BIG DATA AND THE ENTERPRISE
A PERSPECTIVE FROM FEATURED GARTNER ANALYST DONALD FEINBERG

Donald Feinberg, Vice President and Distinguished Analyst, Gartner Inc.

Mark Simmonds, IBM Enterprise Architect and Senior Product Marketing Manager, IBM Software Group

July 11, 2013

Featured Analyst Firm
Gartner.

Agenda

- Big data in the enterprise

- An analyst's perspective

- Making big data a reality - IBM System z Information Management solutions

- Summary and Call to Action

- Q&A

On a Smarter Planet, Unprecedented Changes are Occurring

- Business models under constant **pressure**

- Customers are more **demanding** and connected

- Great **relationships** trump great products

Big Data is All Data and All Paradigms

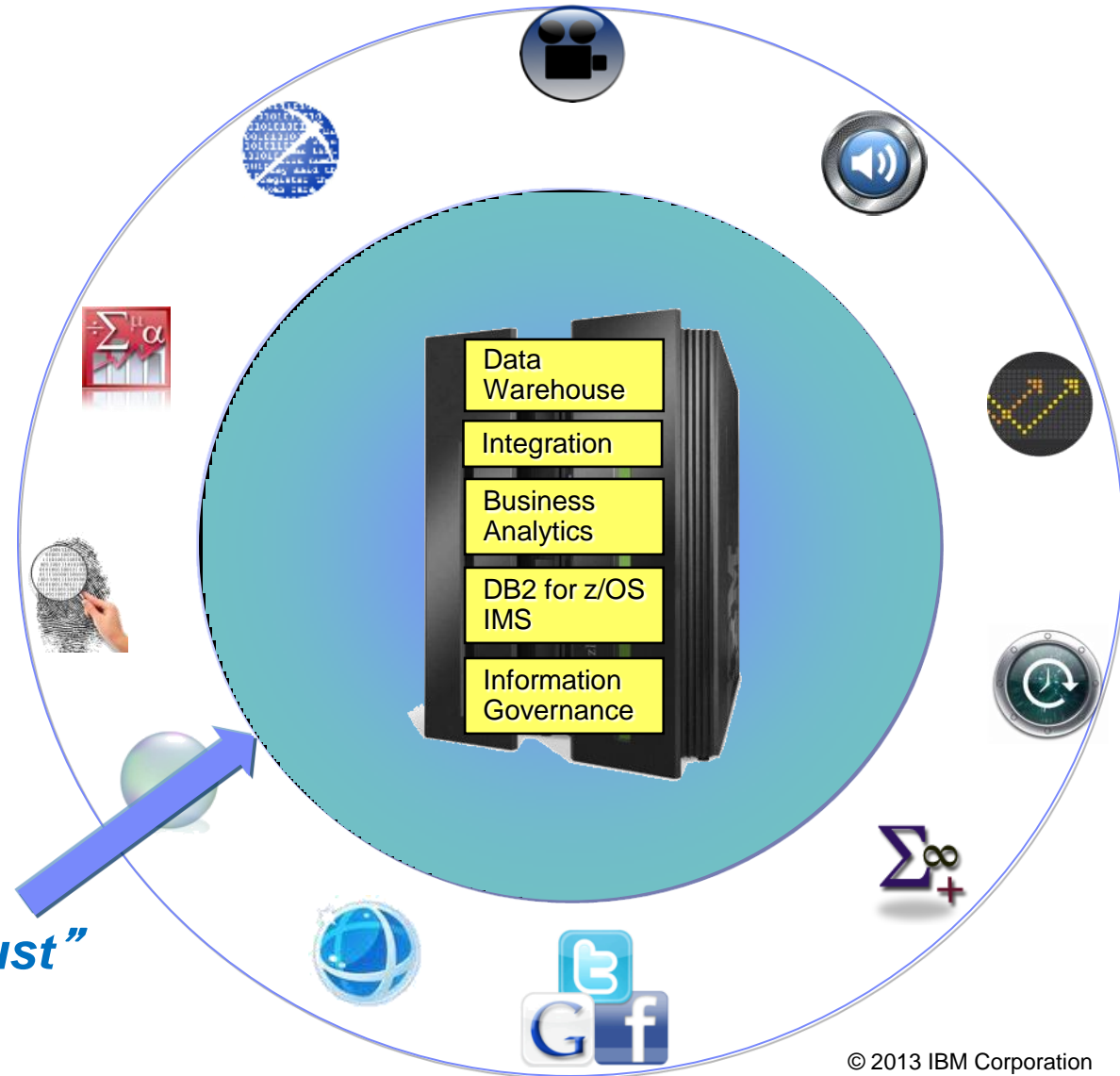| **Transactional & Application Data**	**Machine Data**	**Social Data**	**Enterprise Content**
• Volume	• Velocity	• Variety	• Variety
• Structured	• Structured	• Unstructured	• Unstructured
• Throughput	• Ingestion	• Veracity	• Volume

"We have for the first time an economy based on a key resource [Information] that is not only renewable, but self-generating…..
….Running out of it is not a problem, *but drowning in it is.*"
– John Naisbitt

Majority of today's analytics based on relational / structured data

- Analytics and decision engines reside where the DWH / transaction data is

- "Noise" (veracity) surrounds the core business data

  - Social Media, emails, docs, telemetry, voice, video, content

- What data are you prepared to **_TRUST?_**

- Where do you put your trusted Data?

"Circle of trust"

Data Warehouse

Integration

Business Analytics

DB2 for z/OS IMS

Information Governance

Demand for _differently_ structured data to be seamlessly integrated, to augment analytics / decisions

- Analytics and decision engines reside where the DWH / transaction data is

- "Noise" (veractity) surrounds the core business data

  - Social Media, emails, docs, telemetry, voice, video, content

- Expanding our insights – getting closer to the "truth"

  - Lower risk and cost

  - Increased profitability

"Circle of trust" widens

Data Warehouse
Integration
Business Analytics
DB2 for z/OS IMS
Information Governance

zEnterprise and Big Data
A significant data source for today's business critical analytics

- **Data that originates and/or resides on zEnterprise**
  - 2/3 of business transactions for U.S. retail banks
  - 80% of world's corporate data

- **Businesses that run on zEnterprise**
  - 66 of the top 66 worldwide banks
  - 24 of the top 25 U.S. retailers
  - 10 of the top 10 global life/ health insurance providers

- **The downtime of an application running on zEnterprise = apprx 5 minutes per yr**

- **1,300+ ISVs run zEnterprise today**
  - More than 275 of these selling over 800 applications on Linux

Agenda

- Big data in the enterprise

- *An analyst's perspective*

- Making big data a reality - IBM System z Information Management solutions
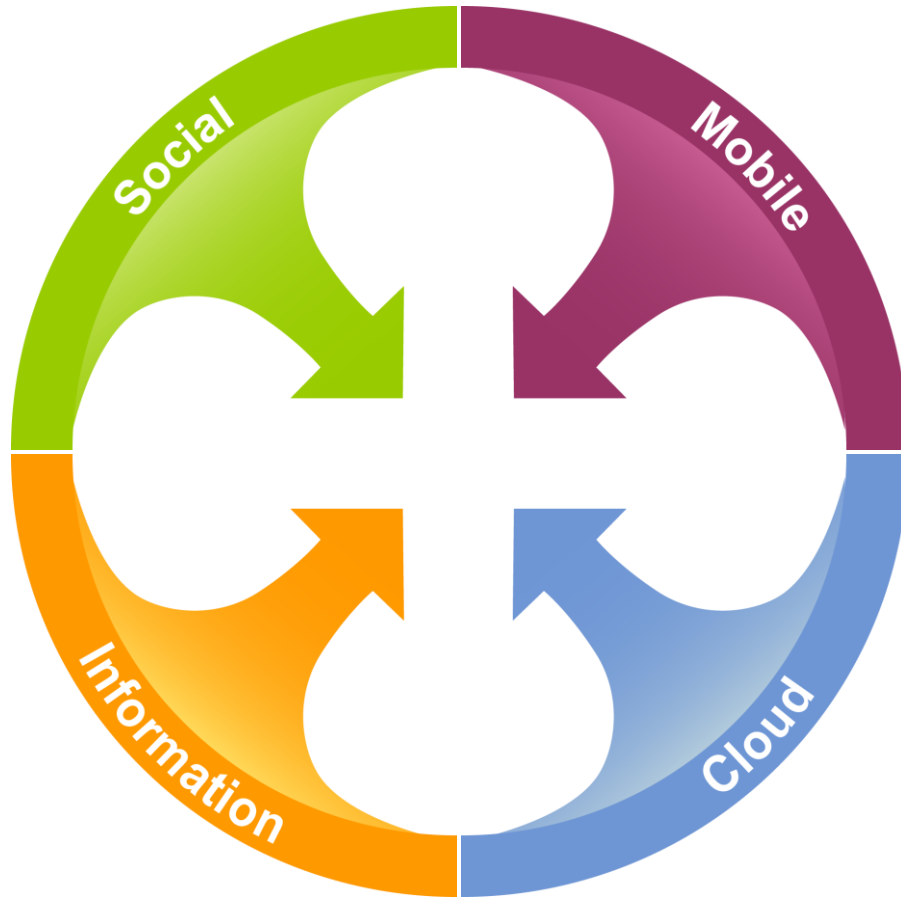
- Summary and Call to Action

- Q&A

Information 2020: Beyond Big Data

Donald Feinberg

VP & Distinguished Analyst

Gartner, Inc.

Gartner.

The NEXUS … Biggest Information Technology Inflection Point in 35 Years!

- What happened the last time?
  - From accounting and finance to operations management and market analysis
- What were the answers?
  - Client/server
  - Relational data
  - Operating systems
  - Server technology
  - Business Intelligence & DW
  - Not "unstructured" but different structure

This time the challenges include the variety and velocity of data in social systems and mobility platforms.
The answers are Information Management and Cloud.

Excitement   **Confrontation**   **Fear**

Emerging Analytic Linkages

Analysis

Descriptive	Diagnostic	Predictive	Prescriptive
What Is Happening?	Why Did It Happen?	What Is Likely to Happen?	What Should I Do About It?

Structured   **Hybrid**   **Content**
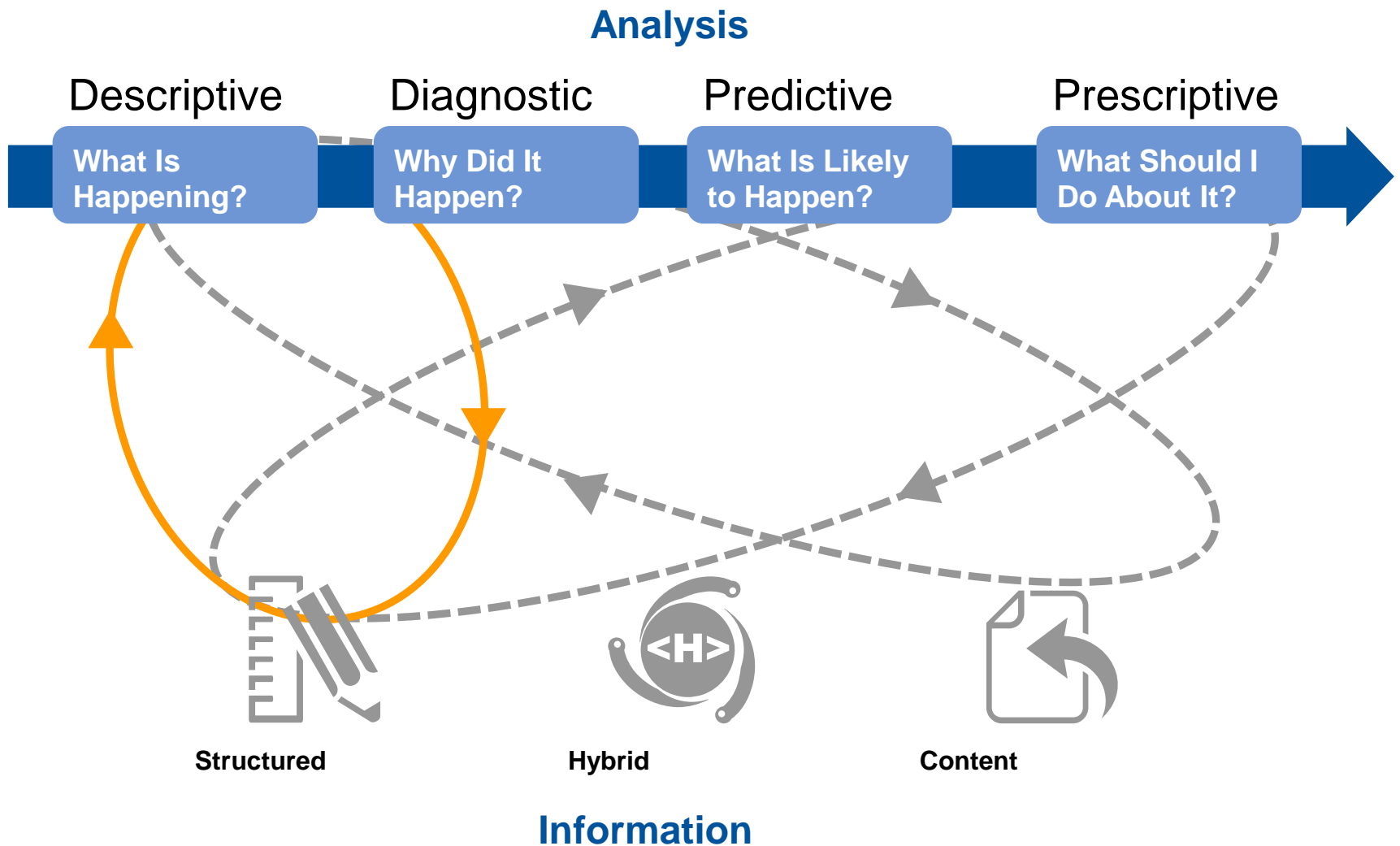
Information

The New World of Information: Fragmenting Technical Landscape

STORE

File, in-memory, graph, column … ← EDW/CMS

MANAGE

Centralized → Federated

SHARE/COLLAB.

Web → Any device

ACCESS

OT, social, mobile, syndicated, open ← Bus app/OLTP

Structured → Content Hybrid

ANALYZE

Fragmented/ Focused → Asset leverage /Distributed

GOVERN

Information Driving Business Performance — Cause or Effect?

Through 2016, organizations integrating high-value, diverse, new information types and sources into a coherent information management infrastructure, will outperform their industry peers financially by <u>more than 20%</u>.

Supporting the SPA:

- It is generally accepted that unstructured data comprises 60% to 80% of all information assets in an organization (e.g., documents, email and images).

- Adoption rates for text mining and analytics as well as e-discovery are rising. Anecdotally, 5% to 15% of analytics already utilize some form of sentiment or social analysis to augment structured data analytics.

- The total worldwide volume of data is growing at 59% per year, with the number of files growing at 88% per year.

- Organizations that focus on risk management, measure/monitor business performance metrics and use integrated information in their business insights ("Value Integrators," IBM) exhibit >30% ROIC, >49% CAGR and orders of magnitude superior EBITDA relative to their peers.*

Alternate position to the SPA:

- Innovative business processes and superior market awareness are the driving differentiators of leading companies, and integrated information assets are merely a byproduct of advanced leadership in other areas of organizations.

- Unstructured data yields a lower value-to-noise ratio than structured data.

- Effective analysis of information permits lagging organizations to close the gap toward leaders, but will continue to be overwhelmed by superior senior leadership skills.

- It is the analysis of information that leads to high levels of performance, and integration alone is ineffective.

*Source: "The New Value Integrator," Insights from the Global Chief Financial Offers Study. IBM, 2010.

Treat Information as an Asset

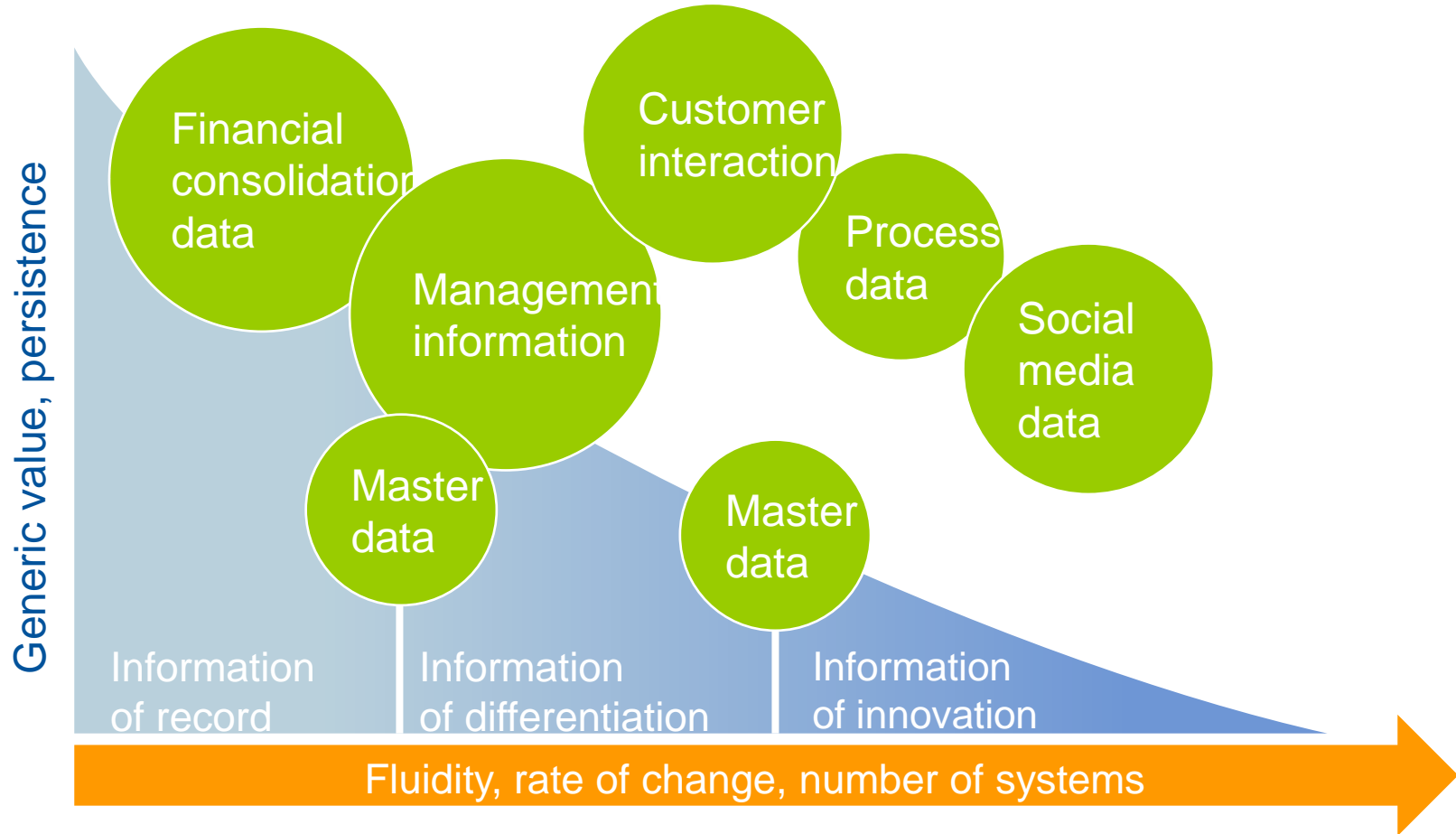
Gartner

Traditional IM

- Requirements based
- Top-down design
- Integration and reuse
- Technology consolidation
- World of DW and ECM
- Competence centers
- Better decisions
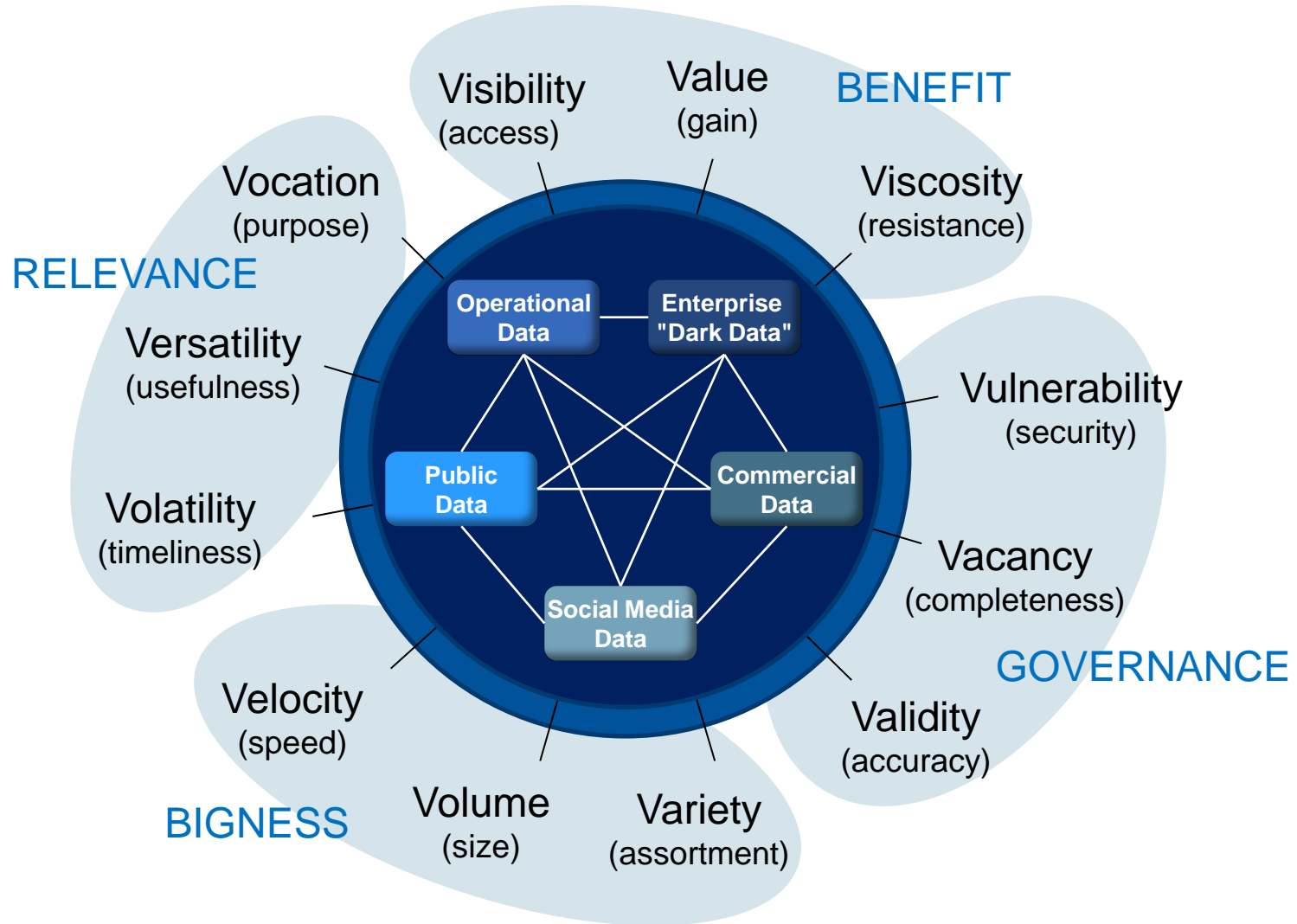- Enterprise

"Big Data" Style

- Opportunity oriented
- Bottom-up experimentation
- Immediate use
- Tool proliferation
- "World of Hadoop"
- Hackathons
- Better business
- Marketing (+)

Gartner

Not All Information Requires a Big Data Approach

Beyond Big Data and Beyond Gartner's Original "3Vs"

Visibility (access)

Value (gain)

BENEFIT

Vocation (purpose)

RELEVANCE

Viscosity (resistance)

Versatility (usefulness)

Vulnerability (security)

Volatility (timeliness)

Operational Data

Enterprise "Dark Data"

Public Data

Commercial Data

Social Media Data

Vacancy (completeness)

GOVERNANCE

Velocity (speed)

Validity (accuracy)

BIGNESS

Volume (size)

Variety (assortment)

Gartner.

Liberating Your Information With the Logical Data Warehouse

Structured　　　**Content**　　　**Hybrid**

Specialized Capabilities

Information Semantic Styles

Manage

Common Capabilities

| Describe | Organize | Integrate | Share | Govern | Implement |

Social　　　Transactional Data　　　Image　　　Text　　　Mobile

Documents　　　IT/OT　　　Audio　　　Video　　　Search Engine

Externalizing Capabilities Leads to Agnostic Information Infrastructure.

Recommendations

- ✓ Prepare for the disruption that will be caused by the Nexus of Forces

- ✓ Research and begin pilot projects using the Logical Data Warehouse

- ✓ Do not jump into Big Data at the sacrifice of management of the other data in the organization

- ✓ Remember, there is more to Big Data beyond Volume, Velocity and Variety

Information 2020: Beyond Big Data

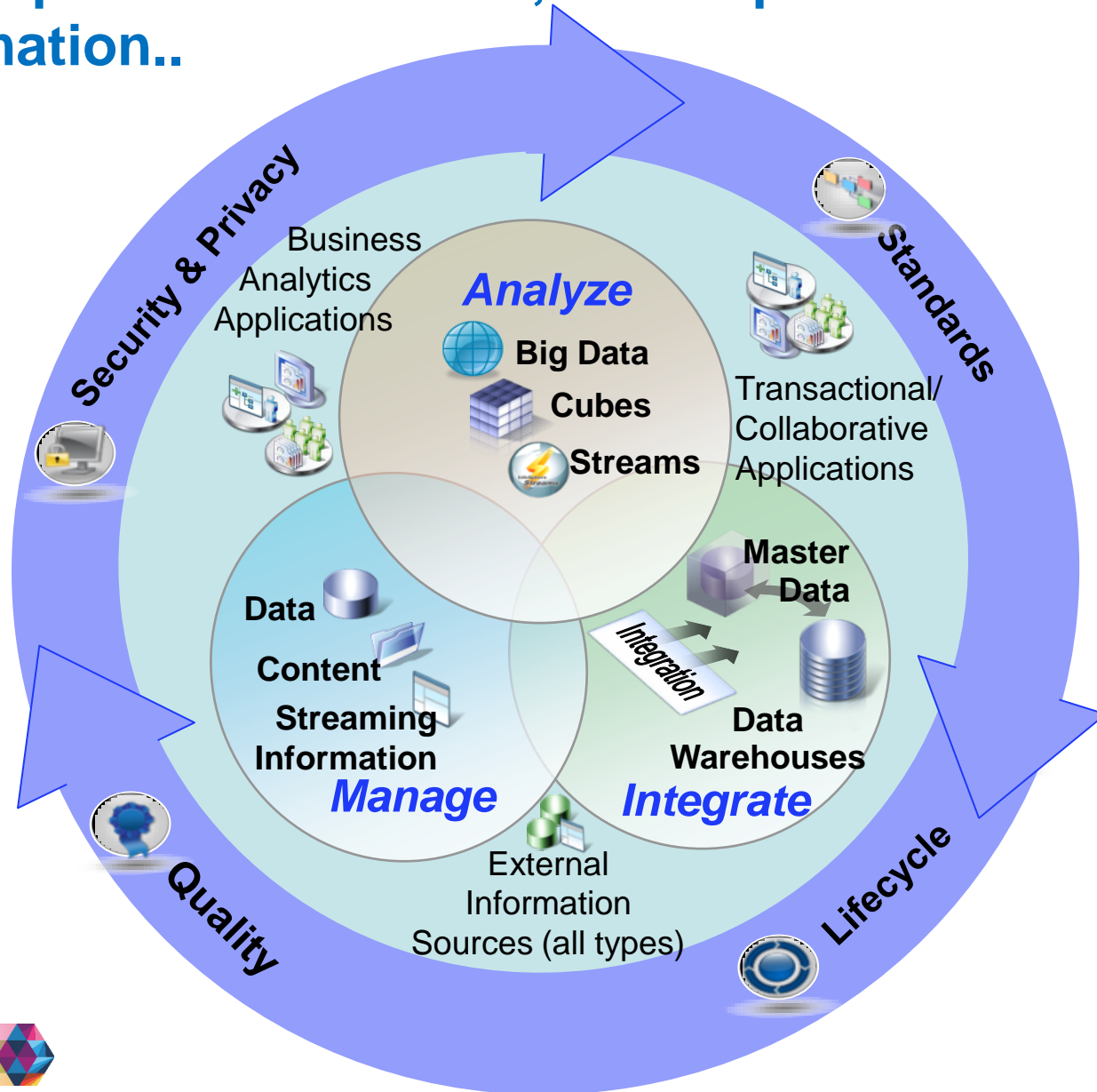Donald Feinberg

VP & Distinguished Analyst

Gartner, Inc.

Gartner®

Agenda

- Big data in the enterprise

- An analyst's perspective

- *Making big data a reality - IBM System z Information Management solutions*
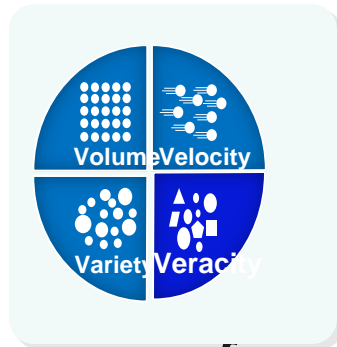
- Summary and Call to Action

- Q&A

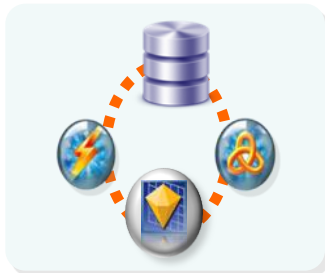# The Enterprise : A federated, interdependent eco system of information..

Forward Thinking Organizations are creating value from Big Data

The power of Data coming together…

…to deliver Improved Business Outcomes

VolumeVelocity

VarietyVeracity

…with the power of Technology…

1. **Enrich your information base**
   with Big Data Exploration

2. **Improve customer interaction**
   with Enhanced 360º View of the Customer

3. **Optimize operations**
   with Operations Analysis

4. **Gain IT efficiency and scale**
   with Data Warehouse Augmentation

5. **Prevent crime**
   with Security and Intelligence Extension

Implementing a Mission Critical Big Data Application

Big Data Exploration
Find, visualize, understand all big data to improve decision making

GPS & sensor information volumes exceeded the capabilities of the existing system. It was redesigned as an enterprise mission critical application using DB2 for z/OS and System z data sharing to now provide the availability and scalability to meet the current and future requirements for this solution.

JOHN DEERE

IBM

What sets zEnterprise apart for Big Data & Analytics Infrastructure

Timely, accurate and secure information	Superior availability, scalability and performance	Reduced costs and complexity	Rapid deployment and expansion
• **Co-location** of data warehousing, business analytics, transactional data • **Reduced data movement** • **Lower latency** and **near real time** data • **Rapid acceleration** of complex queries • High **security** (EAL5+)	• High **availability** (99.999%) • Performs at **100% capacity** • **Prioritization** of critical queries & workloads • Integrated **disaster recovery**	• **Processors, disk, memory added dynamically** without outage • **Pre-install then activate** as needed • **Flexible deployment options**	• **Centralized, scalable** infrastructure • **Virtualization** • Start with your **final architecture**

Data Warehouse – storing multi structured data for greater insight.

Multi-Structured Data

Production Data

Low Latency Copy

Data Warehouse System z

DB2 Analytics Accelerator

Transactional Information

Operational Analytics

Operational Applications

Delivering insights from Multi-Structured data pervasively across the organization embedded in the operational applications
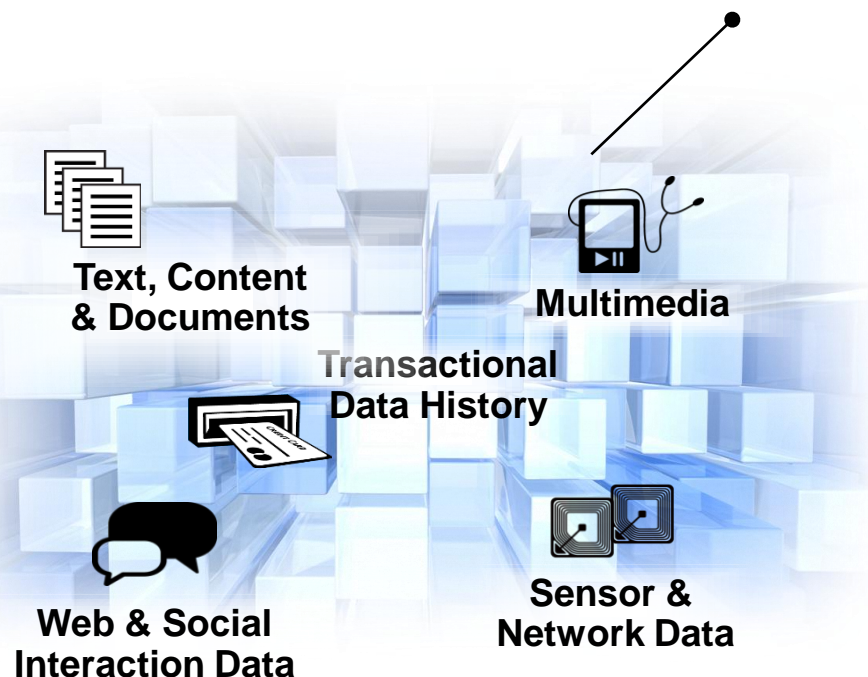
Benefits

- Deliver new insights from multi-structured data such as sensor, social, and clickstream to make fact-based decisions
- Combine multi-structured data with historical data warehouse information to increase understanding
- Provide analytic information at the point of decision enabling fact-based decisions
- Pervasively enable decision makers and other end users across the organization
- Accelerate long running DB2 for z/OS queries from minutes to seconds for greater business value with Analytics Accelerator.

ACCESS, COMBINE & MANAGE a relevant mix of information

IBM Analytics on zEnterprise enables you to blend different sources and integrate them together

Text, Content & Documents

Multimedia

Transactional Data History

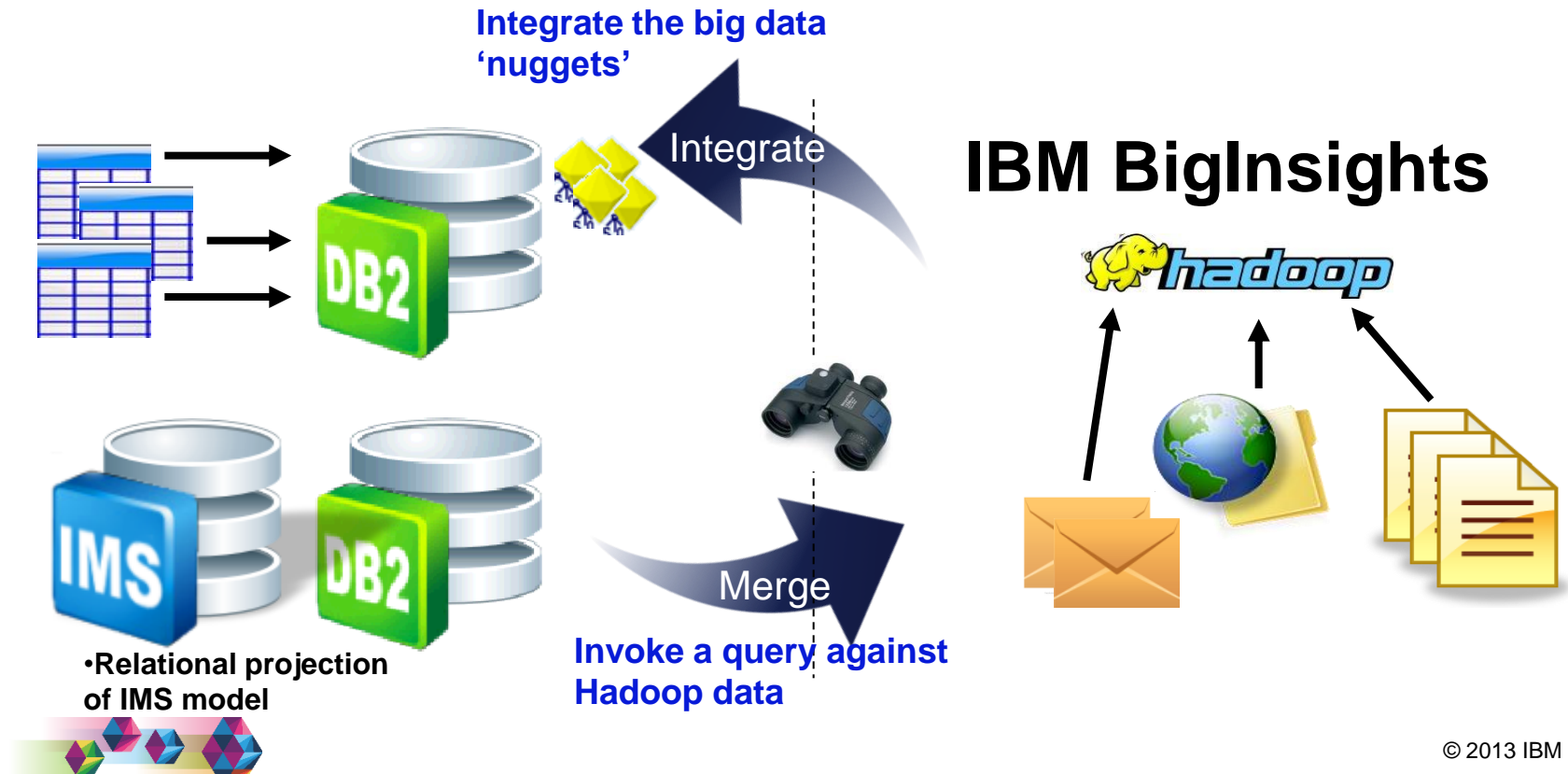**Sensor & Network Data**

Web & Social Interaction Data

- Access big data sources with a single analytics solution

- Integrate big data results with structured enterprise data without the requirement to move source data off of zEnterprise for faster analysis

- Accelerate and cost effectively store large volumes of data, simply and cost competitively for more timely analysis

- Maintains the security and control of your most important corporate asset – your data

Enhancing Big Data Analytics with IMS and DB2 for z/OS

▪ **Much of the world's operational data resides on z/OS**	▪ Unstructured data sources are growing fast

1. Merge this data with trusted OLTP data from zEnterprise data sources
2. Integrate this data so that insights from Big Data sources can drive business actions
- IMS & DB2 - connectors & DB capability to allow BigInsights to easily & efficiently access each data source
- DB2 - connectors & DB capability to allow DB2 apps to easily and efficiently access Hadoop data sources

Integrate the big data 'nuggets'

Integrate

IBM BigInsights

Merge

•**Relational projection of IMS model**

Invoke a query against Hadoop data

DB2 11 Major Themes
Building on the Success of DB2 10

- Performance Improvements
  - Improving efficiency, reducing costs, no application change
  - 0-5% for OLTP, 5-15% for update intensive batch
  - 5-20% for query workloads
  - Exploitation of new zEC12 hardware features
- Continuous Availability Features
  - Improved autonomics which reduces costs and improves a
  - Making online changes without affecting applications
  - Online REORG improvements, less disruption
- Enhanced business analytics
  - Faster, more efficient performance for query workloads
  - SQL improvements and IDAA enhancements
  - Temporal and SQLPL enhancements
  - Transparent archiving
- Simpler, faster DB2 version upgrades
  - No application changes required for DB2 upgrade
  - Access path stability improvements
  - Product stability: support pre GA customer production

**IBM announced DB2 11
Early Support Program**
→ Learn more

DB2 11 Support for Big Data

- Goal: integrate DB2 for z/OS with IBM Hadoop based BigInsights Bigdata platform
  - Enabling traditional applications on DB2 z/OS to access Big Data analytics.
- Analytic jobs can be specified using JSON Query Language (Jaql)
  - Submitted to BigInsights
  - Results stored in Hadoop Distributed File System (HDFS).
- A table UDF (HDFS_READ) reads the Bigdata analytic result from HDFS, for subsequent use in an SQL query.
- Must have a variable shape of HDFS_READ output table
  - DB2 11 supports generic table UDF, enabling this function

TECHNOLOGY
Deliver dramatically faster analysis

IBM DB2 Analytics Accelerator

A high performance appliance that integrates Netezza technology with zEnterprise technology, to deliver dramatically faster business analysis

- Speeds up complex queries up to 2000x

- Lowers the cost of long term storage

- Minimizes data latency

- Improves security and reduces risk

- Complements existing investments

Big Data & Analytics on IBM zEnterprise

Bringing it all together > Driving greater insight

IBM Big Data Platform

| Stream * Computing | Hadoop Systems | Data Warehouse |

Other Big Data Sources

cloudera · SAP · ORACLE · Hortonworks · TERADATA · Greenplum

Enterprise Data Sources *

Business Intelligence, Predictive Analytics

Data Warehousing

Business System / OLTP

IMS   DB2

A Single Solution to:

- ACCESS, COMBINE & MANAGE a relevant mix of information
- TIMELY ACCESS for more accurate answers

IBM

Big Data Innovation
with IBM zEnteprise

Operations Analysis

Analyze a variety of machine data for improved business results

Banco do Brasil purchases the **largest ever** DB2 Analytics Accelerator solution to drive customer insight from operational data. The 120-way system can hold 1.28 Petabytes of data. Queries that previously took *11 hours to run now complete in 26 seconds*, over 1500 times faster!

Banca Carige chooses System z to provide real time analytics as part of their Big Data client solution
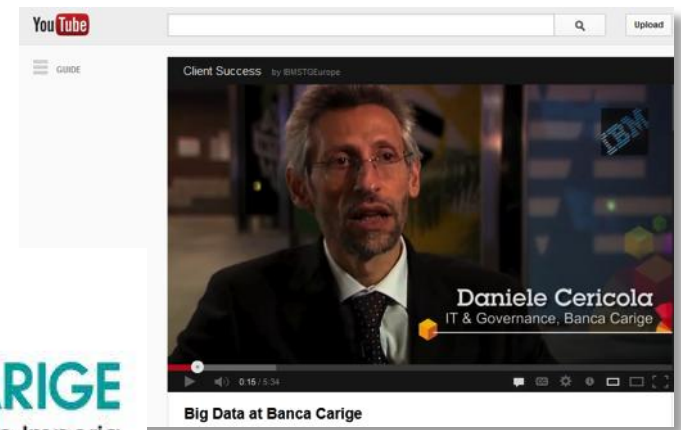
Enterprise wide information Governance solutions multiple platforms, data stores and data types

Business Analytics

Master Data Management

Customer Retention

Data Warehousing

Risk management

Up-sell and Cross-sell

Cost reduction

Business Compliance

SYBASE · MySQL · IBM Informix · NETEZZA · PostgreSQL · Microsoft Office SharePoint · IBM VSAM · Microsoft SQLServer · hadoop · IBM DB2 · ORACLE · TERADATA · IBM IMS

1 *Information Quality*

2 *Lifecycle Management*

3 *Information Protection*

Information Governance underpins and is foundational to your Information Management projects

Agenda

- Big data in the enterprise

- An analyst's perspective

- Making big data a reality - IBM System z Information Management solutions

- *Summary and Call to Action*

- Q&A

IBM understands your big data requirements

IBM understands all kinds of data

- **Game-Changing Innovation** – such as Watson, IBM DB2 Accelerator for z/OS, streaming analytics, Hadoop implementations , expert integrated systems; 20 years of patent leadership, best of breed databases / data warehouses – *DB2 for z/OS, IMS….*
- **Business-Ready Capabilities** – big data and analytics capabilities, integrated and hardened for serious use, with mission and business critical qualities of service unmatched in the industry

IBM knows how to turn data into value

- **Client Expertise** – deep industry know-how and solutions with global reach
- **Strong Ecosystem** – growing investment with 360+ business partners & 100+ universities
- **Build on Current Investments** – enhance existing analytics and information infrastructure with unparalleled breadth and depth of new capabilities

IBM continued investments in big data and analytics processing

- **$16B+ in Acquisitions** – coupled with game-changing innovations
- **IBM zEnterprise -** The platform with the highest levels of integrity, qualities of service and governance for your business critical applications and processes.
- **Analytics Solution Centers** – visited by 4000+ organizations accessing global expertise

Take Action Now!

Next steps:

- **For additional information including whitepapers and demos, please visit:**
  - IBM big data for z web site
  - Analytics on zEnterprise
  - Information Governane
- **Education**
  - Free online education at bigdatauniversity.com
  - 70,000+ registered students
- **Community**
  - **Join Big data and World of DB2 for z/OS on LinkedIn**
- **Further developments:**
  - Watch for future webcast and announcements

- **Wanting to experiment on a big data integration project but not sure where to start?** *Partner with IBM Silicon Valley and other laboratories nearest you..*

Q&A

Donald Feinberg, Vice President and Distinguished Analyst, Gartner Inc.

Mark Simmonds, IBM Enterprise Architect and Senior Product Marketing Manager, IBM Software Group

Program Note

This webcast is sponsored by IBM. Any editorial supplied by IBM is independent of Gartner analysis. All Gartner research is © 2013 by Gartner, Inc. and/or its Affiliates. All rights reserved.

All Gartner materials are used by IBM with Gartner's permission and in no way does the use or publication of Gartner research indicate Gartner's endorsement of IBM products and/or strategies. Any other reproduction of this webcast in any form without prior written permission is forbidden. The information contained herein has been obtained from sources believed to be reliable.

Gartner disclaims all warranties as to the accuracy, completeness or adequacy of such information. Gartner shall have no liability for errors, omissions or inadequacies in the information contained herein or for interpretations thereof. The opinions expressed herein are subject to change without notice.

Thank You

To learn more, please visit:

www.ibm.com