Solutions Big Data IBM
The BIG Data Challenge

• Manage and benefit from massive and growing amounts of data
• Handle uncertainty around format variability and velocity of data
• Handle unstructured data
• Exploit *BIG Data* in a timely and cost effective fashion
By 2015, 80% of all available data will be uncertain.

By 2015 the number of networked devices will be double the entire global population. All sensor data has uncertainty.

The total number of social media accounts exceeds the entire global population. This data is highly uncertain in both its expression and content.

Data quality solutions exist for enterprise data like customer, product, and address data, but this is only a fraction of the total enterprise data.

Multiple sources: IDC, Cisco
Uncertainty arises from many sources

**Process Uncertainty**
Processes contain “randomness”
- Uncertain travel times
- Semiconductor yield

**Data Uncertainty**
Data input is uncertain
- Intended Spelling → Text Entry → Actual Spelling
- GPS Uncertainty
- Testimony
- {Paris Airport}
- Ambiguity
- {John Smith, Dallas}
- {John Smith, Kansas}
- Rumors
- Conflicting Data

**Model Uncertainty**
All modeling is approximate
- Fitting a curve to data
- Forecasting a hurricane
(www.noaa.gov)
**The fourth dimension of Big Data: Veracity – handling data in doubt**

<table>
<thead>
<tr>
<th>Volume</th>
<th>Velocity</th>
<th>Variety</th>
<th>Veracity*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data at Rest</td>
<td>Data in Motion</td>
<td>Data in Many Forms</td>
<td>Data in Doubt</td>
</tr>
<tr>
<td>Scale from terabytes to petabytes (1K TBs) to zettabytes (1B TBs)</td>
<td>Streaming data, milliseconds to seconds to respond</td>
<td>Structured, unstructured, text, multimedia</td>
<td>Uncertainty due to data inconsistency &amp; incompleteness, ambiguities, latency, deception, model approximations</td>
</tr>
</tbody>
</table>

* Truthfulness, accuracy or precision, correctness
Big Data Technologies

Variety
Non-traditional/internet data

Volume

At-Rest Data
Traditional Data

In-Motion Data

InfoSphere BigInsights to perform exploratory analytics
InfoSphere Streams to score incoming data against analytic models

InfoSphere Warehouse
Reuse Warehouse Analytic models

Netezza

Hadoop
Marketeer’s objectives

1. Single view of customer
   - Clickstream
   - Transactions
   - Events
   - CRM
   - Support calls
   - All in one place

2. Increased Targeting Precision
   - Clustering
   - Scoring
   - Feature Selection
   - Associations

3. Improved Relevance
   - Personalized message
   - Matching algorithms
   - Matrix computations
   - Single Value Decomp.

4. Higher campaign profitability
   - Forecasting
   - Predictive algorithms
   - Decision trees
   - Linear Regression

| Page 7 |
**Marketeer’s objectives**

1. Single view of customer
2. Increased Targeting Precision
3. Improved Relevance
4. Higher campaign profitability

**Platform Capability**

- **Consolidation**
  - Clickstream
  - Transactions
  - Events
  - CRM
  - Support calls
  - All in one place

- **Segmentation**
  - Clustering
  - Scoring
  - Feature Selection
  - Associations

- **Matching**
  - Personalized message
  - Matching algorithms
  - Matrix computations
  - Single Value Decomp.

- **Optimization**
  - Forecasting
  - Predictive algorithms
  - Decision trees
  - Linear Regression

**Big Data**

- Clustering
- Scoring
- Feature Selection
- Associations

**Complex Analytics**
IBM Offers a Comprehensive Set of Solutions for BIG Data

- Non-Traditional Data
  - InfoSphere Big Insights
  - Streams filters incoming data
  - Streams reuses Warehouse Analytic models

- Traditional Data

- Persistent Data

- In-Motion Data
Un nouveau mode d’exploration des données

**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**
Big Data: integration points

IBM Big Data Solutions

Client and Partner Solutions

Rules / BPM
- iLog & Lombardi

Data Warehouse
- InfoSphere Warehouse

Warehouse Appliances
- IBM & non-IBM

Master Data Mgmt
- InfoSphere MDM

Database
- DB2 & non-IBM

Content Analytics
- ECM

Business Analytics
- Cognos & SPSS

Marketing
- Unica

Data Growth Management
- InfoSphere Optim

© 2012 IBM Corporation
Typical IBM Big Data Solution Use Cases

Unstructured data

Create context (classification, text mining)

Analyze

Semi-structured data

Parse, aggregate

Analyze, report

Structured data

Analyze, report

Active archival
Long running queries
Big Data Use Case
Customer Experience
(Call Center)
Business Value Hypotheses - Summary

- Analyze Customer Survey Data to Advocates vs. Antagonist
- Analyze Customer Complaint Data to Understand Key Customer Issues
- Understand Impact of Social Networks on Customer Behavior & Influence
- Analyze Agent-to-Agent Internal Memos
- Analyze Customer Interaction Notes (i.e. emails, etc.)
- Analyze Policy & Procedures

- Cost Efficiencies
- Performance
- Revenue Lift
- Regulatory Compliance
Addressing the Customer Experience Issues

**Step 1**
CSR Agent asks customer questions to understand situation and customer need.

**Step 2**
CSR Agent researches question by consulting policies & procedures, looking up data, consulting other agents, or transfers the call to another agent.

**Step 3**
CSR Agent spends time clarifying question based on research and providing best possible answer.

**Step 4**
CSR Agent ensures customer is satisfied & probes for other needs before closing inquiry.

---

**Future State Knowledge Management Solution**

**Phase 1**  
Establish a searchable knowledge base database organized by need with structured actions and procedures for CSR to follow.

**Phase 2**  
Watson: Incorporate transactions and interactions into generating a list of potential reasons for call and prioritize potential next best actions.

**Phase 3**  
Watson 2: Automatically provide suggestions for specific customer inquiries, as well as suggestion for customer treatments based on transaction & interaction history.

---

Search Highly Organized Knowledge, Better Predict Customer Needs & Drive Differentiated Experience
Account / Product specifications, Customer Transactions and Interactions, Policies and Procedures are loaded into Big Insights and relevant information context linkages are established in an indexed Knowledge Base (KB). The KB has natural language capabilities to link issues with treatments.

A customer transfers from IVR to an Agent, and gets identified.

A customized UI helps the agent access the Content search engine to research the customer’s issue, locate appropriate treatment and apply it.
Phase 2: Automate Issue Detection & Treatment

1. The Knowledge Base is compiled into an Active Profile
2. Customer transfers out of IVR to the call center and gets identified
3. Agent is provided with prioritized potential issues by the Rules Engine, together with treatments. Where needed the agent can research the Knowledge Base. Treatment application is automated where possible, reducing manual time required.

**Analytics**
- Watson / Big Insights Analytics
- Predictive Engine

**Knowledge Interface**
- Search Engine
- Knowledge Base Search & Retrieve

**Active Profiling**
- Predict Issues
- Prioritize Treatments

**Agent Assist**
- Caller Identified
- Prioritized Potential Issue(s) Presented to Agent
- Issue Selected, or where necessary, researched
- Treatments for Issue are applied through automation, or where necessary, manually
- Cross or Up Sell
- Closure

**Call Center IVR**