Big Data: Essential Elements to a Successful Modernization Strategy

Ashish Verma
Director, Deloitte Consulting Technology Information Management

Deloitte Consulting

#pbls14
Big Data and its Implications for Analytics

Ashish Verma, Director
Deloitte Consulting

October 23, 2014
Analytics and “Big Data”: Nothing new, totally revolutionary
Analytics and Big Data
The current “hype cycle” around Analytics and Big Data is booming, driven by both user expectations and vendor hype.
An on-going paradigm shift is transforming the landscape of companies with data volumes growing exponentially…

With huge amounts of data generated per day coming from a variety of sources capturing growing volumes of transaction detail about customers, suppliers and operations

- 6 billion mobile phones
  - 34% of adults use tablets

- 1.2 billion mobile web subscribers

- 60 billion intelligent devices

- 30 billion of content pieces shared every day
  - 500TB

- 340 million tweets generated every day @Twitter
  - 57% of all Tweets being generated in the US

- 200+ TB stored data in every sector

Changing the status quo of structured information analysis to unstructured and semi structured data known as Big Data Analytics
Which implies that Big Data cannot be ignored…

With the explosive growth of data in web, social media, and within organizations, there is a vast, rich component of underutilized data sources that executives can no longer ignore in determining the potential competitive advantages it can provide.

### Data Growth

- **~81%** in Big Data ecosystem software growth in 2013
- **650%** or greater growth in enterprise data between 2010–2014
- **80%** of new enterprise data between 2010–2014 will be unstructured
- **39.4%** compound annual growth rate in Big Data technology and services in 2015

### Revenue

- **$50 Billion** or more revenue from Big Data by 2020

### Talent

- **4.4M** Big Data jobs will be in demand in 2015
Analytics and data — Big or “small”
Analytics is the application of a broad set of capabilities looking at data in new and interesting ways with an eye towards driving value that you can measure, and should measure in dollars.
Material players deliver within and across value chain components and are expanding into analytics.

Sources: IDC, CIBC, Wikibon, Company websites, Press releases, Annual reports, Deloitte Analysis
Business needs drive IT capability requirements

Process structured to un-structured data, low to high volumes, in near real time

<table>
<thead>
<tr>
<th>Variety</th>
<th>Volume</th>
<th>Velocity</th>
<th>Technology</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traditional EIM</td>
<td></td>
<td></td>
<td>Traditional Data Warehouse / Analytical Applications</td>
</tr>
<tr>
<td>Structured</td>
<td>Low</td>
<td>Batch</td>
<td>Massively Parallel Processing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Near Real Time</td>
<td>Massively Parallel Processing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Real Time</td>
<td>In-Memory Appliances</td>
</tr>
<tr>
<td>Semi-Structured</td>
<td>Low</td>
<td>Batch</td>
<td>Massively Parallel Processing</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Near Real Time</td>
<td>Massively Parallel Processing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Real Time</td>
<td>In-Memory Appliances</td>
</tr>
<tr>
<td>Un-Structured</td>
<td>Low</td>
<td>Batch</td>
<td>Distributed Clusters</td>
</tr>
<tr>
<td></td>
<td>High</td>
<td>Near Real Time</td>
<td>Distributed Clusters</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Specialized Massively Parallel Processing</td>
</tr>
</tbody>
</table>

Massive Parallel Processing Vendor Snapshot
- EMC GreenPlum
- IBM Netezza
- Teradata Aster
- Oracle Exadata

In-Memory Appliances Vendor Snapshot
- SAP HANA
- Oracle Exalytics
- Kognita

Distributed Clusters Systems Vendor Snapshot
- Hadoop MapReduce
- Aster SQL-MapReduce
- LexisNexis HPCC
Monetizing Big Data for a Telecom Operator
Sanitized Client Example
Three phases to creating the Big Data business

**Identify and Prioritize Opportunity Areas**
- Examine value chain structure, competitors
- Identify current data assets and applications
- Size market segments
- High-level comparison of assets and use cases

**Determine Product Platforms and High-Level IT Capabilities**
- Conduct Deep Dive competitor analysis
- Compare Client data and analysis capabilities, identify differentiation opportunities
- Refine business model elements and assumptions
- Conduct concept and business model validation
- Translate platform business requirements into IT functional requirements
- Conduct deep dive into IT capabilities (current, planned)
- Assess IT capability gaps
- Develop organizational development plan

**Build out Detailed Product and IT Requirements**
- Develop detailed product “crunchy questions” and business requirements
- Interview customers for second-level product validation
- Architect IT solution and develop roadmap
- Engage alpha/beta prospective partners
Client could play across all the components of the Big Data value chain but considerations should inform prioritization

Key considerations for Client’s play in Big Data

- How does the strategy create synergies with existing assets: network, distribution channels?
- How can Client leverage its current assets and relationships to enter the Big Data space?
- What sequencing and plays would be appropriate given core assets and use cases already funded?
- How does the strategy support other growth areas such as IoT, M2M and Security?
- Can Client competitively differentiate and assume a leadership position? How does the track record support the estimated benefits?
In addition to market size, the team considered the importance of capabilities to each macro opportunity.

For the most attractive macro opportunities, Deloitte identified more detailed business opportunities.

Analysis of these opportunities identified key capabilities that the Client would need to either integrate internally, or source from external partners.

Based on this assessment and fit with the client’s capabilities, certain business opportunities were prioritized for deeper business case, product, and technology analysis.

### Opportunity prioritization and needs assessment indicate strong correlation of critical capabilities across use cases

<table>
<thead>
<tr>
<th></th>
<th>Data Sources</th>
<th>Technology</th>
<th>Analytics</th>
<th>Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Internal</td>
<td>External</td>
<td>Infrastructure</td>
<td>Data Mgmt</td>
</tr>
<tr>
<td></td>
<td>Consumer</td>
<td>Business</td>
<td>Strategic</td>
<td>Alliance</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Network</td>
<td>Connectivity</td>
<td>Computing</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Policy</td>
<td></td>
<td>Analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Batch</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Real-Time</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Predictive</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Visualization</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Customer</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: How important is the need for the use case?
- Not Required
- Needed for Strategy
- Critical
- Strong Enabler

6. Internal Improvements
   - Chem Analytics
   - Cust. & Industry Dashboard

1. Infrastructure
   - Big Data add

2. Data Analytics Services
   - Network Insights
   - MDM Analytics
   - Threat Mgmt

3. Data Insights & Monetization
   - Enhanced Data

4. Advertising Services
   - Ad Services (Digital)
   - Ad Services (TV)
Product development approach is informed by customer needs, competitor offerings, and Client’s strength

- High value potential customer segments
- Marketing intelligence needs of the customers

CUSTOMERS
What Problem are we trying to solve?

COMPETITORS
How are those problems being solved now?

COMPANY
What can Client do to solve the problems better?

- Competitor offerings
- Competitor capabilities and strengths
- Competitor weaknesses

- Under/unmet customer needs
- Client’s unique assets
- Client’s product offerings for solving the unmet needs
Three segments show higher propensity for Big Data driven insights

Notes:
1. Based on velocity, variety and volume of data generated in the industry and its relevance for consumer insights
2. Based on top 100 companies in the US by advertising dollar spend

Source: Gartner, AdAge, Deloitte analysis
### Competitors differentiate on a limited set of individual data points

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Transactional (PoS etc.)</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>2. Behavioral and attitude</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>3. Social influence</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>4. Decision path</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>5. Audience tracking</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>6. Digital usage</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>7. Mobile (device/app/network)</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>8. Call, video, text tracking</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>9. Demographic/Lifestyle</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>10. Customer ID</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>11. Location</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
<tr>
<td>12. Financial (credit, income, etc.)</td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
<td><img src="#" alt="Differentiation" /></td>
</tr>
</tbody>
</table>

- Most competitors derive their strength from generation of or access to 1-3 distinctive data sources
- Many competitors are procuring data, or purchasing competitors, to keep up with marketplace changes
- Location-based data is a relatively new area; however, many competitors – both large and small – are pursuing opportunities in this space
Competitor data limitations create opportunity for Operators

**Competitor Limitations**

**Location Insights**
- **Static** location information, updated once or twice a year
- **Home address** (Zip+4) most common form of available location information
- Limited ability to track real-time location

- **Fragmented view** of customer across screens (TV, internet, and mobile)
- Based on **panel data**
- Limited visibility into mobile internet and application usage

**Audience Measurement**
- Visibility in communication patterns **limited to online social media**
- Network map does not necessarily identify influencers or active users

**Social Intelligence**
- Ability to supplement online social media information with **voice and text usage**
- Can be a better predictor of who can influence whom

**Operator Advantages**

- **Dynamic** location, available approximately every 15 minutes
- Location tracked **within a radius of ≤X m**; could be less with WiFi
- **Real time location** (home, office, shop, cafe, road, social venues, etc.)

- **Integrated view** of customer across multiple screens
- Based on **linked users** across TV, internet & mobile for triple play subscribers
- Mobile internet and application usage visibility into mobile subscriber base
Operator advantages in data and capabilities pointed to differentiated product platforms, anchored in how to solve real client problems in a differentiated offering.

High-level requirements were defined for data, integration, analytics, and delivery model dimensions.

Operator advantages were turned into three product platforms:

1. **Product platform description: Audience Measurement**

   **Product Description**
   - **What is the Product?**
     - Audience measurement products track the impressions generated across three primary media networks – TV, digital, mobile.
     - Could be enhanced with other data.
     - Insights may be sold in isolation for each media or combined into an integrated, cross-media view.

   - **What client problem does it solve?**
     - With current solutions, marketers are forced to measure the efficiency of their media buys in silos across TV, digital, and mobile.
     - They are unable to get an integrated view of the customer’s media and advertising consumption.
     - There is no industry standard for mobile ratings.

   - **Who will be our target clients?**
     - Industries:
       - Consumer goods
       - Financial services
     - Companies:
       - P&G
       - Kraft
       - Coke
       - Levi's

   - **How will it differentiate us from the competitors?**
     - Will provide a single, integrated view of customer based on their consumption across various media; competitors provide view by media.
     - More precise TV ratings compared to the competition.
     - Industry standard mobile ratings.

   **Product Requirements**
   - **Data**
     - Marketing impressions for each of AT&T’s primary media.
     - Full spectrum of customer profile information (e.g., demos, financial, behaviors, etc.).

   - **Data Integration**
     - Integrate company and procured/public data into a segmentation of media user groups.
     - Integrate Client media impression ratings with clients’ data to make allocation decisions.

   - **Analytics**
     - Match different media types together.
     - Extrapolate from Client’s existing TV, digital, mobile footprints into a national media view.
     - Allow clients to scenario-plan and calculate optimal media mixes.

   - **Delivery**
     - Access: 24/7 user access for data and analytical tools, pass-code protected.
     - Delivery method: 24/7 dashboard, supplemented by ad-hoc consulting analysis.

Copyright © 2014 Deloitte Development LLC. All rights reserved.
Team identified current market offerings and customer use cases

Based on Client’s competitive strengths, opportunities to bring a differentiated solution were identified (3 major platform areas, each with multiple initial product concepts)

Team developed storyboards, tested with Deloitte Subject-Matter Advisors and selected CPG and Retail client executives

Based on feedback, team pruned less powerful concepts and built out more detailed product requirements for the remainder

**Use case: Cross-media targeting and buying efficiencies**

**Context**
A consumer goods company has over the years learned to optimize targeting and buying efficiencies, but they have had to do so in silos, medium by medium, because of shortcomings in media ratings data.

**Existing Practice**
- Multi-media campaign purchases – silo
  - Advertiser engages and pays different vendors for different media ratings; in order to optimize the parts, (but unclear on the whole)
  - Can only improve media in their silos, not the entire campaign. Potentially holding back an additional 13%+ of buying efficiency

**Client X Approach**
- Multi-media campaign purchases - Integrated
  - Advertiser engages and pays Client X for an integrated perspective on media targeting and efficiencies.
  - Strategies based on Product insights:
    - Optimize ad spend allocation across multi-channel
    - Maximize marketing ROI by creating more unique impressions within the given budget
    - Optimize communications based on consumers media consumption pattern

**Current Results**
- Disparate views: media purchasers have to buy TV, digital, mobile and other media, based upon disparate sets of information.
- Hard to evolve to a customer view: in addition to buying inefficiencies, siloed media data inhibits campaign strategists from taking a customer viewpoint.

**Improved Results**
- Single view: media purchasers find 13% improvement in basic buying efficiency, by looking across multiple media in an integrated fashion.
- Customer view: campaign strategy and management begin to evolve to a customer orientation. For example, media is purchased based upon the daily consumption habits across media.
Business case built upon market sizing, share projections, and preliminary business model.
Starting point: Big Data reference architecture

Data Path

- Data Visualization/Serve
- Data Exploration and Analysis Services
- Analytical Services Delivery Platform
- Data Storage Services
- Data Integration & Mash Up Services

Real Time and Batch

- Voice & Data
- Content Providers
- 3rd Party Aggregators
- Customer Domain

Internal & External Data Services

- Social Media
- Sensor Data
- Web
- Enterprise Applications

Connected Services

- Mobile Apps
- 3rd Party Apps
- Customer Portal

Common UI Framework

- Private API
- Public API
- Security Services
- Content Mgmt.

Anonymized & Aggregated

Deloitte Proprietary and Confidential
Deloitte Pre-built Big Data Ecosystem

Data Sourcing
- Internal
  - ORACLE
  - SAP
  - All Major Databases
- External
  - Geographic
    - REST API
  - Web Logs
    - Apache
  - Social Media
    - Twitter
    - Facebook
  - Others

Data Storage
- Hadoop Clusters / Relational
  - HDFS
  - hadoop MapReduce
- Batch
  - Spark
  - Sqoop
- No SQL/Columnar
  - cassandra
  - Apache HBase
  - VERTICA
- Analytics and Discovery
  - mahout
  - IBM SPSS
  - Apache Kafka
  - Hive QL / SQL
  - Spark
  - Vertica
  - Hive QL / SQL
  - VERTICA

Data Preparation / Calculation
- Analytics and Discovery
  - R Studio
  - Tableau
  - SPSS
  - IBM SPSS
  - R
  - No SQL/Columnar
  - Hive QL / SQL
  - Apache Kafka
  - Storm

Analytics
- Customized Dashboards and APIs
  - QlikView
  - Tableau
  - Spotfire
  - TIBCO Software

Visualization and Discovery
- Apache Kafka
  - Storm
  - Vertica
  - Hive QL / SQL
  - Spark
  - Vertica
  - Hive QL / SQL
  - Apache Kafka
  - Storm

Data In Motion Path
- Data Replication
- Data Sourcing
- Data Storage
- Data At Rest Path

Data At Rest Path
- Hadoop Clusters / Relational
  - HDFS
  - hadoop MapReduce
- Batch
  - Spark
  - Sqoop
- No SQL/Columnar
  - cassandra
  - Apache HBase
  - VERTICA
- Analytics and Discovery
  - mahout
  - IBM SPSS
  - Apache Kafka
  - Hive QL / SQL
  - Spark
  - Vertica
  - Hive QL / SQL
  - VERTICA

Data At Rest Path
- Hadoop Clusters / Relational
  - HDFS
  - hadoop MapReduce
- Batch
  - Spark
  - Sqoop
- No SQL/Columnar
  - cassandra
  - Apache HBase
  - VERTICA
- Analytics and Discovery
  - mahout
  - IBM SPSS
  - Apache Kafka
  - Hive QL / SQL
  - Spark
  - Vertica
  - Hive QL / SQL
  - VERTICA

Data In Motion Path
- Data Replication
- Data Sourcing
- Data Storage
- Data At Rest Path