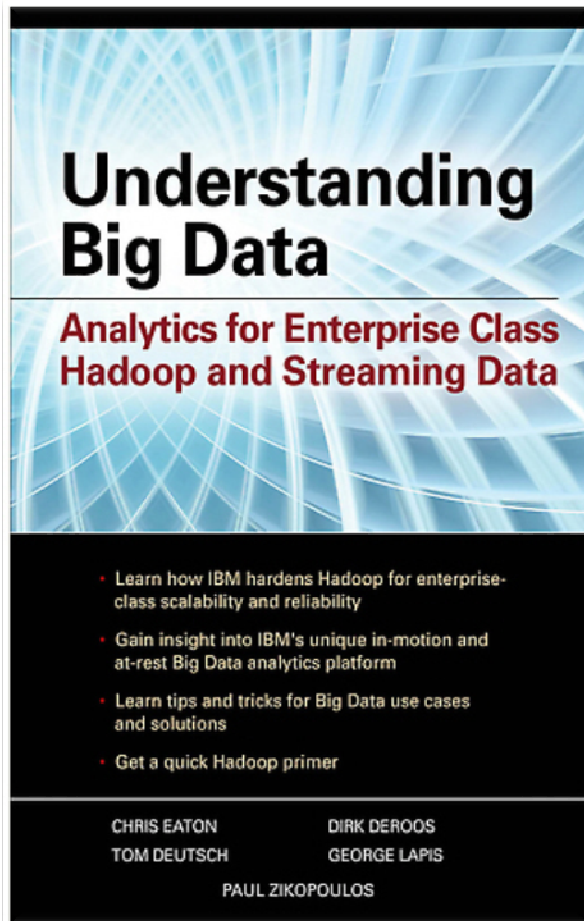
The screenshot displays the BigData University website interface. At the top, the logo "BigData University BETA" is visible, along with a language selector set to "English" and buttons for "login" and "sign up". A navigation bar includes links for HOME, LEARN, DOWNLOAD, RESOURCES, JOBS, and LEARN Hadoop. The main content area features several promotional banners: "Hadoop Fundamentals" with a video player, "Why register?" listing benefits like "Easy and Affordable" and "Latest industry trends", and "Study Made Easy!" with a "sign up now" button. A "Student Testimonials" section highlights a testimonial from Balázs (USA) about the quality of the training material. At the bottom, a footer contains links for "about us", "legal", "contact", and "bug reports".

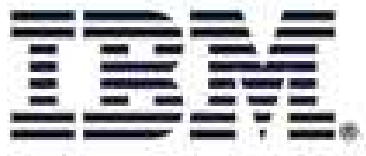




¿Dónde Encontrar Más Información?

<http://www.bigdatauniversity.com>

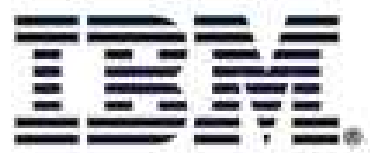


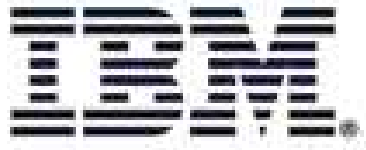


Estar actualizado en el Big Data Hub



www.ibmbigdatahub.com





¡ Muchas
Gracias !

Leonardo González Barceló
lgonzal@cl.ibm.com

