Innovating Using Big Data: A Social Capital Perspective

Onkar Malgonde
University of South Florida
omalgonde@usf.edu

Anol Bhattacharjee
University of South Florida
abhatt@usf.edu

Abstract

This paper examines if and how big data initiatives can help firms realize innovative outcomes, and the
type of capabilities that firms need to maximize their return from such initiatives. We employ social
capital as a theoretical lens to examine this problem by augmenting this theory with recent development
in the absorptive capacity literature. We posit a research model and a set of hypotheses for future
empirical testing.

Keywords

Innovation, Big Data, Social Capital Theory, Absorptive Capacity, Knowledge Management.

Introduction

In today’s competitive business environment, firms face tremendous pressure to continually innovate and
optimize their processes, products, and services. However, the process by which firms can continually
innovate remains unclear. While innovation has traditionally been viewed as a function of internal
research and development processes, firms are increasingly relying on ideas or information sourced from
external sources such as their customers, suppliers, and business partners for identifying new product or
service features (Morrison et al. 2000). As an example, Walmart’s legendary collaborative planning,
forecasting, and replenishment (CPFR) system, that provided the company with a distinctive competitive
advantage in the highly competitive discount retailing industry, was sourced from its supplier Proctor &
Gamble. Walmart, in turn, passed on this innovation to its suppliers to help them improve their
manufacturing and distribution efficiency.

While many firms have access to information, as a resource, from external sources, they are not equally
capable of identifying useful information, absorbing it, assimilating it, or applying it to improve
performance (Kogut and Zander 1992). The idea of resources and capabilities stems from the resource-
based view of the firm (Barney (1991) and dynamic capabilities theory (Teece et al. (1997) in the
management literature. Today, many firms are increasingly investing in “big data” as an information
resource and big data analytics as a valuable capability (McGuire et al. 2012). Big data is defined as “data
sets whose size is beyond the ability of typical database software tools to capture, store, manage, and
analyze” (Manyika et al. 2011:1). It includes structured internal business data generated from typical
business transactions, such as sales, customer, and inventory data, and non-transactional and
unstructured external data such as customers’ internet search behaviors, clickstreams, social media
interaction, and financial transactions, as tracked by search engines such as Google, social media
companies such as Facebook, mobile applications, Internet service providers, banks, and credit card
issuers. Big data is usually characterized with three V’s: volume, velocity, and variety. Volume refers to the
fact that the data is typically in the petabyte range (1 petabyte equals 1 million gigabytes). Velocity implies
that the data is generated as a very fast pace - 90% of this data has been created over the past two years
alone (IBM 2012). Variety indicates that it includes both structured data (from relational databases) and
unstructured data (clickstreams, social media behaviors, etc.). Big data analytics refers to the collection,
analysis, and generation of insights from this diverse range of data from multiple sources. Strong demand for heterogeneous external data has spawned companies such as Topsy Labs and Gnip that provide such data on demand (either in bulk or analyzed form) to firms, industries, researchers, and so forth.

It has been suggested that firms can leverage big data capability to discover new and useful insights about their products, processes, customers, markets, and competitors. For example, Vera Bradley, a handbag accessory firm, analyzed customer purchase histories to segment its customer base for targeted e-mail marketing. This segmentation yielded an 101% increase in click through rate, 275% increase in conversion rate, with overall 63% lesser emails sent (Thau 2014). Furthermore, the firm analyzed consumer demand, competitor prices, and margins to strategically determine its product prices (Keenan 2014). In another example, the Industrial and Commercial Bank of China (ICBC), in its attempt to increase its business and customer base, used big data analysis to identify strategic locations for new branches and generate individual branch configuration, resulting in an increase in customer deposits at its Suzhou facilities by USD 1.04 billion (Bell 2013). In a third example, Roadnet Technologies Inc. used real time data from more than 100 million vehicles and mobile devices, and combined it with other data such as weather reports, historical traffic reports, and sporting events to help trucking companies save fuel, time, and costs (Rosenbush 2013).

At the same time, big data is posing unique challenges for firms in terms of their ability to process, manage, and extract value from this ever-growing volume of data. Businesses today create 2500 petabytes of data every day (IBM 2012), and this volume is expected to double every 1.2 years (KnowWPC 2012). This enormous influx of data is causing a severe strain in firms’ IT infrastructures, as evidenced from a recent survey, where 55% of executives said that data is slowing down their IT systems (Avanade 2010). Firms also acknowledge that poor data management can cost 15% to 25% of their operating revenue (Luckie 2012). Given the significance and challenges of big data for organizational performance and competitive advantage, his paper examines whether and how big data can help firms generate innovative outcomes. In this paper, we view innovative outcomes as actionable insights such as problem identification or solution discovery, in contrast to generating new product or services.

To explore the role of big data, we employ the social capital perspective (Nahapiet and Ghoshal 1998). Social capital theory is relevant to our study because it emphasizes the importance of a firm’s social network in the procurement of external information, and the importance of this information in the production of innovations. This theory also elaborates the process by which information is converted to innovative outcomes. This analysis is used to postulate theoretical propositions and formulate a research model for future empirical testing. Theoretical and practical implications of our propositions are discussed.

**Literature Review**

**Knowledge Management**

The role and management of information in organizations has been examined in the knowledge management (KM) literature. In one of the earliest works in this area, Nonaka and von Krogh (2009) distinguished between explicit knowledge (knowledge that can be transferred) and tacit knowledge (experiential knowledge that cannot be transferred) and described four processes by which tacit knowledge can be made explicit and vice versa (internalization, combination, externalization, and socialization). In the information systems literature, Alavi and Leidner (2001) differentiated between data as raw numbers, information as processed/interpreted data, and knowledge as personalized information, and examined the role of knowledge management systems (KMS) in four phases of KM: knowledge creation, knowledge storage/retrieval, knowledge transfer, and knowledge application. the role of in these phases . Focusing on knowledge transfer, Wasko and Faraj (2005) observed that people share knowledge with strangers in an electronic network if they believe that such sharing can enhance their professional reputation. In the context of knowledge application, Kankanhalli et al. (2005) reported that organizational employees’ use of electronic knowledge repositories (EKR) is shaped by their perceived costs and benefits of EKR use.

KM research provides evidence that knowledge (especially, explicit knowledge) can be shared between participants on a social network. However, it does not elaborate how such knowledge sharing contributes
to organizational performance. Our research complements KM research by (a) focusing specifically on the sharing of big data through a social network, that, to the best of our knowledge, has not yet been explored in KM research, and (b) by elaborating the process by which knowledge sharing contributes to innovative outcomes in an organization. The next section presents a theory to further elaborate how knowledge sharing leads to organizational innovation.

**Social Capital Theory**

Nahapiet and Ghoshal (1998) suggested that social capital is the primary driver of intellectual capital, which is key to firm innovations. Social capital is defined as the “sum of the actual and potential resources embedded within, available through and derived from the network of relationships possessed by the firm” (Nahapiet and Ghoshal 1998:243). Initial ideas on social capital were developed by Pierre Bourdieu and Robert Putnam (Siisiäinen 2000). Bourdieu discussed the importance of building social capital in the harmonious functioning of a pluralistic society, characterized by conflict, social agency, and political intervention. Putnam described the formation of social capital in terms of three core components: a social network, social values (especially trust), and social norms or obligations.

Nahapiet and Ghoshal (1998) applied social capital contexts to the firm level of analysis. At the firm level, social capital may relate to information of potential value derived via a social network of firms, competitors, business partners, and so forth. Information derived from such sources may include customer preferences, competitive trends, new technologies, emergent services, and the like that may generate new insights for improving firm performance. The utility of the social capital framework has been empirically validated by Tsai and Ghoshal (1998), who found that flow of information in a social network and ability to absorb knowledge are related to innovation at the business unit level within firms. Tsai (2001) reported that not only is the firm’s social network important for innovation, but the firm’s position within the network is also important. Ahuja (2000) noted that the nature of ties within a network influence the sharing of information. In a longitudinal study of firm innovativeness, Corey (2010) reported that access to diverse information through a firm’s network increases its ability to generate exploratory innovations. Key constructs in social capital theory, with their constituent dimensions and conceptual definitions, are listed in Table 1.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Conceptual Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social Capital</td>
<td>Network and the assets that may be mobilized through that network (Nahapiet and Ghoshal 1998).</td>
</tr>
<tr>
<td>Structural Capital</td>
<td>The properties of the social system and of the network of the relations as a whole (Nahapiet and Ghoshal 1998)</td>
</tr>
<tr>
<td>Cognitive Capital</td>
<td>Resources providing shared representations, interpretations, and systems of meaning among parties (Nahapiet and Ghoshal 1998)</td>
</tr>
<tr>
<td>Relational Capital</td>
<td>Relationships firms have developed with each other through a history of interactions (Nahapiet and Ghoshal 1998)</td>
</tr>
<tr>
<td>Big Data Readiness</td>
<td>State of the firm to adopt and apply Big Data insights (LaVelle et al. 2011)</td>
</tr>
<tr>
<td>Absorptive Capacity (ACAP)</td>
<td>The ability to value, assimilate, and apply new knowledge (Cohen and Levinthal 1990)</td>
</tr>
<tr>
<td>Exchange Capability (Potential ACAP); includes:</td>
<td>Capability of the firm to value and acquire external knowledge (Zahra and George 2002).</td>
</tr>
</tbody>
</table>
Table 1. Conceptual Definitions

Dimensions of Social Capital

Nahapiet and Ghoshal (1998) posited that social capital has three dimensions: structural, cognitive, and relational. The structural dimension refers to the properties of a social network, such as the types, configuration, strength and appropriability of ties within the network. Network ties relates to ‘who you know’ in the network. The greater and varied is the network composition, the greater is the influx of information. Network configuration deals with the positioning of the firm within the network. Appropriability of a network enables members to use the network for other purposes like family, religious beliefs, etc. This is an important property since not all network ties are established for professional purposes.

The cognitive dimension refers to a system of representation and interpretation shared by parties exchanging information in a social network. It includes common beliefs, shared knowledge and codes, shared narratives, and customs that help the recipient understand information communicated by its network partners and evaluate its value within a proper context. Shared codes and languages include the use of specific technical jargon understood by both parties involved in information exchange. Shared narratives refers to the use of stories, myths, or metaphors that can be used by firms to make sense of the communicated information (Nahapiet and Ghoshal 1998).

The relational dimension refers to the trust, norms, obligations, and identification that define the relationships between firms involved in information exchange. Trust refers to the faith a firm has in the information that is communicated to it by its network partners. If the information is trusted, it is more likely to be accepted and assimilated in internal operations. Norm is the shared system of collective decision-making and behavior in a network (Coleman 1990). Norm allows the network to evaluate and establish an accepted system for information exchange among its members. Obligations are implicit expectations that network partners will share information in exchange them receiving information from other partners. Identification allows a firm to associate itself with other firms based on common interest,
goals, industry or contracts. This collective existence strengthens firms’ propensity and frequency of sharing information.

**Exchange and Combination**

Nahapiet and Ghoshal (1998) posited that the two processes that converts social capital into intellectual capital (innovation) are exchange and combination. Exchange is the process of acquiring new information from external sources, while combination is the process of synthesizing acquired information with and prior knowledge for purposes of exploitation or application. Nahapiet and Ghoshal (1998) also suggested that for successful exchange and combination, four conditions must occur: access to parties for combining/exchanging intellectual capital, anticipation of value, motivation, and combination capability.

Exchange is a precondition for combination to occur. Exchange validates external information, weights it based on relevance, and reads it for combining it with preexisting knowledge (Moran and Ghoshal 1996). Prior knowledge, also, is therefore a prerequisite for combination to occur. Prior knowledge may include understanding of a firm’s processes, problems, and opportunities, which can provide the baseline and the opportunity for integrating externally acquired information. Prior knowledge is cumulative and built through numerous iterations of combining external knowledge with existing knowledge, which may then lead to a cumulative set of innovative outcomes.

**Absorptive Capacity**

Nahapiet and Ghoshal (1998) noted that firms must possess combinative capability to successfully exchange and combine external information with internal knowledge to create innovative outcomes. This concept has been popularly labelled absorptive capacity in the literature, defined as “the ability [of firms] to value, assimilate, and apply new knowledge” (Cohen and Levinthal 1999), and has been used to study a wide range of organizational phenomena.

Seeking greater clarity about the domain and operationalization of the construct, Zahra and George (2002) defined absorptive capacity as “dynamic capability pertaining to knowledge creation and utilization that enhances a firm’s ability to gain and sustain competitive advantage” (p. 185), and reconceptualize it as consisting of two subsets of potential and realized absorptive capacities. Potential absorptive capacity (PACAP) refers to the capabilities of the firm to acquire and assimilate, while realized absorptive capacity (RACAP) refers to firms’ capabilities of transforming and exploiting knowledge.

PACAP consists of two sub-dimensions: acquisition and assimilation. Acquisition is the firm’s capability to identify and acquire external knowledge. This includes valuing external knowledge for its relevance to internal operations. Assimilation allows a firm to analyze process, interpret, and understand the knowledge obtained from external sources. Since the exchange process in social capital theory relates to acquiring and assimilating new, external knowledge, PACAP can be viewed as the capability that helps firms exchange information from external sources. Therefore, we refer to this capability as exchange capability.

Drawing on Kogut and Zander (1992) definition of combinative capability as a firm’s ability to synthesize and apply existing and acquired knowledge, Zahra and George (2002) defined realized absorptive capacity (RACAP) as the ability of the firm to leverage the knowledge that has been absorbed. Both these definitions emphasize the synthesis and application of acquired and prior knowledge. Furthermore, RACAP consists of two sub-dimensions: transformation and exploitation. Transformation is the firm’s capability to develop and refine the routines that facilitate combining existing knowledge with the newly acquired external information, while exploitation is the application of the combined knowledge to a firm’s internal goals. Therefore, combinative capability is the capability that enabled the combination process in social capital theory.

**Research Hypotheses**

Synthesizing Nahapiet and Ghoshal’s (1998) social capital theory with Zahra and George’s (2002) reconceptualization of absorptive capacity, we posit a research model to explore how big data can be acquired and applied to a firm’s innovative activities. This model is illustrated in Figure 1 (next page) and described next.
Social Capital

Just as social capital refers to the network and assets that may be mobilized through that network (Nahapiet and Ghoshal 1998), in a big data context, users of big data must rely on a network of partners to provide them with the data and/or the analytic capability to generate insights from big data to generate innovative outcomes. The greater, diverse, and appropriate is the network, the more alternatives and information can be obtained from external sources for firm use.

The structural capital concerns the configuration of the network (Nahapiet and Ghoshal 1998), which in a big data setting, refers to the network of big data vendors or partner firms that can provide external data in a timely manner. This network can not only facilitate the exchange of big data, but may also generate tips and referrals about possible avenues of future data exchange or collaboration with other firms. These expectations lead us to hypothesize:

Hypothesis 1a: A firm’s structural capital has a positive effect on its exchange of big data with external partners.

Cognitive capital deals with the shared representations, interpretations, and meanings that are shared within the network (Nahapiet and Ghoshal 1998). Such terminologies and technical jargon associated with big data include technical terms such as online analytical processing (OLAP), drill-down analysis, NoSQL, Hadoop, and data mining. The use of such technical terms makes it possible to succinctly communicate between information sharing partners, and therefore facilitates the exchange of big data. Hence, we hypothesize:

Hypothesis 1b: A firm’s cognitive capital has a positive effect on its exchange of big data with external partners.

Relational capital related to the relationships among firms within a social network that have developed over years of interaction through mechanisms such as trust, norms, obligations and identification (Nahapiet and Ghoshal 1998). In the big data context, trust between network partners facilitates the
process of data exchange by reducing any friction associated with such exchange. The presence of norms, such as standard data sharing practices, and obligation also enabled the data exchange process. Hence, we posit:

**Hypothesis 1c**: A firm’s relational capital has a positive effect on its exchange of big data with external partners.

### Exchange

Externally sourced big data is of little use if it is not combined with preexisting internal data for generating innovative insights for the firm through synergy. Since the exchanged data is often in varying quantity, quality, and format, the combination process may include cleansing, formatting, structuring and readying the external data for integration with internal data. However, the combination process is predicated on the prior completion of the exchange process, per Zahra and George (2002), which leads us to hypothesize:

**Hypothesis 2**: A firm’s exchange of big data has a positive effect on its combination of big data with preexisting knowledge.

### Combination

Social capital theory says that the exchange and combination processes culminate in the formation of intellectual capital such as new product or service ideas, new processes, or new business models. The generation of innovative insights is also the goal of most big data initiatives. For example, in ICBC’s case, the bank combined external geographic and demographic databases with its internal databases to produce insights about where to locate branches and how to staff these branches. This observation leads us to our hypothesis:

**Hypothesis 3**: A firm’s combination of external and internal data has a positive effect on its innovative outcomes.

### Exchange Capability

Exchange capability of an firm is its ability to value and acquire external knowledge that is useful for its operations (Zahra and George 2002). Firms lacking such capability may face significant odds in internalizing big data, even if they are presented with such data by external data sources by virtue of their social capital. On the other hand, firms with high levels of exchange capability are more likely to be successful in appreciating the value of big data available to them, and internalizing it for generating insights. This expectation leads us to posit:

**Hypothesis 4**: A firm’s exchange capability has a positive effect on its exchange of big data with external firms.

### Combinative Capability

Similar to exchange capability, firms also need combinative capability to effectively combine big data acquired from external sources and preexisting internal knowledge about their operations, processes, or resources. Firms that are lacking such capability will struggle to exploit the combined data to gain critical insights or innovative outcomes, even if they are successful in procuring big data from partners through an exchange process. For example, Kohl’s Corporation, a department store chain, is testing new targeted marketing technique where customers will be provided with real-time online coupons on their mobile phones during their store using live information about their presence in certain sections of the store and product browsing behavior, integrated with their preferences and interests as previously stored in their online user profiles (Thau 2014). Such a high level of combinative capability is expected to increase the probability of sales at Kohl’s, leading us to postulate:

**Hypothesis 5**: A firm’s combinative capability has a positive effect on its innovative outcomes.
**Big Data Readiness**

Lastly, we define *big data readiness* as the extent to which an adopting firm is willing and able to source and apply big data. LaVelle et al. (2011) identified three categories of readiness: aspirational, experienced and transformed. An aspirational firm is at an initial stage of building big data capabilities and vision. Experienced firms have invested in the process of sourcing external data, combining it with existing data, and analyzing the data. Transformed firms have already realized considerable gains from their big data initiatives. Naturally, firms that are more ready (i.e., transformed firms) are more likely to realize their expected innovative outcomes than those that are less ready. This expectation leads to our last hypothesis:

*Hypothesis 6: A firm's big data readiness has a positive effect on its innovative outcomes.*

Hypotheses H1-H5 represents our application of social capital theory to the context of big data. These hypotheses describe how transfer and combination of big data may lead to innovative outcomes for firms. Hypothesis 6 presents a new construct, firm’s readiness, and its role in big data initiatives.

**Methods**

Our proposed social capital theoretic model of big data innovation will be tested via a survey of the top executives of around 1000 US based firms. We will try to identify publicly available lists of big data adopter firms as the sampling frame for our survey. We have already developed measures for each of our constructs of interests that are not included in this paper for lack of space, but are available from the authors upon request. This instrument will be pilot tested with a convenience sample of big data users (MBA and EMBA students at the authors’ university) prior to survey administration. Following multiple reminders, we expect to have a response rate of 20-25%, which will be adequate for statistical analysis.

**Conclusion**

Our study has several theoretical implications for research. First, to the best of our knowledge, this is the first rigorous academic study to examine the impact of big data initiatives for firms and the types of capabilities needed to extract the most value from such initiatives. Second, our study will demonstrate how to theorize the nature and impacts of big data initiatives and its relationships to firm innovations. Although we used social capital theory as the starting point for our analysis, future research may examine other theories and/or employ a multi-theoretic approach to gain a better understanding of the role and impacts of big data in firms. Third, given that prior research has focused primarily on the role of information systems on firm efficiency and/or productivity, our study may draw attention to the potential role that information systems may play in building firm’s innovative abilities.
References


