Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 100 (2016) 19 – 28

Procedia Computer Science

Conference on ENTERprise Information Systems / International Conference on Project MANagement / Conference on Health and Social Care Information Systems and Technologies, CENTERIS / ProjMAN / HCist 2016, October 5-7, 2016

Technological, organizational and environmental security and privacy issues of big data: A literature review

Khairulliza Ahmad Salleh[a,b]*, Lech Janczewski[a]

[a]*The University of Auckland, Business School, 12 Grafton Road, Auckland 1142, New Zealand*
[b]*Universiti Teknologi Mara Perak, Tapah Campus, Tapah 35400, Malaysia*

Abstract

This paper provides a literature review on security and privacy issues of big data. These issues are classified into three contexts; technological, organizational and environmental that is meant to facilitate future research. The main objectives of the review are to identify security and privacy issues of big data and to categorize the issues into a classification framework. The outcome of this review reveals that security and privacy issues of big data not only originate from technological deficiencies, but it may also be the outcome of organizational culture and environmental influences. At the end of review for each of the contexts, main issues were extracted and presented as potential factors that may affect organizational intention to adopt big data.

Peer-review under responsibility of the organizing committee of CENTERIS 2016

Keywords: big data; TOE framework; security and privacy; big data adoption

1. Introduction

Big data is a term that frequently appears in current business and academic discussions on recent technology trends. In publications, this term is rarely discussed without the inclusion of its unique characteristics; the 3Vs. The first 'V' refers to 'Volume', which describes large amount of data, the second 'V' is for 'Variety'- different types

* Corresponding author. Tel.: +64 2108270573
E-mail address: k.salleh@auckland.ac.nz

Peer-review under responsibility of the organizing committee of CENTERIS 2016
doi:10.1016/j.procs.2016.09.119

and sources of data and the final 'V' refers to 'Velocity' – the speed of data transfer and creation[1]. Else, other V's have also been described as forming the unique characteristics of big data, such as 'Value' and 'Veracity'[2]. These characteristics notably differentiate big data from the traditionally known methods used to capture, store, and analyze data. At present, big data is gaining greater attention due to increasing number of connected devices that generate very large amount of data. With proper use of big data technologies and applications, organizations will be able to exploit these data and transformed it into valuable information.

While the benefits of big data may reach diverse functions in organizations and individuals' life in digitized world, this extent of reach however, introduces far greater exposure to security and privacy risks. Although it is undeniable that big data sources may be utilized to derive better insights[3]; the underlying security and privacy concerns remain. These concerns may possibly be amplified by big data's volume, variety and veracity when deploying system infrastructure in supporting big data applications [4]. Organizations today are already confronted with overwhelming tasks of protecting their information assets, hence to some organizations; the idea of having big data applications deployed will invite further security issues and larger number of breaches. In fact, security and privacy issues have been cited in several big data survey done by technology providers and market research companies as one of the hindering factors in big data adoption[5 6 7].

Even though these issues have been reported multiple times as one of big data adoption's hindering factors, the specific security and privacy related issues that is of concern to organizations considering big data adoption are rarely discussed in publications. This study therefore intends to derive the possible security and privacy issues that may be influencing big data adoption by reviewing literatures in information systems domain. The following sections proceed as follows: section 2 briefly presents the motivation/objectives of the study and section 3 described the research methodology. The following sections present the findings of the literature review by classifying it into three contexts (section 4, 5, 6). The final sections draws a conclusion and provides future research direction.

2. Motivation, scope and objectives

Existing scholarly literature on big data were written from different perspectives to highlight the various applications of big data and its associated challenges in today's data driven era. Majority of literature on big data at present can be grouped into the following categories: big data overview, big data processing algorithm, big data applications, big data infrastructure and big data security, privacy and trust [8]. The largest number of publication can be found under the big data overview category, where scholars provide a general overview of big data, its challenges, the framework, techniques and technologies as well as other issues related to big data and its future direction in research. Examples of publications that fall under this category are those written by Chen and Zhang [9], an article that discusses on big data's impact on privacy, security and consumer welfare by Kshetri [10] and an article by Abbasi et.al.[11] that critically discuss the research agenda for big data research in information systems (IS).

This study chooses to add to the body of knowledge in the area of big data adoption/application and its associated security and privacy related concerns. While there are numerous publications that highlight the application of big data, the ones that specifically present the relation between security and privacy issues in big data adoption are still fragmented and scarce. Most discussion on security and privacy issues of big data exist as a sub-section in articles that surveyed big data challenges and opportunities in general.

Hence, this study aims to contribute to big data domain by conducting a literature review on big data's security and privacy related concerns and to present on how these security concerns may affect big data adoption by organizations. The main objectives of this study are: 1) To identify studies that discuss on security and privacy concerns of big data, and, 2) To categorize the security and privacy concerns/issues found in the articles into a classification framework (TOE – Technological, Organizational, Environmental).

3. Methodology

References to big data, analytics, big data technologies and certain combination of these terms can be found in most popular publication - in both online and physical form of publication. For the purpose of this study, the initial literature search were made on top IS academic journals. The IS journals selected were the eight leading journals under the Association for Information Systems (AIS) Senior Scholars' Basket of Journals; European Journals of