Role of Trust, Privacy Concerns and Data Governance in Managers' Decision on Adoption of Big Data Systems

Selen Balkaya

Bogazici University, Istanbul, Turkey

This paper investigates the role of trust, privacy concerns, and data governance on managers' intention to use big data systems. In literature, trusting beliefs, such as functionality, helpfulness, and reliability were found to be antecedent of trust in technological artifacts. Notice, access, choice, and security principles were found to be crucial in eliminating privacy concerns. On the other hand, this paper focuses on data storage and data collection which have been significant criterion for managers in evaluating companies' data governance policies. A model depicting the relationships amongst all these factors and their relation to users' intention to adopt big data systems and a scale was proposed in the paper.

Keywords: big data systems, privacy concerns, data governance, trust, technology acceptance

Introduction

Big data systems cover some of the youngest technologies of our age with plenty of unknowns. In its simplest form, big data are collection of data sets so large and complex that traditional database and software applications are not capable to process it. In addition to the underlying data processing and analytical technologies, business intelligence and analytics include business-centric practices and methodologies that can be applied to various high-impact applications, such as e-commerce, market intelligence, e-government, healthcare, and security (Chen, Chiang, & Storey, 2012). Wide applications of big data systems and opportunities associated with data and analysis make big data systems as one of the major technology trends in the 2010s. With the evolvement of new technologies around big data, managers face decisions to make on adoption of technologies.

Trust is one of the primary reasons why users intend to use a specific technology. Whether the object of trust is another person or an information technology, one trusts the other to the extent that one chooses to depend on the other and reconciles away fears by being willing to become vulnerable to the other without controlling the other (Mayer, Davis, & Schoorman, 1995). In general, Internet research provides evidence that trust in another actor (i.e., a Web vendor or person) and/or trust in an agent of another actor (i.e., a recommendation agent) influences individual decisions to use technology (Mcknight, 2011). As a result, trust is expected to have a significant role on adoption of big data systems as well.

Selen Balkaya, master, Department of Management, Faculty of Economics and Administrative Sciences, Bogazici University, Istanbul, Turkey.

Correspondence concerning this article should be addressed to Selen Balkaya, Department of Management, Faculty of Economics and Administrative Sciences, Bogazici University, Bebek, Istanbul 34342, Turkey.

On the other hand, privacy is one of the main issues in collection of data. With massive amounts of our personal data now being routinely collected and stored, privacy breaches are almost inevitable and there are few effective controls over how data are used or secured (The Guardian, 2014). Also, it is unclear how all this collected data will be used in the near or long-term future. Therefore, creation of standards in big data systems is essential to better protect data. Users' concerns based on the privacy of the data collected are expected to affect their trust in the big data systems and hence will have a role in their adoption decision.

With the evolvement of big data, another challenge that managers face is data governance programs. Protection of data and information systems is not a small task. Companies need to adapt new data governance policies for all data storage and collection processes. This includes understanding and meeting legislative and regulatory requirements. Data governance of the companies is expected to affect users' trust on big data systems and applications.

This study examines the role of trust, privacy concerns, and data governance in users' (users in managerial level) intention to use big data systems by integrating trust, privacy, and data governance constructs into Technology Acceptance Model (TAM).

Theoretical Background

Big Data

Big data are a fairly newborn concept that covers new set of techniques and technologies requiring new forms of approach to extract knowledge from massive amount of data.

Dhar (2013) explained big data as data science which are the systematic study of the extraction of knowledge from data. Data science aims to discover and extract actionable knowledge from the data, that is, knowledge that can be used to make decisions and predictions, not just to explain what's going on (Dhar, Jarke, & Laartz, 2014).

Therefore, big data have different technological aspects that are new to users. As Gartner Inc. defines: Big data are high-volume, high velocity, and high-variety information assets that demand cost-effective and innovative forms of information processing for enhanced insight and decision-making. All these new technological aspects require new skills in the organizations and demands managers to face operational and strategic challenges, such as reevaluating company's information sources; human resources' skills in analyzing such data; security, competency, and reliability of current technological systems, data governance of the company, and other possible sensitivities, such as privacy and accessibility that emerge with the nature of data.

As Chen et al. (2012) explained the emergence of business analytics and intelligence in their study; key characteristics and capabilities change with evolvement of data science and new challenges and opportunities associated with data science arise. In summary, data science technologies and applications were mostly structured. Data were collected by companies through various legacy systems, and often stored in commercial relational database management systems (Chen et al., 2012).

Web intelligence, web analytics, and the user-generated content collected through Web 2.0-based social and crowd-sourcing systems (Doan, Ramakrishnan, & Halevy, 2011; O'Reilly, 2005) have led to a new era in data science in the 2000s, centered on text and web analytics for unstructured web contents (Chen et al., 2012).

As Chen et al. (2012) stated in their study that next coming era is Web 3.0 (mobile and sensor-based) era which underlies mobile analytics and location and context-aware techniques for collecting, processing, analyzing, and visualizing such large scale and fluid mobile and sensor data are still unknown.

Data become more unstructured and massive with evolvement of data collection technologies. Privacy of people whose data are collected has also been an issue which has no standards and mature control mechanisms yet. Therefore, sharing and using the data for companies became a sensitive issue. On the other hand, technologies related to big data are various and their competency, maturity, reliability, and usability are questionable. Data governance policies of the companies also became crucial. While opportunities arise with big data era cannot be ignored, managers need to face the challenges around it by evaluating both company resources and the technology itself meticulously.

Trust in Technology

Because of risks and uncertainty associated with new technologies, users hesitate to start using them immediately. Trust in technological artifacts has been studied in several studies. Many researchers found support for attribution of trustworthiness to technology.

For instance, Benbasat and Wang (2005) in their study stated that people respond socially to technological artifacts and perceive that they possess human characteristics (e.g., motivation, integrity, and personality). In particular, research findings have demonstrated that components of trust in humans and in technological artifacts do not differ significantly. This indicates that people not only utilize technological artifacts as tools, but also form social and trusting relationships with them (Benbasat & Wang, 2005).

On the other hand, to differentiate trust in technology from trust in people, McKnight, Carter, Thatcher, and Clay (2011) in their study stated that their position is that trust situations arise when one has to make oneself vulnerable by relying on another person or object, regardless of the trust object's will or volition. Perhaps the most basic dictionary meaning of trust is to depend or rely on another (McKnight & Chervany, 1996).

Trust (more accurately called trusting beliefs) means beliefs that a person or technology has the attributes necessary to perform as expected in a situation (Mayer et al., 1995). Three trusting beliefs—competence, benevolence, and predictability/integrity—in human were adapted to technology in McKnight et al.'s (2011) study.

Functionality. With technology, users consider whether the technology delivers on the functionality promised by providing features sets needed to complete a task (McKnight, 2005). Functionality belief in technology corresponds to competence belief in humans. Functionality belief is studied in big data systems as well, since big data systems cover technologies which need have features different then traditionally used data processing technologies that are not competent to process massive amounts of collected data.

Helpfulness. Users do hope that a technology's help function will provide advice necessary to complete a task (McKnight, 2005). Helpfulness belief in technology corresponds to benevolence belief in humans. Since big data systems are totally new to companies, all the employees who are expected to use these systems in company will need help at different points of time to finish the task they have. Managers need to evaluate if the big data systems provide adequate, effective, and responsive help that enables users' task performance

Reliability. Trustees are expected to be consistent, predictable, or reliable (Giffin, 1967; McKnight, 2005). With people, predictability refers to the degree to which an individual can be relied upon to act in a predictable manner. This is risky due to peoples' volition or freedom to choose (McKnight et al., 2011). Although technology has no volition, it still may not function consistently due to built-in flaws or situational events that cause failures (McKnight et al., 2011). Big data systems reliability is also crucial to managers since they will

use the system output for decision support. Failures in processing data will end up in fallacious output which may lead to serious business failures. As a result, it is believed that trusting beliefs in technology that are adopted from trusting beliefs in humans, will also be valid for users to develop trust in big data systems before intend to use these systems.

Privacy Concerns

With big data applications, organizations will start storing more sensitive information which includes personal data from social media, transactions, sensors and locations, etc. This information is important for organizations in order to provide customized services to their customers. According to Liu, Marchewka, Lu, and Yu (2004), even customers benefit from sharing personal information, privacy concerns also become an important issue because of the organizations are able to access, share, and store this information.

Fukuyama (1996) stated that trust is important in a high tech environment. Hoffman, Novak, and Peralta (1999) suggested that the main reason for many people not to provide personal information to a vendor is due to a fundamental lack of trust. It can be concluded that a person's behaviors are effected by that person's attitudes.

To protect privacy, organizations are the responsible ones who should care about the users' privacy concerns via policies. The Federal Trade Commission (FTC, 2000) proposed that fair information practices include four dimensions in their report to US Congress:

Notice—providing notice to consumers that personal information is collected prior to the collection of that information. Data collector should ensure the confidentiality.

Access—providing access to the data that is collected about them by ensuring adequate authentication methods.

Choice—providing a choice to allow an organization to use or share information collected about consumers. Data collectors must afford consumers an opportunity to consent to secondary uses of their personal information.

Security—providing guarantee that personal information will be protected against unauthorized access, use, or disclosure.

FTC aims to increase protection of privacy with those principles. What this paper wants to investigate is how enabling that privacy will affect the decision makers' trust to big data applications in an organization. In other words, caring about others' privacy may be an antecedent to build trust.

Data Governance

Big data propose a great area for new innovations and analytics, but they are mandatory for organizations and those who are responsible to develop internal governance mechanisms in order to prevent any unsecured access to personal data (Gold, 2014). Forrester Consulting (2013) conducted a survey to the senior executives of organizations which deal with big data and asked how they govern big data. The report includes concerns of decision-makers about securing and protection of sensitive data with the addition of masking and quality of data. This result shows that the data governance is increasingly important for organizations and they should take protection and privacy as main issues of big data.

Dean, Kalapesi, and Rose (2013) also stated that there is a convergence between the agendas within organizations for data management and data protection and privacy.

This paper positioned data governance as a mediator factor between privacy and trust to see how decision makers' adoption of big data effected. "Data storage" and "data collection" will be analyzed as two critical factors of data governance.

Data Storage

Amount of data that organizations store and process growing exponential by time (Tallon, 2013). According to a study by Center for Large-Scale Data Systems, especially in some sectors, like telecommunications, healthcare, and transportation data growth per annum is around 100%, reported by managers (Beath, Becerra-Fernandez, Ross, & Short, 2012). So, growing data bring more responsibility to decision-makers in organizations while planning to implement big data systems.

One of the biggest challenges is to improve governance policies and structures according to this fast growing environment because data governance is not a luxury for companies anymore (Tallon, 2013). According to a published white paper by International Data Corporation (IDC), big data approach has some advantages due to its unstructured and informal nature, but using sensitive data brings more regulatory requirements to enterprises. Traditional database systems provide various security policies and they are align with governance for data protection but big data software generally has no pre-built structures like that so, it is enterprises' responsibility to provide enough protection mechanisms for securing data like other database management systems (Villars, Olofson, & Eastwood, 2011).

Data Collection

Big data are used to make decisions about individuals and one of the key requirements is how this personal data collected (Gold, 2014). Gold (2014) also stated that organizations should be transparent to the customers while collecting data and should explain how their data will be used, processed, and stored according to the protection principles.

According to Information Commissioner's Office Report (2014), the transparency of personal data will be fair for individuals if they will be notified at the collection of data and should be informed the processing purpose. If organizations want big data applications to comply with data governance and protection principles, then data collection process should be considered by decision makers as well.

Gold (2014) also expressed that if a company misuse the personal data then they can be a subject of a media story which will conclude a trust and reputation loss for the organization, since it is very easy for consumers to share their views to the world. Moreover, being seen as a trustworthy company in the consumers' perspective can bring competitive advantage as well. Rose, Barton, Souza, and Platt (2013, p. 10) gave an example:

Two companies in the same industry, using the same data in the same new ways, will likely achieve fundamentally different results, with the more trusted organization able to access at least five to ten times more data than the less trusted one. This, in turn, will lead to better online recommendations, more accurate targeting, faster development of new products and services, and other tangible benefits to consumers. This is the trust advantage.

It can be understood that data collection process can be very crucial for organizations in this situation as well.

Technology Acceptance Model

TAM provides a theoretical framework to explain user acceptance of information technology products/systems. TAM (Davis, 1989) suggests that when users are presented with a new technology, a number

of factors influence their decision about how and when they will use it. TAM (Davis, 1989) consists of perceived usefulness (PU), perceived ease of use (PEOU), attitude toward using, behavioral intention to use, and actual system use. In TAM (Davis, 1989), PU and PEOU are the two most important determinants of system use. However, numerous empirical studies have found that TAM consistently explains a substantial proportion of the variance (typically about 40%) in usage intentions and behavior (Venkatesh & Davis, 2000). Therefore, in this study, keeping the nature of big data systems in mind, this study aims to address how variables, such as privacy, data governance and trust may influence usefulness, ease of use, and intention to use.

This study aims to reveal the relative importance of privacy, company's data governance, trust with other use antecedents in TAM (PU and PEOU), in users' adoption of big data systems. By testing the proposed extended TAM model, this paper aims to empirically examine the significance of relationships amongst the constructs in the context of big data systems.

Research Model

Objective of this paper is to propose a model which aims to explain how trust effects the big data adoption and whether "privacy" has any effect on "trust" with the mediator effect of "data governance". Privacy can be an important antecedent to build the trust for decision makers and with this model this paper also suggests that data governance concerns also effect executive's intention to adopt big data applications. In a further research, this model can be tested with the scale (see Appendix) proposed in this research.

Based on the arguments discussed previously, following research model and hypotheses are proposed in Figure 1.

Figure 1. Proposed research model.

Hypotheses:

H_{1a}: PEOU of a big data system will positively affect users' intentions to use the system.

H_{1b}: PEOU of a big data system will positively affect PU of the system.

H₂: PU of a big data system will positively affect users' intentions to use the system.

H_{3a}: Trust in big data systems will positively affect PEOU of the system.

H_{3b}: Trust in big data systems will positively affect PU of the system.

H_{4a}: Privacy of the big data system will positively affect the trust of users for big data system.

H_{4b}: Privacy of the big data system will positively affect the beliefs about data governance of the company.

H₅: Data governance will positively affect the trust of the users for big data systems.

Methodology and Further Study

This paper reviewed the literature to propose a model and a scale (see Appendix) to test the effect of trust, privacy concerns, and data governance on the managers' intention to use big data systems. In a further study to measure constructs and their relations, and to test the research model, a survey can be administered, using the questionnaire developed in this study.

Since big data are relatively a new area and usage level of big data systems and applications amongst companies can be low, convenience sampling should give us more effective results for the purpose of the study. Survey can be designed in a way that eliminates companies who have never experienced any big data systems or its applications. Survey can be sent to managers who had active role in decision stages of adoption of big data systems in their company.

All of the items used in the survey are on a seven point Likert scale. Factor analysis can be conducted for all variables to evaluate construct validity and internal consistency of the scale. Reliability of items in measuring the constructs in the model should also be assessed with Cronbach's alpha.

Testing of hypotheses that means values and matrix inter-correlations among the constructs will be calculated. Also, Structural Equation Modeling (SEM) method can be used to test the relationships among the variables in the proposed model in a further study.

Conclusion

For this study, technology acceptance and big data literature were reviewed to see how the relations were developed and tested previously.

Trust was found to be one of the primary factors that affect users' intention to adopt a technology. On the other hand, privacy concerns and data governance are major factors that affect users view on and trust which affect their intention to use big data systems as found in literature.

In this study, a model and a scale (see Appendix) were developed based on the literature to test the relation between trust, data governance and privacy concern constructs, and managers' intention to use big data systems. The model was based on one the most frequently used models in technology acceptance literature, Technology Acceptance Model. Trust was added as outer factor that affect users' perception of ease of use and usefulness of big data technologies. Data governance and privacy concerns constructs are believed to have effect on trust construct. All the validated items from literature are used in the scale (see Appendix). In a further study, a survey should be conducted to a selected sample to see how significant is the relation in the proposed model.

References

- Beath, C., Becerra-Fernandez, I., Ross, J., & Short, J. (2012). Finding value in the information explosion. *MIT Sloan Management Review*, 53(4), 18.
- Benbasat, I., & Wang, W. (2005). Trust in and adoption of online recommendation agents. *Journal of the Association for Information Systems*, 6(3), 4.
- Chen, H., Chiang, R. H. L., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big Impact. *Management Information Systems*, 36(4), 1165-1188.
- Davis F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly*, 13(3), 319-340.

- Dean, D., Kalapesi C., & Rose, J. (2013). Unleashing the value of consumer data. *BCG Perspectives*. Retrieved from <https://www.bcg.com/publications/2013/big-data-advanced-analytics-marketing-sales-unleashing-the-value-of-consumer-data.aspx>
- Dhar, V. (2013). Data science and prediction. *Commun ACM*, 56(12), 64-73.
- Dhar, V., Jarke, M., & Laartz, J. (2014). Big data. *Business & Information Systems Engineering*, 6(5), 257-259.
- Doan, A., Ramakrishnan, R., & Halevy, A. Y. (2011). Crowdsourcing systems on the world-wide web. *Communications of the ACM*, 54(4), 86-96.
- Fukuyama, F. (1996). Trust still counts in a virtual world. *Forbes ASAP Supplement*, 1337051, 33-34.
- Federal Trade Commission Report to Congress. (2000). *Privacy online: Fair information practices in the electronic marketplace*. Retrieved from <http://www.ftc.gov/reports/privacy-onlinefair-information-practices-electronic-marketplace-federal-trade-commission>
- Forrester Consulting. (2013). *Big data needs agile information and integration governance*. Retrieved on December 31, 2014 from <http://www.ibmbigdatahub.com/whitepaper/big-data-needs-agileinformationand-integration-governance>
- Giffin, K. (1967). The contribution of studies of source credibility to a theory of interpersonal trust in the communication process. *Psychol. Bull.*, 68(2), 104-120.
- Gold, S. (2014). Big data and data protection paper from ICO. *Journal of Direct, Data and Digital Marketing Practice*, 16(2), 135-137.
- Hoffman, D. L., Novak, T. P., & Peralta, M. (1999). Building consumer trust online. *Communications of the ACM*, 42(4), 80-85.
- Information Commissioner's Office. (2014). Information Commissioner's Annual Report and Financial Statements 2014/15. Retrieved from <https://ico.org.uk/media/about-the-ico/documents/1431982/annual-report-2014-15.pdf>
- Liu, C., Marchewka, J. T., Lu, J., & Yu, C. S. (2004). Beyond concern: A privacy-trust-behavioral intention model of electronic commerce. *Information & Management*, 42(1), 127-142.
- Mayer, R. C., Davis, J. H., & Schoorman, F. D. (1995). An integrative model of organizational trust. *Academy of Management Review*, 20(3), 709-734.
- McKnight, D. H. (2005). Trust in information technology. In G. B. Davis (Ed.), *The Blackwell encyclopedia of management, management information systems* (Vol. 7, pp. 329-331). Malden, MA: Blackwell.
- McKnight, D., Carter, M., Thatcher, J., & Clay, P. (2011). Trust in a specific technology: An investigation of its components and measures. *ACM Transactions on Management Information Systems (TMIS)*, 2(2), 1-25.
- McKnight, D. H., & Chervany, N. L. (1996). The meanings of trust. *University of Minnesota MIS Research Center Working Paper Series*, pp. 96-04. Retrieved from https://www.researchgate.net/publication/239538703_The_Meanings_of_Trust
- O'Reilly, T. (2005). What is web 2.0: Design patterns and business models for the next generation of software. *Social Science Research Network Working Paper Series*. Retrieved from <http://www.oreillynet.com/pub/a/oreilly/tim/news/2005/09/30/what-is-web-20.html>
- Rose, J., Barton, C., Souza, R., & Platt, J. (2013). The trust advantage: How to win with big data. *BCG Perspectives*. Retrieved from http://image-src.bcg.com/Images/The_Trust_Advantage_Nov_2013_tcm9-92206.pdf
- Tallon, P. P. (2013). Corporate governance of big data: Perspectives on value, risk, and cost. *Computer*, 46(6), 32-38.
- The Guardian. (2014). *Little privacy in the age of big data*. Retrieved on December 16, 2014 from <http://www.theguardian.com/technology/2014/jun/20/little-privacy-in-the-age-ofbig-data>
- Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model: Four longitudinal field studies. *Management Science*, 46(2), 186-204
- Villars, R. L., Olofson, C. W., & Eastwood, M. (2011). Big data: What it is and why you should care. *White Paper, IDC*. Retrieved from http://www.admin-magazine.com/HPC/content/download/5604/49345/file/IDC_Big%20Data_whitepaper_final.pdf

Appendix

Variables	Items (1: Strongly disagree, 7: Strongly agree)	Source
Intention to use big data	I am willing to use XXX Big Data Systems to extract knowledge from data collected.	
	I am willing to use XXX Big Data Systems to support company's decision making stages.	
Perceived ease of use	My interaction with XXX Big Data Systems is clear and understandable.	
	It is easy to get XXX Big Data Systems to do what I want it to do.	
	It was easy for me to extract the information using the XXX Big Data Systems.	
	Overall, I found that XXX Big Data Systems is easy to use.	
	Using XXX Big Data Systems enabled me to reach suitable information more quickly.	
	Using XXX Big Data Systems improved the quality of data analysis I perform regularly.	
	Using XXX Big Data Systems made the data collection and analysis tasks easier to complete.	
	Using XXX Big Data Systems enhanced effectiveness of data collection and analysis.	
Perceived usefulness	Using XXX Big Data Systems gave me more control over the data collection and analysis tasks.	
	Using XXX Big Data Systems allowed me to accomplish more data collection and storage and more data analysis than would otherwise have been possible.	
	Using XXX Big Data Systems greatly enhanced the quality of company's decision support systems.	
	Using XXX Big Data Systems supported all the analysis the company use on regular basis.	
	Overall, I found that XXX Big Data System is useful.	
	XXX Big Data Systems have the functionality that the company needs.	
Trust in big data systems functionality	XXX Big Data Systems have the features required for company's activities.	
	XXX Big Data Systems have the ability to do what the company requires it to do.	
	XXX Big Data Systems have the overall capabilities the company needs.	
Trust in big data systems predictability	XXX Big Data Systems behave in a predictable way.	
	XXX Big Data Systems functions in the same way each time we use it.	
	As a work tool, XXX Big Data Systems very predictable.	
	The help function of XXX Big Data Systems provides competent guidance (as needed).	
Trust in big data systems helpfulness	The help function of XXX Big Data Systems provides whatever help user needs.	
	The help function of XXX Big Data Systems provides very sensitive and effective advice, if needed.	
	The help function of XXX Big Data Systems supplies users' need for help through a help function.	
Privacy-notice	I believe the XXX Big Data Systems notifies users about information collection.	
Privacy-access	I believe the XXX Big Data Systems allows users to review and change their personal information.	
Privacy-choice	I believe the XXX Big Data Systems do not release personal information without users' permission.	
Privacy-security	I believe the XXX Big Data Systems gives users a choice before disclosing personal information to third parties.	
	I believe the XXX Big Data Systems make effort to keep users' personal information secured.	
	I believe the company has policies regarding sensitive data protection.	*
Data governance-data storage	I believe the company has necessary tools and applications in order to prevent unauthenticated access to personal data.	*
	I believe the company has roles within organization which are responsible for assessing and resolving privacy risks.	*
	I believe the company is transparent about the purpose of data collection.	*
Data governance- data collection	I believe the company explains how personal data of users will be used for service customization and analysis.	*
	I believe the company will notify users if it changes the purpose of using personal data.	*

Note. * Data governance items are developed via interviews with IS professionals. After doing the pre-test, items can be modified according to the feedbacks given.