Proceedings of Business and Economic Studies

Review Article

Which Aspects of Big Data Usage is Creating Information Security Concern?

Hira Batool, DU Rong

Xidian University, School of Economics and Management, Department of Management Science & Engineering. Xian, China

Abstract: The present research study proposed some of the big data usage perspectives for testing either they have a role in creating information security concern or not. The researchers first dig out some of the theoretical support for filling the gap regarding big data and information security bridge that was previously noted in literature. The present researches approached big data analytics manager in the Pakistani banking industries for validating the proposed model. The data were analyzed using SPSS Andrews approach due to the nature of the research study. The findings revealed that the proposed perspective including perceived benefits, cloud storage, and online behavior monitoring should be tested in the future studies by proposing their indirect affect in the creation of information security issue. The study brings a new aspect in literature of management regarding big data usage practices.

Keywords: *big data; security; cloud storage; behavior; monitoring*

Publication date: October 2018

Publication online: 31st October 2018

Corresponding Author: Hira Batool, drbatoolhira@gmail.com

0 Introduction

Recent investigations from[1] highlighted that in Pakistan the information security concern has gaining importance but strongly ignored from the aspect consumers end. The previous investigation on information security concern was studied from the perspective of knowledge sharing where the role of trust factor found responsible for dealing with the information security concern, but the Pakistani banking industries are more interesting in finding the consumers end information security concerns in the present world of knowledge sharing. On the other hand, the review research studies in the knowledge sharing world like presently big data in the Pakistani business industry highlighted that the use of big data for the improvement of companies' performance is giving rise to information security concern among the consumers end[2], and this is due to the young generation load toward employing 2.0 services. However, the connection between linking big data and information security was not previously proposed, so the present researcher tries to work on this domain to fully express the thoughts in the final dissertation work for proposing better understanding toward consumer end information security concern. In the recent era, the Pakistani society is getting hits with the proliferation of business industry social networking and web 2.0. This enables the individuals and businesses to engage with digital technology on an unprecedented scale producing huge amounts of data, also known as "big data." People are utilizing online medium for voluminous distinct reasons, for example, socialization, free time enjoyment, convenient interaction with friends and family, and also for preserving "social capital." The popularity of online medium has twisted the real level of structured and unstructured information. Definitely, the current focus that "big data" is creating everywhere just due to the accelerated evolution of the online medium. Such as online medium has provided further stations for business and communication, massive volumes, and selection of information are all increasingly being generated for storage in the online medium. The stored data easily usable in the form of statistics by government bodies, business officialdoms, and researchers and marketing houses, etc. Firms in various business sectors such as online banking sector,

entertainment, advertising, and marketing have initiated to give vigilant devotion toward big data phenomenon as one of the main means for the big information[3]. Its likewise critical to remember that even with several benefits of online medium in the big information medium, the ability of organizations to assemble, save, and investigate big data stances security and privacy associated risks for customers. The connections generated in the online medium are obtained by several parties such as government, large institutions, third parties, and customer, and service companies. The increase of social media and the tradition of sharing data also have fueled the growth of online media such as companies' online applications and Google+ facilities for individuals' daily life. The exponential development of online medium has conveyed an exciting emphasis on the privacy concerns of users. Online medium has been afflicted with problems of privacy risks[4]. Due to the role of the online medium in the generation of data, this previously multifaceted issue becomes more complicated. Any user individual might feel unsafe due to the perception of losing control over the confidential data. The environment of the internet, with the user, believes over the quantity of privacy and with the liability of online medium suppliers do have the authority to disseminate the privacy easily. The privacy and protection hitches might be of paramount importance. The environment of social networking, information privacy indicates that the identifiable amount of data congregated by the companies for their own analysis. Such perception creates privacy issues from data extortions such as digital accumulation and inappropriate use of consumer data by third parties. These risks arising perception are coming from the social environment talking about big data usage among companies[5]. The sophistication usage of technology used for the analysis of the huge data produced by the online medium has enlarged significantly during a couple of past years. Furthermore, advantageous and innovative ways of collecting large volume information with speed and variety have appealed the privacy and security issues into the forefront[6]. The present research begins capturing the life through the lens of electronic reality in medium, which makes easier for the people and the organizations for capturing the consumer behavior to improving their services or the targeted goals. Internet giants like Google cloud storage to banking consumers have been criticized for quite a while because of the absence of transparency reporting on what happens to accumulated data. Various hidden tools and methods are in use by online based businesses for collecting big data that are usually hide in the various technical layers[7], which makes the business transparency questionable regarding the stored information. The top of these is several third party based software listed over the online medium like online banking application facilitate in collecting real-time user information serving over the internet. Therefore, concurrent unstructured and structured information through online banking applications, Google+ cloud storage services normally carry privacy dangers. However, although the social medium designed for carrying the role of primary communication, people, especially with the concern of information security might show indifferent behaviors[8]. Hence, the majority of the Pakistani people on online medium might not aware of the risk associated with the doing things online. While[9] reported that some of the people stated perception related to data confidentiality could be leading factor for avoiding the use of big data technology. Others may have undergone privacy invasion and so, may not even consider their data to be confidential[9] becoming apathetic toward their own privacy as time passes. The latest research in the information systems has concentrated toward improving the worth of technology usage by knowing which big data perspective can cause information security concern among the users. Still, the addition in the recent literature failed to report the answer to the question not only from theoretical support but also through the evidence. Batool *et al*., Batool and Rong[1,2], and Ion *et al*.[10] reported that the gap is related to the understanding toward what kind of big data perspective should be studied in reporting information security concern. The present research study tries to raise a few of the big data usage perspective in relation to build its connection with the information security concern from the help of big data analyst in Pakistan.

1 Literature review

The researchers Bulgurcu, Sharma and Crossler[5,6] and Boyd and Ellison[4] added that the online medium is design on the basis of public access layout programs so; it can easily give the sense of risk carrying with it. Hence, all the government ruled bodies or organizations based on the public services can easily access toward this public access layout programs. Majority of the research study based on the privacy testing behavior

contradict with each other due to vary nature of human. Feuchtl and Kamleitner and Ion[9,10] reported people who are comfortable even with sharing of their private information depending on their level of trust on the online medium[10]. Started the vagueness toward the understanding of the use of big data as the result of the online medium on the perception of consumers toward the privacy issues and ultimately can affect trust and engagement. With any doubt, it is vital to observe privacy in relation to large data in the Pakistani context where hacking and security defilements are the majority of the time reported in the online medium. Considerable data have legitimate individuals to extract identified information about the people.[16]

2 Theoretical framework and proposed hypothesis

Kshetri[11] reported that the theory of prospect explained the usage or the adoption of choice on the basis of satisfaction associated with it. The present literature adds in this respect that choice for the private data storage also considers the level of risk associated with the choice of data base[12]. Further broadening the concept by rely on the principle of usefulness. This occurrence has intriguing consequences for individuals' choice to talk about their private info on online medium, as customers can share the feelings about the usefulness of their choice. Often times, people are rationally apathetic toward an effect. In Pakistan Batool *et al.*, Batool and Rong[1,2] highlighted the discussion about the information security as a result of using online banking applications. In line with the previous research findings, Hogben[13] directed that researcher own observations can initiate a link in studying big data aspects. Metzger[14] initiated from the apathy theory perspective which enables individual concern or thoughts as understanding any social existence. In the same way, researchers using online social medium can demonstrate a "non-pathological" absence of attention in data security matters due to the absence of previous supported believe Feuchtl and Kamleitner[9]. Azam[20] said that the Pakistani society is just in believe that privacy matters are currently too early to discuss with the firms until they start collecting information on the basis of take caring of it. Drawing on the prospect concept that is an expansion of expected utility theory and by the terrible apathy concept, we imagine our study model for this research. Within this paper, we explore whether the online applications which are big data drivers, such as online banking services through giving the personal information to banks, storing personal data in banks cloud storage, and being monitored by banks for maintaining the transparency can influence consumer's security belief. We study whether or not these constructs will influence the consumer's inclination toward online medium information security concern. Consequently, the researchers attempt to answer the following research questions: 1. What is the association between privacy issues and its antecedents in the circumstance of big data within internet banking? What is the association of cultural values explained with context to big data within internet banking? To control for an explanation of outcomes as a result of extraneous elements, an earlier study on online medium and data processes notorious a variety of variables that can affect the true security concern of respondents. Assessing the effect of other factors is vital for a research version as it eliminates any troublesome factors[17]. Therefore, for this particular study version, cultural values, and beyond privacy invasion were contained as the moderator factors to find out whether they affect the dependent factor.

2.1 Proposed hypothesis

The privacy apathy dimension is a fairly newer theory in the information systems research[6], highlight it is the momentum of progress toward measuring information privacy concern. Due to the age of Web 2.0 technologies, the gathered and saved millions of time data by different web sites, programs, agencies, and third parties; people may believe that there is no such information security concern. Similarly, a recent poll revealed that almost half of the bank keeping a sharp eye on the consumers' things of doing online. Hence, they consider that they are being tracked throughout the Internet by companies such as "Google and Facebook" (csmonitor.com 2013). Thus, it is safe to hypothesize that users with privacy apathy place lower value and price to their private information and so, care less about data disclosure[20].

> H_1: Big data perspective of online behavior monitoring would have a lesser role in creating an information security concern.

Privacy security belief "is the subjective possibility that consumers think that their personal information is protected as anticipated"[14]. In the online medium, users that demonstrate greater security beliefs are thought to have wise control over private information. Hence, the

more in control over the private information disclosure the more likeness toward the disclosure of private data[19]. Perceived advantages state to a customer's overall prospect of auspicious outcomes from an online medium without any insignificant privacy threats[5]. Therefore, as the results, individuals are most prospective toward giving up the level of privacy for the sake of possible advantages linked to the online medium. In online medium settings, the consumer's dread in the form of losing control over their personal data is remunerated by the numerous paybacks likewise information, pleasure, and ease[13]. Accordingly, its predicted that:

H_2: Big data perspective of perceived benefits would have a lesser role in creating an information security concern.

Privacy threat belief suggests the likelihood of likely loss due to the release of private data[15]. Thus, it regarded as the price of solitude as demonstrating data is frequently deliberated insecure, so the cost and risks related to the online medium can vary from accidental middle parties getting users' private information to the hacking of personal account as per the information shared on online medium[5]. Numerous studies have inveterate the adverse outcome of apparent privacy threat on an individual's intention to unveil personal particulars over the internet transactions and actions[12].

H_3: Big data perspective of cloud storage would have a greater role in creating an information security concern.

Furby[8] defined perceived cultural values as imply a feeling of ownership and claim toward some behavioral practices. Over the medium of the internet, the researchers[18], Sharma and Crossler[6] identified possession entails as the sensitivity toward privilege, ownership, and attachment toward the information shared in the society over the online medium. When people trust that the information shared on the online medium is their own information and contains some degree of attachment using their individuality and solitude, it definitely affects their privacy risk belief[6]. Thus, its hypothesized that:

H_4: The cultural values would play a moderation role in improving companies' trust and engagement factor.

3 Data collection and sampling procedure

The data collection for this research study was based on taking opinions from the big data analysts working in Pakistani banking sector. The reason was as the present study aims to provide a framework for measuring whether the use of big data by the companies can create information security concern among the users and ultimately could contribute in dropping the companies' engagement and trust factors among their users. Hence, the big data analysts were the right person who can add the opinions for the proposed research framework for future investigations. The male and female big data analyst was separately contacted to add the separate gender perspective on current study proposed framework. Data were collected through adopted scales on each of the proposed variables from previous studies. The items on perceived benefits of big data were adopted by Kshetri[11], online behavior monitoring from Ion *et al*.[10], perceived cloud storage from Kshetri[11], information security concern from Bansal and Gefen,[7] and finally the cultural values were assessed through Hofstede cultural dimensions. All of the adopted scales were a test from the validation and reliability aspect [Tables 1 and 2].

3.1 Data analysis

The study used SPSS on the bases on Andrew's macro procedure that is a statistical application aiming to examine information collected as it is used in many studies that conduct quantitative study[4]. Moreover, data were first collected through Google Doc, and then transmuted into Excel to sort out information correctly. Subsequently, a codification was made to transmuted information from Excel into SPSS, for example, 1 for "Female," 2 for "Man" and for all dimensions utilized in our survey.

Regression analysis results for research question 1

Table 1. Reliability reporting

Proposed variables	Cronbach's Alpha	Number of items
Big data perspective of benefits	0.993	4
Online behavior monitoring	0.894	4
Online cloud security	0.789	3
Information security	0.819	4
Cultural values	0.892	4

Table 2. Validity

	BDPOBM	BDPPB	BDPCS	SC	CV
BDPOBM	1	0.279	0.010	0.123	0.20
BDPPB	0.320	1	0.187	0.002	0.430
BDPPB	0.001	0.034	1	0.021	0.065
SC	0.742	0.050	0.080	1	0.050
CV	0.062	0.003	0.004	0.007	1

Table 3. Male respondents based regression analysis based on correlation Andrew concept

Research question 1 proposed variables	BDPOBM	BDPPB	BDPCS
Dependent variable information security concern	−0.192 (0.765)	0.009 (0.067)	−0.103 (0.087)
−0.281 (0.083)*			
0.021 (0.082)			
−0.088 (0.077)			
R^2	0.077	0.077	0.100
Adjusted R^2	0.037	0.028	0.067
Standard Error	0.93985	0.87763	0.91980
F statistics	2.104	1.689	2.589*

*Significant at 0.05, **Significant at 0.01 (values in the parenthesis are representing beta values)

Table 4. Female respondents based regression analysis

Research question 1 proposed variables	BDPOBM	BDPPB	BDPCS
Dependent variable information security concern	0.489 (0.065)**	0.109 (0.067)	0.218 (0.087)*
−0.355 (0.083)**			
−0.018 (0.082)			
−0.188 (0.077)*			
R^2	0.76	0.278	0.786
Adjusted R^2	0.037	0.0280	0.037
Standard error	1.09398	1.07763	1.01980
F statistics	2.104	9.689**	1.589

*Significant at 0.05, **Significant at 0.01 (values in the parenthesis are representing beta values)

From the results of Table 3, the male data analyst in Pakistan depicts that the identified variables in the proposed research model as the big data perspective in creating information security concern do have some role. However, the lesser role as proposed in the literature would not found for Pakistani consumers where the privacy-related things are now becoming important due to the awareness and open talks in the online medium. The negative correlational value depicts that the use of big data as getting benefits, monitoring of online behavior and cloud storage usage can easily become game changer related to information security concern. The data analyst further comments that if Pakistani consumers receive any negative thoughts or opinions about the big data perspective than it can easily create an information security concern. The values of the Andrews test suggest for altering the current form of hypothesis statement. This means that in future researcher should propose an indirect relationship among the big data perspective and information security concern. On the other hand the factors highlighted by male data analyst for information security concerns are; perceived benefits and data monitoring.

The thoughts from the female data analyst Table 4 concluded that all of the proposed factors as big data perspective in creating information security concern definitely have a role in the creation of security concern. However, like male respondents, the results of the female respondents showed that all of the factors have an indirect affect in the creation of security concern which depends on the how female society members perceive about the big data usage across from the society members.

From the responses of male and female data analysts, it can be seen that the female comes out with high f-test value which states that the proposed model is more attractive for female as in Pakistan the female is more concerned toward protecting their online data to avoid any deeds that could affect their respect and image in the society. The male data analysts' thing that in Pakistani society if male consumers know that their companies are using big data that could be affect their privacy concern, the all of the proposed factors could be the cause of information security concern with total contribution value of 10% and big data perspective of online behavior monitoring and big data benefits could account for 7% in creating information security concern. On the other hand, the female data analyst vote big data benefits could be the second main trapper for females to create information security concern with the r square value of 27%. Similarly, the main factor in female opinion is cloud storage as they can directly hit the girls respect in society. Hence, it could count for 78% in creating an online security concern. The study also noticed that the male respondent's study results support for studying big data perspective on information security due to the value of adjusted r square that was not noticed from the female respondent's group. Hence, the proposed research question could be answer here with the approved research variables including online behavior monitoring, perceived data benefits and cloud storage are not only the drivers of big data but also the driver of online information security concern. Next, comes toward measuring the proposed research hypothesis. The result of the male sample group depicts

Table 5. Regression analysis results for research question2

Research question 1 proposed variables	BDPOBM	BDPPB	BDPCS
Moderator variable Information security concern	0.213 (0.109)*	281 (0.189)*	112 (0.137)
Moderation role of cultural values			
		−0.103 (0.135)	−0.199 (0.181)
R^2	0.76	0.278	0.786
Adjusted R^2	0.037	0.0280	0.037
Standard error	1.09398	1.07763	1.01980
F statistics	2.104	9.689**	1.589

*Significant at 0.05, **Significant at 0.01 (values in the parenthesis are representing beta values)

that online behavior monitoring can trigger the security concern as the result of companies big data usage. While on the other side of the female group sample results pointed out the online behavior monitoring and cloud storage can easily trigger security concern among the Pakistani female consumers. The regression analysis for this research question included both from the perspective of correlation and beta's values. The value row-wise representing beta's while the column-wise represents the correlational influence. In the male group the beta and correlational stands out more for the big data perspective of behavior monitoring. Hence, with a change in the online customers' perception about big data perspective of behavior monitoring the 76% change can be noticed regarding information security concern. The correlational value depicts that this change could be negative toward the organizations and it also found statistically significant; hence, H1 hypothesis should propose in a way that tests the indirect relationship between big data perspective and security concern. On the female side, the big data perspective of behavior monitoring plus Google+ services could be 7% and 8% change information security concern and the correlational value directed that this impact might be negative.

The role of cultural values from female sample group Table 5 was assessed because in Pakistani society the females are more restricted to bound with their cultural values. For example, the researcher being female bound to avoid online uploading of any personal data as compared to the male members in the society. Hence, on the personal experiences of the researcher, the group of female members was involved to study the moderation role of cultural values in information security concern. The findings confirm that the moderation could be possible in a future research study with related to benefits and cloud storage aspect of big data. The Pakistani cultural value related to risk avoidance can easily affect the usage of big data benefits and cloud storage due to the risk associated with these factors. Hence, the relationship with the organization can be drop from female side as the usage of big data technologies. It can be concluded that the researcher proposed hypothesis for answering research question three should be investigated in relation to studying the relationship between big data usage and information security concern in future research studies.

The present research study contributed to literature for bridging the gap of big data and information security concern. The purpose for this research study was to highlight the Pakistani big data using companies or who are searching for indulging the big data into their organization must be aware from the aspect of its linkage with the security concerns among the company service consumers.

4 Conclusion

The present research study provided the evidence that the big data usage is not only responsible for companies' efficiency and effectiveness but can also easily create information security issue that can directly drop the trustworthiness and relationship with its customers. The Pakistani big data analysts are actively supporting the proposed framework for bridging how big data related perspective can create information security concern, and the cultural values also need to be considered in the framework in the view of data analyst for understanding the risk and benefits associated with the usage of big data adoption in Pakistani companies. The researcher provides the future direction for the detailed study in Pakistan that will not only be a major contribution in literature of big data and information security but also present the real facts from the consumer perspective. This will be only possible if, this study targets the online consumers in understanding whether the current users are fully aware of the perspectives of big data in creating security concerns or not. Hence, the Pakistani companies can plan toward dealing with this issue which is still not being noticed.

4.1 Limitations

The present study is the part of researcher future research in progress on information security and

big data. Hence, the detail from the perspective of implementing this model into consumers for checking the actual implications of the proposed model will be reported after the project has been completed. The purpose of the research was to stands the pillars of theoretical framework from data analyst expertise that needs to further validated from cross-cultural studies.

Acknowledgement

This research is supported by Humanities and Social Science Talent Plan in Shaanxi through grant ER42015060002.

References

[1] Batool H, Rong DU, Ullah K. A model on information security through knowledge sharing attitude: Evidence from Pakistani banking industries' behaviors. Sci Int ;29:553.

[2] Batool H, Rong D. A general review on big data management challenges and themes. Int J Manag ;6:13-22.

[3] Tan W, Blake MB, Saleh I, Dustdar S. Social-network-sourced big data analytics. IEEE Internet Comput ;17:62-69.

[4] Boyd DM, Ellison NB. Social network sites: Definition, history, and scholarship. J Comput Mediat Commun ;13:210-230.

[5] Bulgurcu B. Understanding the Information Privacy-Related Perceptions and Behaviors of an Online Social Network User (Doctoral dissertation, University of British Columbia); 2012.

[6] Sharma S, Crossler RE. Disclosing too much? Situational factors affecting information disclosure in social commerce environment. Electron Commer Res Appl ;13:305-319.

[7] Bansal G, Gefen D. The impact of personal dispositions on information sensitivity, privacy concern and trust in disclosing health information online. Decis Support Syst ;49:138-150.

[8] Furby L. Possession in humans: An exploratory study of its meaning and motivation. Soc Behav Pers ;6:49-65.

[9] Feuchtl S, Kamleitner B. Mental Ownership as Important Imagery Content. ACR North American Advances Magazine; 2009.

[10] Ion I, Sachdeva N, Kumaraguru P, Čapkun S. Home is Safer than the Cloud!: Privacy Concerns for Consumer Cloud Storage. In: Proceedings of the Seventh Symposium on Usable Privacy and Security ACM; 2011. p. 13.

[11] Kshetri N. Big data's role in expanding access to financial services in China. Int J Inf Manag ;36:297-308.

[12] Li H, Sarathy R, Xu H. Understanding situational online information disclosure as a privacy calculus. J Comput Inf Syst 2010;51:62-71.

[13] Hogben G. Security Issues and Recommendations for Online Social Networks. Vol. 1. ENISA Position Paper; p. 1-36.

[14] Metzger MJ. Privacy, trust, and disclosure: Exploring barriers to electronic commerce. J Comput Mediat Commun 2004;9:942.

[15] Malhotra NK, Kim SS, Agarwal J. Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model. Inf Syst Res ;15:336-355.

[16] Nadkarni A, Hofmann SG. Why do people use facebook? Pers Individ Differ ;52:243-249.

[17] Ormond DK. The Impact of Affective Flow on Information Security Policy Compliance (Doctoral Dissertation, Mississippi State University); 2014.

[18] Richards NM, King JH. Three paradoxes of big data. Stanford Law Rev ;66:41.

[19] Raschke RL, Krishen AS, Kachroo P. Understanding the components of information privacy threats for location-based services. J Inf Syst ;28:227-242.

[20] Azam A. Model for individual information privacy disclosure in social commerce environment. Int J Bus Environ;7:302-326.