DECISION SCIENCES INSTITUTE

An Examination of Big Data Capabilities in Creating Business Value

(Full Paper Submission)

Shiwei Sun  
Auburn University  
szs0100@auburn.edu

Suning Zhu  
Auburn University  
szz0012@auburn.edu

Xiaoye Cheng  
Auburn University  
xzc0017@auburn.edu

Terry Byrd  
Auburn University  
byrdter@auburn.edu

ABSTRACT

As the era of big data has arrived, using big data capabilities to create business value has become organizations’ strategic consideration. In current data-rich environments, how to make a good decision and create business value has caused scholars and practitioners’ attention. From the perspective of creating business value, we identified three main capabilities by content analysis, including analytical capability, predictive capability, and decision support capability.

KEYWORDS: Big data capabilities, Business value, Content analysis

INTRODUCTION

Because of the fierce competition and turbulent market environments, many firms adopt state-of-the-art information technologies to develop competitive advantages, which has become a part of organizational information systems (IS) strategy (Low et al., 2011; Sultan, 2010). For example, digital business strategy has gained great attentions both from scholars and practitioners (Bharadwaj et al., 2013).

Data-driven decision-making refers to the practice of basing decisions on the analysis of data, rather than purely on intuition (Provost & Fawcett, 2013). Data has been likened to the new oil to create value for business. The refined data enable firms to identify profitable activities, which in return drives firms’ growth (Rotella, 2012). However, after getting reliable information sources, how to extract and analyze useful results from data becomes the focus for firms’ decision-making. Big data technology is one specific form of digital technologies, which can be used to support the quality of decision-making, and, in turn, create business value.

Scholars have widely acknowledged that big data is a business solution rather than a technical one (e.g., McGlinchey, 2013). In recent years, characterized as volume, variety and velocity (3Vs) (Manyika et al., 2011), big data has been fashionable and critical in business and IT fields. Its role has been well recognized for firms in decision-making, marketing, and many other aspects. However, most industries are still at the early stage of big data adoption. The process of how big data technology creates business value has not yet been explored. The capabilities of big data have become the focus both in information systems (IS) research and industry.
Currently, the business environment is becoming more volatile. Firms face the challenge of making effective strategic decisions so as to cope with the market unprecedented turbulence and create the emerging business opportunities. In this dancing, rugged, and competitive landscape, internal operations and external environment of the firm exacerbate the uncertainty for prediction.

The emergence of big data technology makes the business process more agile in response to the current turbulent environment. In the era of big data, data-driven decision-making is proved to have an important impact on firm performance. For example, the study from Brynjolfsson, Hitt, and Kim (2011) statistically indicates that the more data-driven a firm is, the more productive it is after controlling for a wide range of possible confounding factors. McAfee and Brynjolfsson (2012)'s analysis also show that the more companies characterized themselves as data-driven, the better they performed on objective measures of financial and operational results.

The fundamental purpose of our study is to explore the main big data capabilities in the business world from the extant academic and practical publications. To serve this end, we need a method that can help us extract the elements of big data capabilities from the text in current literature. Content analysis method is to derive the content and address our specific research question.

In this study, we first define big data and identify the general role of big data. Following that, we make an inductive content analysis and summarize the main capabilities of big data. Based on the analysis results, we will discuss how big data capabilities create business value.

THEORETICAL BACKGROUND

Defining Big data

Although scholars have not reached an agreement on the definition of big data, it is commonly believed that big data is derived from business intelligence and analytics (BI&A). From data processing perspective, some scholars see it as the ever-increasing information from different sources that is very “big” to process (e.g., Liu, 2013). From business solutions perspective, some other scholars consider it as an approach that assists organizations to analyze the huge amount of data (e.g., Boubeta-Puig et al., 2014). However, in general, most scholars accept the viewpoint that big data mainly refers to big sets that are overpowering in size (e.g. Chen et al., 2013).

According to the diffusion of innovation theory, big data can be considered an innovation for organizations (Rogers, 2003). As a new technology, it is characterized with its advanced BI&A functions to deal with the huge amount of data. Its unique capabilities, such as analytical and predictive capabilities, allows firms to create business values (Sun et al., 2014).

To summarize the above definitions and our understandings, big data is different from other new innovations because of its unique capabilities. It serves for a business solution rather than a technical function. Indeed, the real-world use of big data is to generate business value. Not only does it deal with the amount of data that couldn’t be processed by the traditional approach, but also it can provide business solutions due to its unique capabilities. The analysis results from big data might be used to help organizations identify potential profitable business activities to achieve competitive advantages. The analysis results might be also used to help organizations develop a strategic blueprint.

Big data capabilities in organizations

Big data can do a lot of things for organizations. For example, from the view of strategy-as-practice, Whittington (2014) argued that big data have an effect on strategizing. In other words, big data influences the distinctive practices that make organizational strategies come out. From
a macroscopic view, big data also serves the foundation of organizational innovation, competition and productivity (Lycett, 2013).

The emergence of big data technology provides a good solution to the quality of decision-making. Prior research discussed the role of big data capability from the aspect of its contribution to a better strategic decision-making. For example, Demirkan and Delen (2013) argued that the capabilities of service-oriented decision support systems could be acquired through cloud computing in big data analysis. McAfee and Brynjolfsson (2012) argues that using big data leads to better predictions which yields to better decisions and precise interventions. This could also lead to revolution of organizational decision-making culture.

Overall, the adoption of big data can improve organizational performance. Big data capabilities can be identified from various aspects. For example, Schroeck et al. (2012) emphasized that big data could be used to target customer-centric outcomes. In current turbulent business environment, the capability of analyzing the unstructured data (e.g., social media data) is extremely important for organizations’ development.

METHODOLOGY

We adopted content analysis to identify and categorize big data capabilities. Content analysis refers to “a research technique for making replicable and valid inferences from texts (or other meaningful matter) to the contexts of their use” (Krippendorff, 2012, p. 18). Content analysis is helpful to reveal the focus of attention from a large amount of text using both qualitative and quantitative approaches (Neuman, 2006). To achieve the predictive and inferential intent, content analysis processes the text from the words to numbers, and back to words, through factor analyses, word frequencies and other approaches. By the support of content analysis results from this process, the analyst can interpret and implicate for practice.

Content Analysis Process

Content analysis can be categorized to inductive content analysis and deductive content analysis based on the purpose of the research (Elo and Kyngäs, 2008). The inductive method is used when there is not a solid basement of knowledge or knowledge about the phenomenon is fragmentary (Lauri and Kyngäs, 2005). On the contrary, the deductive content analysis is applicable when former knowledge or theories exist (Kyngas and Vanhanen 1999). In order to identify big data capabilities in our research, which are dispersed in prior literature, an inductive content analysis can well serve our study purpose.

The inductive analysis process is to abstract categories of knowledge from particular text data, which is a process from the specific to general (Chinn and Kramer, 1999). Elo and Kyngäs (2008) proposed a process of inductive content analysis with three steps: preparation phrase, organizing phrase, and the phrase of reporting the analyzing process and the results (as shown in Figure 1). Following this process, we conduct content analysis of the former literature on big data as stated in the next paragraphs.
First, we defined the criteria to include and/or exclude a paper and determined the specific search items for a review of the literature.

Second, we developed the coding protocol and collected the data that addressed our research question. All the factors identified from the text should answer our research question. In order to increase the reliability assessment, the human coding and text mining software SAS Enterprise Miner will be both employed for our data collection work. The coders were trained on our protocol.

Third, we presented the results and explained the findings. We inferred results and demonstrated the validity. The indicator Krippendorff’s alpha will be used to evaluate the inter-coder reliability.

Target journals

Chen et al. (2012) listed the top journals, conferences, and industry magazines with BI&A publications. We believe these basket journals are good target articles for analysis. Additionally, in order to explore big data capabilities in business and management world, we generally consider high-quality information systems (IS) and general management journals as well as some distinguished practical journals. The eight in the “‘Senior IS Scholars’ Basket of Journals”, elite management journals, and five distinguished practical journals are deemed appropriately to be selected as the target journals. We went through three rounds to finalize the list of target articles for our analysis.

In round one, the criteria of selection was that the main content of the papers should focus on big data. In order to avoid omitting the articles that may contribute to our research purpose, we firstly downloaded the papers in the selected journals that contained the keywords “big data” within the text from January, 2005 to February, 2015. Those articles were believed to be relevant to our research interest. Among those articles, any comments or opinions towards big data capabilities were our major content focus. 329 articles were identified in this round.

In round two, after downloading all the potential articles with PDF format in round 1, we then used Adobe Reader’s “search” function, inputting keyword “big data”. Papers are deemed to be relevant to big data if their titles or abstracts contain “big data”. 150 articles were identified in this round.

---

Figure 1: Inductive Content Analysis Process

<table>
<thead>
<tr>
<th>Preparation Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Select the unit of analysis</td>
</tr>
<tr>
<td>• Making sense of the data and whole</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Organizing Phase</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Open coding</td>
</tr>
<tr>
<td>• Categorization</td>
</tr>
<tr>
<td>• Abstraction</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Reporting the Analyzing Process and Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>• Detailed description of process and methods</td>
</tr>
<tr>
<td>• Validity and reliability discussion</td>
</tr>
</tbody>
</table>
In round three, the full text of the 150 articles were reviewed through a consensus process to ensure that the article is closely related with big data capability. In order to keep a high reliability, a further check was made by the other coauthors. When there was a disagreement, we discussed together until getting a consensus. Finally, 95 articles were generated for the content analysis.

Based on Chen et al. (2012)’s recommendation and our selection criteria, the basket selected paper distribution is shown in Table 1.

<table>
<thead>
<tr>
<th>ACADEMIC PUBLICATION</th>
<th>#</th>
<th>INDUSTRY PUBLICATION</th>
<th>#</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIS Quarterly</td>
<td>1</td>
<td>Communications of the ACM</td>
<td>5</td>
</tr>
<tr>
<td>Academy of Management Journal</td>
<td>1</td>
<td>Harvard Business Review</td>
<td>7</td>
</tr>
<tr>
<td>Decision Support Systems</td>
<td>2</td>
<td>MIT Sloan Management Review</td>
<td>3</td>
</tr>
<tr>
<td>European Journal of Information Systems</td>
<td>1</td>
<td>Computer World</td>
<td>11</td>
</tr>
<tr>
<td>Journal of Information Technology</td>
<td>1</td>
<td>CIO Insight</td>
<td>27</td>
</tr>
<tr>
<td>Marketing Science</td>
<td>1</td>
<td>Computer Weekly</td>
<td>16</td>
</tr>
<tr>
<td>Management Science</td>
<td>1</td>
<td>KM World</td>
<td>8</td>
</tr>
<tr>
<td></td>
<td></td>
<td>CRM Magazine</td>
<td>7</td>
</tr>
</tbody>
</table>

Organizing

Each coder read all the selected papers carefully and made notes in styles of short phrases at the same time. The notes manifested the effects of adopting big data. After this, all notes were collected and put into a spreadsheet.

All coders gathered to discuss about categorizing big data capabilities based on all notes we took in the previous step. The number of categories is not limited. In other words, every coder can come up with as many categories as possible at this point. Then, all the categories were analyzed again by coders. Based on different aspects of effects that big data have, authors abstracted big data capabilities from categories. The main point of categorization was to reduce the number of big data effects and the goal of abstraction was to generalize big data capabilities. This abstraction process repeated until we obtained concise and reasonable categories of big data capabilities.

Coding method

After paper selection and organizing work, we moved to code the selected articles. Two steps were executed.

First, all the aspects that were associated with big data capabilities from a business perspective were our target of analytical constructs. Big data capabilities may vary greatly in different contexts. Our work was to recognize any potential capabilities cited in those articles that may impact firms’ operations and performance. Additionally, in terms of a given factor, different authors might use different forms of expressions. Because the factor might be mentioned at any section in an article, all the forms in elaboration with similar meanings were viewed as the same factor based on our analysis agreement.

Second, we focused on the frequency intensity for the articles analysis. In one article, the author might mention a specific capability many times. For the analysis convenience, we viewed it as occurring once in terms of calculating the frequency. We obtained the frequency of analytical
constructs using both SAS Enterprise Text Miner software and human face identification. This could help us overcome the limitations of the computer method in finding the associated capabilities. At the same time, during this process, the reliability and validity could also be improved. Then, according to our content analysis process agreement, the capabilities identified during content analysis were categorized based on the characteristics of related capabilities.

RESULTS

In this phase, categories of big data capabilities were reported and explained in details in our paper.

By conducting content analysis of the selected journals, we successfully found the main big data capabilities for business use at organizations. The total frequency of each big data capability identified in the articles was listed below (See results in Table 2).

From the perspective of creating business value, we identified three main capabilities: analytical capability, predictive capability, and decision support capability.

Table 2: Identified Big Data Capabilities and Corresponding Frequencies

<table>
<thead>
<tr>
<th>BIG DATA CAPABILITY CATEGORY</th>
<th>FREQUENCY</th>
</tr>
</thead>
<tbody>
<tr>
<td>Analytical Capability</td>
<td>59/95</td>
</tr>
<tr>
<td>Predictive Capability</td>
<td>29/95</td>
</tr>
<tr>
<td>Decision Support Capability</td>
<td>45/95</td>
</tr>
</tbody>
</table>

Figure 1: Tagcloud Visualization of Big Data Capabilities

We believed that our process of content analysis had reliability and validity. First, we had three coders working separately on all the selected papers which possibly covered all big data aspects in former literature. Second, our summary was based on understanding of all authors which could be considered reliable. The final agreement coefficient is 90%.

**Analytical capability**

We define analytical capability as an organization’s overall ability to capture, store, and process/analyze large volumes of data at or near real-time rates. It is the data processing capability that can deal with the data that could not be processed by the traditional approach.

The analytical capability enable firms to manage unstructured data and real-time data, which is the basic technical function of big data technology. The potential use of analysis results and conclusions could be roughly obtained from this fundamental capability. Some organizations might have already been handling big data for years. However, this current unique big data capabilities makes this era of big data different. For example, with the help of analytical capability, managers can easily get the visualized output and the extracted insights.

According to our content analysis results, analytical capability may include:
• Managing various structured and unstructured data types, including video, text, voice, e-mail, log files, transaction data (e.g., consumer data, sales data, and purchases data etc.), office documentation, and even transient data such as digitally-borne TV broadcasts and telephone conversations.
• Making far greater and faster data analytics.
• Collecting and processing disparate bits of information and data.
• Processing non-traditional data types.
• Unearthing valuable information--passenger data and customer social media information.
• Analyzing data in real time.
• Visualizing the data and output.
• Making a speed and flexibility to handle a high volume data of unprecedented scope.
• Uncovering valuable correlations.

Predictive capability

We define predictive capability as an organization’s overall ability to use the analysis results for prediction at each organizational level. Predictive capability is the capability that use the results for forecast use.

According to our content analysis results, this capability may include:

• Helping find the trends in a specific industry.
• Proactively identifying and unlocking a wide variety of secrets via predictive analytics.
• Mining the internet to yield new insights on customers, optimize search, or personalize web experiences.
• Extracting valuable insights through deep analytics.
• Forecasting the price of commodities to figuring out what consumers want, what the company will respond.
• Guiding the company’s action.
• Revealing a business opportunity or a threat.
• Spotting new opportunities, do things more efficiently or identify problems early.

Decision support capability

We define decision support capability as an organization’s overall ability to use big data analysis results to support managerial decision-making at each organizational level.

Most executives find they don’t have enough structured or unstructured data to support their decision-making when the current problems are going to get more and more complicated (Toguri, 2012). Decision support capability is the capability that use the analysis results to make the business decision better (McGlinchey, 2013).

According to our content analysis results, this capability may include:

• Presenting the analysis results to senior executives in a compelling and realistic way.
• Processing unstructured data (e.g., social media data) for decision-making.
• Conducting a real-time or near real-time decision-making.
• Enabling faster and better business decision-making.
• Extracting information and providing actionable insights for decision-making.
• Leading to business decisions by real-time analysis results.

CONCLUSIONS AND IMPLICATIONS

We conclude that big data serves as a business solution to create business value. The key to
making it unique capabilities of big data, including analytical capability, predictive capability, and decision support capability.

Based on the current literature, this research explored big data capabilities through content analysis. The findings revealed three types of big data capabilities: analytical capability, predictive capability, and decision support capability. This research will help practitioners gain a better understanding of how big data contribute to create business values for organizations.

There are several limitations to this study. First, the generalizability is limited because big data capabilities are various. It is hard to recognize all the capabilities based on our basket journals. Second, the goal of this study is to systematically explore big data capabilities through content analysis, and try to discuss the relationship between big data capabilities and business value. However, empirical research is needed to investigate the relationship between big data capabilities and business value. Case studies may also contribute to the reliability and validity of our research findings.
REFERENCE


Sun et al.  

Big Data Capabilities in Creating Business Value


Shiwei Sun, Casey G. Cegielski, Lin Jia, and Dianne J. Hall. (2014). Understanding the Factors Affecting Organizational Adoption of Big Data in Developing Countries, *Proceedings of Decision Sciences Institute Annual Conference (DSI)*, Tampa, FL.


**APPENDIX 1: 95 ARTICLES USED IN CONTENT ANALYSIS**


