Why Big Data? Why Now?
Information is at the Center of a New Wave of Opportunity...

44x
as much Data and Content
Over Coming Decade

2020
35 zettabytes

2009
800,000 petabytes

80%
Of world’s data is unstructured

... And Organizations Need Deeper Insights

1 in 3
Business leaders frequently make decisions based on information they don’t trust, or don’t have

1 in 2
Business leaders say they don’t have access to the information they need to do their jobs

83%
of CIOs cited “Business intelligence and analytics” as part of their visionary plans to enhance competitiveness

60%
of CEOs need to do a better job capturing and understanding information rapidly in order to make swift business decisions
The Challenge: Bring Together a Large Volume and Variety of Data to Find New Insights

- Multi-channel customer sentiment and experience analysis
- Detect life-threatening conditions at hospitals in time to intervene
- Predict weather patterns to plan optimal wind turbine usage, and optimize capital expenditure on asset placement
- Make risk decisions based on real-time transactional data
- Identify criminals and threats from disparate video, audio, and data feeds
Extracting insight from an immense volume, variety and velocity of data, in context, beyond what was previously possible.

**Variety:** Manage the complexity of multiple relational and non-relational data types and schemas

**Velocity:** Streaming data and large volume data movement

**Volume:** Scale from terabytes to zettabytes
The Traditional Approach: *Business Requirements Drive Solution Design*

- **Business Defines Requirements** – What Questions Should we Ask?
- **IT Designs a Solution** with a set structure and functionality
- **Business executes queries to answer questions over and over**
- **New requirements require redesign and rebuild**
The Traditional Approach: **Business Requirements Drive Solution Design**

**Business Defines Requirements – What Questions Should we Ask?**

- Business executes queries to answer questions over and over
- New requirements require redesign and rebuild
- IT Designs a Solution with a set structure and functionality

**Well-Suited To:**
- High value, structured data
- Repeated operations and processes (e.g. transactions, reports, BI, etc.)
- Relatively stable sources
- Well-understood requirements

**Stretched By:**
- Highly variable data and content
- Iterative, exploratory analysis (e.g. scientific research, behavioral modeling, etc.)
- Volatile sources
- Ill-defined questions and changing requirements
The Big Data Approach: *Information Sources Drive Creative Discovery*

- Business and IT Identify Information Sources Available
- IT Delivers a Platform that enables creative exploration of all available data and content
- New insights drive integration to traditional technology
- Business determines what questions to ask by exploring the data and relationships
Big Data Shouldn’t Be a Silo
Must be an integrated part of your enterprise information architecture
Merging the Traditional and Big Data Approaches

**Traditional Approach**
*Structured & Repeatable Analysis*

**Business Users**
Determine what question to ask

**IT**
Structures the data to answer that question

**Big Data Approach**
*Iterative & Exploratory Analysis*

**IT**
Delivers a platform to enable creative discovery

**Business**
Explores what questions could be asked

- Monthly sales reports
- Profitability analysis
- Customer surveys
- Brand sentiment
- Product strategy
- Maximum asset utilization
The Solution – IBM’s Big Data Platform
Bring together any data source, at any velocity, to generate insight

- Analyzing a variety of data at enormous volumes
- Insights on streaming data
- Large volume structured data analysis

IBM Big Data Platform
- Variety
- Velocity
- Volume
Optimize capital investments based on 6 Petabytes of information

- Model the weather to optimize placement of turbines, maximizing power generation and longevity
- Build models to cover forecasting and real-time operation of power generation units
- Incorporate 6 PB of structured and semi-structured information flows
A Platform Approach Address Enterprise Client Needs

**Enterprise Client Needs**

- Enable creativity and agility
- Focus on outcomes
- Reduce and manage complexity
- Lower development and integration costs

**Big Data Platform Delivers**

- Platform for V³
- Analytics for V³
- Ease of Use for Developers/Users
- Enterprise Class
- Extensive Integration
IBM Watson is a breakthrough in analytic innovation, but it is only successful because of the quality of the information from which it is working.
Big Data and Watson

**Big Data technology is used to build Watson’s knowledge base**

Watson uses the Apache Hadoop open framework to distribute the workload for loading information into memory.

**Watson can consume insights from Big Data for advanced analysis**

Watson’s Memory

- POS Data
- CRM Data
- Social Media

Approx. 200M pages of text (To compete on *Jeopardy!*)

- Distilled Insight
  - Spending habits
  - Social relationships
  - Buying trends

Advanced search and analysis
Imagine the Possibilities … in a World with No Limits

**Information from Everywhere**
- Data & content
- Apps, web & sensors
- At rest & in motion
- Integrated & federated

**Radical Flexibility**
- Virtualization at every level
- Automated administration
- Easy-to-use analytics

**Extreme Scalability**
- “Big data” analytics
- Real-time stream processing
- Efficient parallelism
- Workload-optimized