Demystifying the Big Data Phenomenon for Strategic Leadership

Society for Advancement of Management
2016 International Business Conference
Arlington, VA April 7 – 10th

LauraAnn Migliore, Ph.D.
Faculty, University of Phoenix
School of Advanced Studies (SAS)
lamigliore@email.phoenix.edu

and

Ravi Chinta, Ph.D.
Department Head of Business Administration
Auburn University at Montgomery
ravichinta@hotmail.com
This presentation explains the concept of Big Data using Machine Learning practices as a viable approach to finding new ways to gain insights and add value.

The objective is to demystify the Big Data phenomena through the lens of data as a currency for monetization using a strategic leadership perspective to advance management, inspire innovation, and influence positive change.

This presentation discusses research data from more than 100 C-level retail executives and uses anecdotal evidence to integrate company examples to demonstrate empirically the link between investments in digital technologies for Big Data analytics to improve performance and create new value for customers.
Outline

- Problem Statement & Research Opportunities
- Literature Review
- Methodology & Data Analysis
- Limitations of the Research
- Discussions, Recommendations, & Conclusions
To date, there is limited research on digital leadership understanding to apply IT capabilities effectively for purposes of achieving competitive advantage using Big Data and Machine Learning practices.

There are many professional news-media sources (e.g. case reports, white papers, and magazine articles) to help increase leadership awareness about advanced-computing technologies (McKinsey & Company, 2014).

Few peer-reviewed journal articles on the leadership dimension of digital thinking and learning to achieve success with Big Data and Machine Learning practices
Three Literature Themes

1) Evolution of Data in Organizations
2) Data Currency and Monetization
3) Strategic Leadership for Digital Transformation
Big Picture of Data Analytics

Operational
- Data Strategy
- Reporting
- Dashboards

Competitive Advantage
- Operational Effectiveness
- 360 View of Customer
- Business Agility
- Price Optimization
- Revenue Optimization
- Social Media Analytics
- Demand-driven intelligence
- Fraud detection

Strategic
- Data Mining
- Predictions
- Deep Insights

Using Apache Hadoop (open-source IT environment) for distributed storage & processing of large sets of data on low-cost commodity hardware

- Core technology is the Hadoop Distributed File System (HDFS) – (Java-based; layered platform designed to span large clusters of commodity servers)
- Pluggable architecture for wide variety of data access methods & scalable to data needs
Multiple layers in artificial neural networks (process mapping)
Each neuron is a constrained mathematical function (algorithm)
Neurons determine pattern or feature in a data set
Data Currency & Monetization

Example: Starbucks Customer-Connected-Coffee Experience

Power of data via the Starbucks loyalty-card program:

Results:

(1) **Discovered new monetization opportunities**, such as potentially licensing out a white-label mobile platform to other national retailers (Kaplan, 2014).

(2) **Captured and analyzed data via machine learning models** to obtain new insight for improving the customer experience and increasing revenues.

(3) **Enhanced mobile application** for customers to place orders anywhere from over 7000 locations in the U.S.; quick-orders from past orders on the menus screen and/or can find other offerings; provides an estimated wait time and uses Apple or Google maps to direct the customer to the closest Starbucks using location services within their mobile device.
Zulily: Machine Learning
(Based in Seattle, WA; e-commerce company)

• **Business Need**: Turn vision into reality by providing customers (primarily moms) a unique and engaging online purchasing experience every day

• **Problem**: Silos of data within its current IT infrastructure and with a growing customer base and a system that could not keep up with volumes of new data

• **Action**: Combined IT resources using the Hadoop architecture to build its own Zulily Data Platform (ZDP) for Big Data with Machine Learning applications

• **Results**:  
(1) Gained internal-processing efficiencies to drive personal content to customers  
(2) Optimized timely shipments to customers  
(3) Delivered online content in mobile applications & desktop computer browsers
• **Business Need:** Provide safe, reliable, accessible, and customer-focused service in a fiscally-responsible manner

• **Data Problem:** *Volume* (over 330 data points); *Variety* (electronics/sensors; telematics); *Velocity* (10-second intervals of streaming data); and *Veracity* of uncertain data (various components and systems)

• **Action:** Compressed the data and loaded into Hortonworks Data Platform (HDP) using the Hadoop Distributed Filesystem to analyze data through algorithms, identify pattern recognitions, and correlated events associate with component failures and maintenance inefficiencies.

• **Results:**
  (1) Improved decision making and accuracy in inventory replenishment
  (2) Gained efficiencies in parts and labor costs
  (3) Increased bus lifespans
Methodology & Data Analysis
Model constructed from employee-customer-profit chain dimensions, which included employees, customers, and investors (Rucci, Kirn, & Quinn, 1998) and IT capabilities (Mazidi, Amini, & Latifi, 2014) for 10 different types of digital technology investments (DTI) currently transforming organizational processes (Migliore & Chinta, in press).
Population, Sample, & Data Collection

- **U.S. retail C-level executives** (CEOs, CIOs, CTOs, etc.)

- Data collected through reliable and validated **online-panel samples** using a third-party reputable vendor, who has specialization in digital technologies for online market research.

- The sample (N = 107) is **generalized to U.S. retail firms** having at least 50 employees on payroll and in business for at least five years (Migliore & Chinta, in press).

- Respondents were **currently employed C-level executives** with one of the following job-position titles:

  - Owner
  - Chief Executive Officer (CEO)
  - President
  - Chief Information Officer (CIO)
  - Chief Technology Officer (CTO)
The representation of the retail firms were diverse in retail sector:

<table>
<thead>
<tr>
<th>Category</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>General merchandise</td>
<td>24%</td>
</tr>
<tr>
<td>Department stores</td>
<td>22%</td>
</tr>
<tr>
<td>Technology/electronics</td>
<td>15%</td>
</tr>
<tr>
<td>Food</td>
<td>12%</td>
</tr>
<tr>
<td>Health &amp; beauty/wellness</td>
<td>7%</td>
</tr>
<tr>
<td>Home improvement – construction</td>
<td>5%</td>
</tr>
<tr>
<td>Entertainment</td>
<td>4%</td>
</tr>
<tr>
<td>Automotive</td>
<td>3%</td>
</tr>
<tr>
<td>Luxury</td>
<td>3%</td>
</tr>
<tr>
<td>Other</td>
<td>5%</td>
</tr>
</tbody>
</table>
Data Analysis – Key Findings

- The unit of analysis is at the firm level

- **Data analysis resulted in seven regression models to predict 10 different types of digital technology investments (DTI)**

- **The 10 Digital Technology Investments (DTI) included:**
  
  Data Analytics  
  Mobile Device Application  
  Mobile Commerce (Wallets)  
  Crypto-Currency Acceptance  
  Cloud and Distributed Computing  
  Enterprise application architecture  
  Cybersecurity  
  Data architecture  
  SMART facilities – energy management systems  
  Connected Car (wireless mobile)
Data Analysis – Key Findings

Data analytics and Connected Car showed statistical significance

According to Migliore & Chinta (in press):

“Data analytics showed significance when using Pearson’s test ($r = .198 \mid .041$ sig.), but fell just outside the range of significance when using Kendall Tau ($T = .159 \mid .058$). However, connected car showed significance using both: Pearson’s test ($r = -.275 \mid .004$ sig.) and Kendall Tau ($T = -.189 \mid .018$ sig.). The time frame for data analytics mean was 2.35, which from the survey translates into “need now or within current year” and the connected car mean was 2.95, which from the survey translates into “one to two years.” (p. 23)
Hardware & Software Purchases for Digital Initiatives

Core Tenets

• **The key predictor is understanding IT capabilities**, which was extracted via factor analysis using dimensions of the employee-customer-profit (ECP) chain model and the different data sources that C-level executives use to **reflect** upon to gain understanding – reflection is a **mediating variable** in the model.

• The data sources for understanding IT capabilities included the following from Migliore & Chinta, in press:

  - Information technology audits
  - Technical consultations
  - Technical training classes
  - Trade industry or Professional Conferences
  - Risk assessment
  - Conversations with employees, customers, suppliers, investors and peers
  - Personal observations
  - Gut-felt intuition
Hardware & Software Purchases for Digital Initiatives

Core Tenets

One important finding of the Migliore & Chinta (in press) study is the mediating role of the C-level leader’s focus, which include importance and frequency of reflection.

Here is where thinking preferences for what is considered important for improving financial performance is revealed in the frequency of reflection upon key data sources to obtain understanding about their firm’s ECP chain, IT capabilities, and ability to create value for competitive advantage.

These findings align with the leadership-thinking preferences and readiness to develop digital competencies for the digital age (DeBoer, Bothma, & Olwagen, 2012).
## Model Summary & Coefficients

### Table 1: Regression Model Summary for Hardware & Software purchases in DTI

<table>
<thead>
<tr>
<th>PURCHASE Type</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
<th>R Square Change</th>
<th>F Change</th>
<th>df1</th>
<th>df2</th>
<th>Sig. F Change</th>
<th>Durbin-Watson</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDWR (1)</td>
<td>.534a</td>
<td>.285</td>
<td>.278</td>
<td>.821</td>
<td>.285</td>
<td>41.867</td>
<td>1</td>
<td>105</td>
<td>.000</td>
<td>--</td>
</tr>
<tr>
<td>HDWR (2)</td>
<td>.570b</td>
<td>.325</td>
<td>.312</td>
<td>.801</td>
<td>.040</td>
<td>6.198</td>
<td>1</td>
<td>104</td>
<td>.014</td>
<td>2.184</td>
</tr>
<tr>
<td>SOFTWR</td>
<td>.510b</td>
<td>.260</td>
<td>.253</td>
<td>.767</td>
<td>.260</td>
<td>36.871</td>
<td>1</td>
<td>105</td>
<td>.000</td>
<td>2.094</td>
</tr>
</tbody>
</table>

a. Predictors: (Constant), IT Capabilities Understanding
b. Predictors: (Constant), IT Capabilities Understanding, Customer Perspective Understanding

### Table 2: Coefficients for Hardware & Software purchases for Digital Initiatives

<table>
<thead>
<tr>
<th>PURCHASE Type</th>
<th>Primary Predictor</th>
<th>Unstandardized Coefficients B</th>
<th>Std. Error</th>
<th>Standardized Coefficients Beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>HDWR (M1)</td>
<td>IT Capabilities Understanding (IV)</td>
<td>CNST = 3.991</td>
<td>.079</td>
<td>.534</td>
<td>50.281</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.516</td>
<td>.080</td>
<td></td>
<td>6.470</td>
<td>.000</td>
</tr>
<tr>
<td>HDWR (M2)</td>
<td>Customer Attitudes Understanding (IV)</td>
<td>CNST = 3.991</td>
<td>.077</td>
<td>.347a</td>
<td>51.510</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.335a</td>
<td>.106a</td>
<td></td>
<td>3.154a</td>
<td>.002a</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.265b</td>
<td>.106b</td>
<td>.274b</td>
<td>2.489b</td>
<td>.014b</td>
</tr>
<tr>
<td>SOFTWR</td>
<td>IT Capabilities Understanding (IV)</td>
<td>CNST = 4.075</td>
<td>.074</td>
<td>.510</td>
<td>54.974</td>
<td>.000</td>
</tr>
<tr>
<td></td>
<td></td>
<td>.452</td>
<td>.074</td>
<td></td>
<td>6.072</td>
<td>.000</td>
</tr>
</tbody>
</table>

a. M1 Predictors: (Constant), IT Capabilities Understanding (IV)
b. M2 Predictors: (Constant), Customer Attitudes Understanding (IV)
Regression Model Explanation

For Model 1 hardware purchases [HDWR (M1)], when the C-level executive’s perception increases by one unit in understanding IT capabilities within their retail organization, the ability to predict likelihood to purchase hardware increases by 51.6 percent (Migliore & Chinta, in press).

These empirical results represent an implication to resource allocations by C-level executives in their emphasis on IT hardware investments as they begin to understand the importance of IT capabilities within their organizations.

For Model 2 hardware purchases [HDWR (M2)], when the C-level executive’s perception increases by one unit in understanding customer attitudes towards their retail organization, the ability to predict likelihood to purchase hardware increases by 26.5 percent, holding constant understanding IT capabilities (Migliore & Chinta, in press).

These empirical results represent an implication to resource allocations by C-level executives in their emphasis on IT hardware investments as they begin to understand the importance of understanding customer attitudes coupled with understanding IT capabilities within their organizations.

For software purchases (SOFTWR), when the C-level executive’s perception increases by one unit in understanding IT capabilities within their retail organization, the ability to predict likelihood to purchase software increases by 45.2 percent (Migliore & Chinta, in press).

These empirical results represent an implication to resource allocations by C-level executives in their emphasis on software investments as they begin to understand the importance of IT capabilities within their organizations.

NOTE: The measurement for C-level executive understanding of IT capabilities within the organization was constructed from their responses on the data sources they use to understand the current status (Migliore & Chinta, in press). These data sources were measured using a Likert scale (1 – 5) on effectiveness and included the following: IT audit results, knowledge obtained from technical consultations, skills and competencies developed by taking technical training classes, insights and perspectives learned from trade industry or professional conferences, results from risk assessments, new information and knowledge obtained via conversations with employees, customers, suppliers, investors and peers, personal observations, and intuition.
Limitations of the Research

• The primary limitation in this study is self-reported responses used to construct the variables within the model (Figure 1)

• Next, the construct for the independent variable (understanding the ECP Chain) is C-level executive perception of understanding and **not** an actual test to measure and compare respondent scores to a set standard of what constitutes understanding of ECP

• Another limitation is the timing of the data collection, which occurred in late 2014. Since that time, the term and meaning of *Big Data* has continued to evolve among executive leader understanding to include Machine Learning Practices, as explained in this presentation

• For the phenomenon studied (executive’s understanding of ECP chain upon digital technology investments) there are many other variables that are not in our research frame and thus, may limit our generalization in *sample size (107)* to the greater U.S. retail population, even though our online-panel sample represented the characteristics of the greater U.S. retail population
Discussion Items

• **The results of the Migliore & Chinta (in press) study** show a sample of U.S. C-level retail executives whose perception of understanding, based upon focus of reflection and level of importance in IT capabilities, will lead to the likely investment in hardware and software purchases for digital technology initiatives (DTI)

• Only **data analytics** and **connected vehicle** showed statistical significance among the 10 DTIs; however, with the ongoing evolution of understanding IT capabilities, it is anticipated executive leaders will lean more towards cloud-distributed computing and find cost and processing efficiencies with the Hadoop architecture

• **Data is like currency** that can be used to create new value in customer experiences, increasing revenues through new monetization opportunities – data brokering industry and personal information to increase precision in predicting revenue and/or cost impact

• **Leaders need digital competencies** to understand how to translate meaning of data into action
8 Recommendations

To achieve strategic objectives aligned to your organization’s vision and mission statements, we recommend looking through the lens of learning and discovery when approaching Big Data and Machine Learning practices:

1. Be willing to explore the opportunity in the context of the problem and what it could mean for your organization by aligning it with a value-creating strategy.

2. Remain open to how Big Data/Machine Learning can improve the quality of your data.

3. Test assumptions to identify where gaps in data may be.

4. Gain clarity among the data science team and other key stakeholders by asking questions to address most critical business problem(s).
5. Exchange perspectives by having meaningful dialogue with data science teams and other key stakeholders

6. Structure the solution to identify the best data design and select the most appropriate algorithm(s) for addressing the problem(s)

7. Open the channels of feedback to remove any silos and/or barriers to keep communications loop flowing

8. Execute appropriate actions and capture lessons learned to teach others and build digital competencies within the organization
Conclusion

• The key to demystifying the Big Data phenomena is to see it through the lens of revenue and profit, where data is a currency to be monetized for creating new value

• The strategic leadership implications include improved productivity, collaboration, and innovation via the combination of IT resources and capabilities within cloud-distributed environments

• These IT capabilities are made possible through stacked technology platforms like Hadoop and through use of Machine Learning algorithms to boost competitive advantage for those organizations who understand its potential and power in applying IT capabilities
For Additional Information and/or Questions

For more information on the Migliore & Chinta (in press) study, see the article, “Mobile Technology and the Employee Customer Profit (ECP) Chain” in the 2016 Winter Issue (81), Volume 1 of the SAM Advanced Management Journal (Pages 52-69).

Primary Contact:  Dr. LauraAnn Migliore
lamigliore@email.phoenix.edu
References


• Business Insider (2012). The Digital 100: The world’s most valuable private tech companies.

  designing training programmes to build leadership competencies in the digital age. South

• Green, F. (2012). Employee Involvement, Technology and Evolution in Jobs Skills: A Task-Based

• Hortonworks (2016). Metro Transit of St. Louis using Hadoop.

• Hortonworks (2016). Zulily does Hadoop.

  White paper. Mobidia.

References (continued)


References (continued)


Detailed Discussion Slides
Evolution of Data in Organizations

• Costs of traditional mainframe systems have dropped dramatically with distributed cloud-computing systems (Karlgaard, 2015)

• Significant rise in smartphone adoption and global Internet access

• Expanding universe of data creation in the online

In 2014, the global quantity of information was doubling every eight months.

In 2015, it was doubling every four months.

By 2020, the world will have a quadrillion times more information than it has now.

Source: Chief Executive.Net | March/April 2016
The **data-brokering business** generates over $300 billion dollars per year; employs 3-million employees (Morris & Lavandera, 2012).

**Data brokers** collect personal information about consumers and sells that information to other organizations -- may be referred to as database marketers or consumer data analytics firms.

**Data types:** your name, phone number, residence, buying habits, interests, etc.

**Data collection:** courthouse records, mortgages, voter registration information, website cookies, loyalty card programs, social media platforms, retailer records including credit and debit card purchases, and/or information voluntarily shared online by consumers – surveys, sweepstakes entries, and warranty registrations.
**Data Currency & Monetization**

- **Largest U.S. data-brokering firm:** Acxiom is headquartered in Little Rock, AK

- **CEO leadership perspective (Scott Howe):** It’s all about gaining intelligence to know what people are interested in to deliver appropriate advertising offers for products and services (Morris & Lavandera, 2012). We provide secure, appropriate, and legal data-selling services to eliminate spam emails and irrelevant ads cluttering user website and applications.

- **Critics perspective:** It’s deceptive and increases privacy risks; majority of consumers don’t know their personal data is being sold and machine algorithms are making decisions about them (e.g. employment, insurance rates, credit, etc.)

---

**Data Categories:**
- Plus-size Apparel
- African-American Professional
- Biker/Hell’s Angels
- Allergy Sufferer
- Exercise – Sporty Living
- Working Class Mom

**Data Categories:**
- Burdened by Debt: Singles, X-tra Needy
- Credit Crunched: City Families, Ethnic Second-City Strugglers
- Fragile Families
- Small Town Shallow Pockets
Today, mass amounts of historical data (structured and unstructured) can be ingested into systems where a program (artificial intelligence) learns from the data using an algorithm to evaluate the data and create a model to predict future occurrences.

Examples include: Forecasting demand of sales, personalizing content, flagging suspicious transactions, detecting fraudulent orders, and filtering reviews, etc.

Effective digital leaders will be known by the questions they ask – human in the loop to guide and direct the algorithms:

- How likely is a customer to purchase a particular product based on past behavior?
- Does the algorithm need to be adjusted?
- How does the algorithm interact with other algorithms?