The Relationship between Online Data Collection and Consumer Autonomy

Osama Sohaib
School of Information, Systems and Modelling, University of Technology Sydney, Australia
Osama.Sohaib@uts.edu.au

Celina M. Olszak
Department of Business Informatics, University of Economics in Katowice, Poland
celina.olszak@ue.katowice.pl

Abstract

Recent advancements in the field of Big Data are facilitating various business intelligence activities for businesses. However, we contend that online data collection can generate tensions for consumers. The Big Data collection can compromise consumers' sense of autonomy, the lack of which can be harmful to consumer privacy, data security, data confidentiality, and data ownership. This study presents preliminary results on the relationship between online data collection and online consumer autonomy in Australia. This study identifies open research questions for future research.

1. Introduction

The advancement of web 2.0 technologies has intensely augmented the information flow, giving upsurge to a considerable volume of data called 'Big Data' [1]. The most common concern of the Big Data collection is disclosing consumers' private information to unrelated third parties or its irrelevant access [2]. The personal data leaks in two situations. Firstly, when the provider intentionally shares confidential information, which can be protected entirely [3]. Secondly, when the provider fails to safeguard the data. Privacy is complicated and contains various aspects. The damage caused by privacy invasions is immeasurable. One vital aspect of privacy is autonomy: a space in which one can make decisions on own.

A vast majority of online consumers' lack of awareness that their online web activities are being collected. For example, when doing online shopping or surfing the web, consumers' data are collected at different levels: web browsers, membership companies, social networks sites, ad-serving agencies, etc. Consumer analytics has been the interest of Big Data from a marketing perspective [4]. However, the impact of all these data and sources of information on consumers' autonomy, such as privacy, needs to be fully considered and assessed [5]. The consumer should retain control over their data. There is a need to build consumer trust and confidence in the online environment.

Current literature mainly focuses on the business perspectives of Big Data [4,6]. However, there is a need to expand our understanding of consumers' autonomy at the core of the Big Data collection. There are imminent opportunities regarding consumer behavior that the Big Data revolution will probably bring [7]. There is a need for research on consumers' perceptions of control and autonomy choice in the age of Big Data [8]. Big Data provides a huge benefit for businesses, such as consumer purchasing decisions and behavior. However, consumer autonomy is vital to revealing Big Data's full value potential.

Therefore, this project aims to discover consumer autonomy and control regarding the online data collection in Australia. According to the Consumer Policy Research Centre [5], there is a gap between consumer expectations and the Australian digital economy's current Big Data collection practice. Our research's rationale is two-fold: First, it is essential to understand the Australian consumer understands how their data is being used and shared. Secondly, there is a need to rectify the Australian consumers' lack of awareness about the data. The literature offers an inadequate understanding of how the consumer perceives Big Data.

This research will fill this gap by giving a comprehensive understanding of the relationship between online data collection and consumer autonomy and control. We will then develop managerial implications and areas for future research.

URI: https://hdl.handle.net/10125/70639
978-0-9981331-4-0

2. Big data overview

According to George [9], Big Data is "generated from an increasing plurality of sources, including Internet clicks, mobile transactions, user-generated content, and social media as well as purposefully generated content through sensor networks or business transactions such as sales queries and purchase transactions. Furthermore, genomics, health care, engineering, operations management, the industrial Internet, and finance all add to big data pervasiveness, cited in [1] are not appropriately catered. The ease and importance of personalization have genuine privacy concerns [10]. The increasing development in data collection, including online browsing and buyer attitude and the social network, the history of credit, and the importance of products play a huge role in the online market.

Technology has currently played a role in converting a usual consumer into someone who will continuously generate structured and unstructured type data and the traditional type data. The dimensions mentioned are essential in defining Big Data and are called three V's, i.e., volume, velocity, and variety [11,4]. Volume refers to the enormity feature, and velocity is the speed with which the Big Data is made even bigger via digital processing. Variety defines new formats and data diversity. Data may be in words, images, videos, and non-numeric forms and cannot be subjected to usual statistical analysis.

Many dangers characterize the commercial uses of Big Data. When a large amount of personal data is gathered, it increases the risk of the extent of losses linked to a security breach [12]. The literature related to Big Data consists of many things. Such as the revolutionary effects of Big Data generating new strategies for problem-solving [13], the impact of the digital technologies [14], the philosophy of Big Data [15], the challenges of Big Data [16], privacy concerns [17], and economic trade-offs [18]. The latest advancements in Big Data science have a positive effect on speeding up the growth in the development of new ideas, giving benefits for both the providers and the consumers [2]. However, one disadvantage is that it might affect the wellbeing of the consumer [2]. Data is a product available in every global economy [19], which persuades researchers to enhance the understanding of factual information generated. Recent advances emphasize on the business outlook of the Big Data. The focus should also be on consumer control and autonomy.

3. Research context

As stated by Akerkar [20] and discussed in detail by Sivarajah et al. [16], the Big Data main challenges are grouped into three core groups: data, process, and management challenges. However, the challenges are from the computing and technology perspective. "Big Data has the potential to further our understanding of each stage in the consumer decision-making process" [7]. Therefore, to capture the individual behavior in the Big Data life cycle so that valuable information can be derived, human analysis is also needed. Herein, we will investigate the new data culture of consumer control and autonomy that takes place in the Big Data collection life cycle.

Focusing on the Australian context, the authors formulate the following primary research question: *What is the association between online data collection and online consumer autonomy and control?*

This study considers three critical aspects of understanding consumer's autonomy to go through Big Data. (1) Big Data collection sources, (2) Big Data storage and access, (3) and Big Data management control services. Figure 1 shows the research study overview.

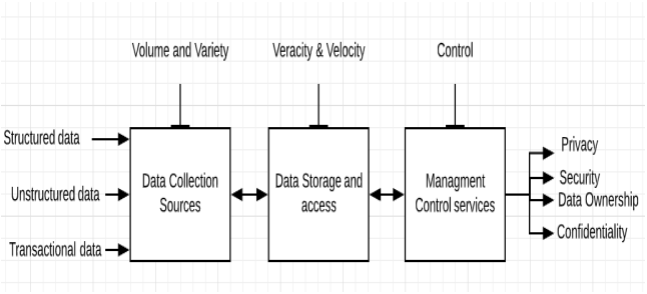

Figure 1. Big Data collection cycle

3.1 Big Data Collection Sources

The consumer uses a variety of data sources such as structured, unstructured, and transactional data. Retailers in the Big Data era now can collect store all of the consumers' record that was searched, purchased, clicked on, or abandoned, etc. from the search engines, social media, advertisement, and retailer websites, etc.

Consumers, which are the driving force for information sharing, are the enablers of Big Data and are respectively characterized by persona. The personas will describe the real people with backgrounds, goals, and values, giving a clear picture of the consumer's expectations and how they are likely to interact with the

online environment. In particular, we address the following research question: *Do different consumer personas have awareness and understanding of a different type of data collection?*

3.2 Big Data Storage and Access

Companies collect and store consumer data for analytics. Depending on the type of data, the data captured is analyzed and presented the results. Consumer awareness and control of the Big data collection and access are associated with consumer trust. For example, a company's privacy policy disclosure may increase consumers' trust in their data collection [22]. In particular, we address the following research question: *How acceptable do consumers find the companies use their data?*

3.3 Big Data Management Control Services

Management control services' concerns are related to managing and accessing Big Data. The central management control concerns are privacy, security, data governance, confidentiality, and data ownership [16,20]. In the Big Data realm, it is generally perceived that these challenges are vital from the provider's perspective. However, addressing the various control issues from the consumer's perspective is fundamental to current and future importance. For example, consumer control of personal data influences their privacy concerns [22, 23]. In particular, we address the following research question: How can consumers identify and trust the level of control over their data?

4. Research Method

The research uses a survey method to collect data from participants in Australia. According to Statista 2019 [21], "E-commerce revenue in Australia is expected to grow to 26.9 billion U.S. dollars in 2023". Sydney has the largest economy in Australia. Data were collected through a survey at the University of Technology Sydney. University students are the majority of online users, and Internet usage is comparatively higher than another age groups. Therefore, adopting students as the sample is considered more applicable to online consumers [24]. We collected data from both students and staff in the Faculty of Engineering and I.T. To ensure participant culture, it is vital to determine that each participant lived the most of their lives in the country and spoke the native language as their primary language [24].

Data collection lasted from March 2019 to May 2019. The survey questions consist of participants' profiles, knowledge, and awareness of Big Data collection in Australia, consumer attitude, acceptability, and consent to their data collection. We designed the survey instrument modified from the literature available such as [1, 5].

5. Results

This study shows the preliminary results of a data set consists of 110 responses. 55% of respondents were males, and 45% were females. The majority (45%) were 18-25 years old, followed by 26-34 years old (34%). Figure 2 shows that most female participants spend more than five hours per day on the Internet, while male participants spend 3-5 hours per day. Figure 3 shows participants' Internet usage.

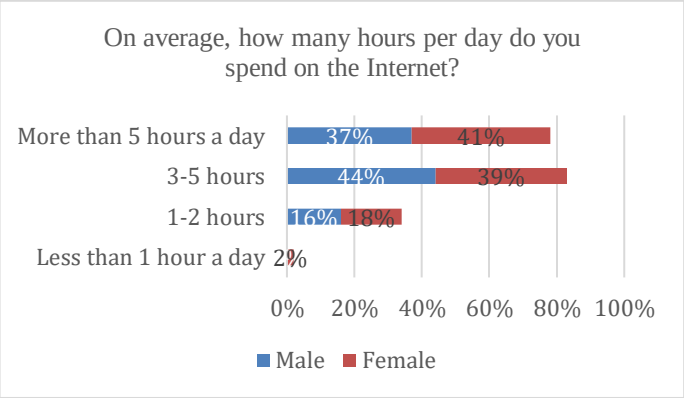

Figure 2. Internet usage

5.1 Big data collection awareness and understanding

Figure 3-4 shows that most participants are well aware of Big Data collection and the Australian data protection law. More than 80% of respondents are well aware of the volume of data being collected online.

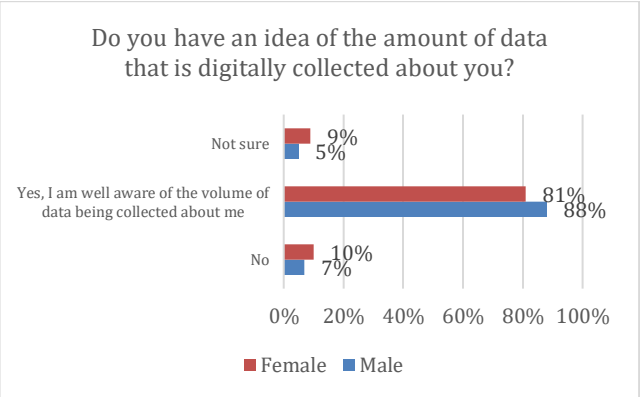

Figure 3. Consumer awareness of data

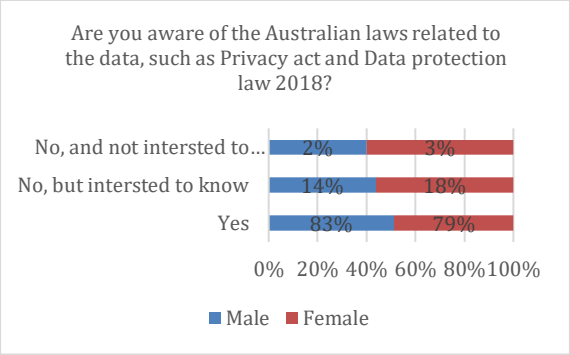

Figure 4. Awareness of Australian data protection laws

5.2 Consumer acceptability to data collection

Figure 5 shows consumers' acceptability of data collection. Table 1 also shows the results regarding the consumers (male and female) acceptability of their data collection. Overall, the results show that participants (both male and female) don't want their activities should be monitored, and information should not be shared without their knowledge.

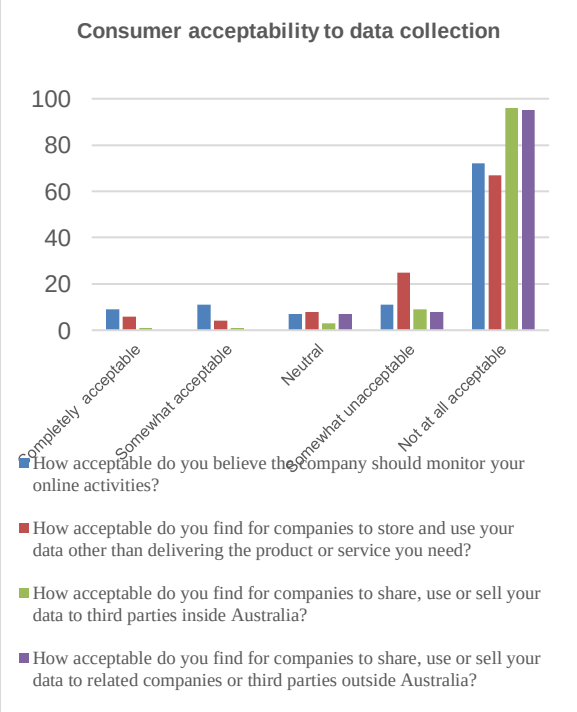

Figure 5. Consumer acceptability to data collection

5.3 Consumer attitude to data collection

Figure 6 shows consumers' attitudes towards their data collection. Table 2 shows the result regarding the consumers (male and female) attitude towards their data collection. Under the condition of this study, we observed a strong positive association between consumer attitude (both male and female) and online data collection. Overall, the majority of participants indicated that companies should give them a choice to control their data. Also, participants indicated that they trust the Australian law and local companies about their data protection.

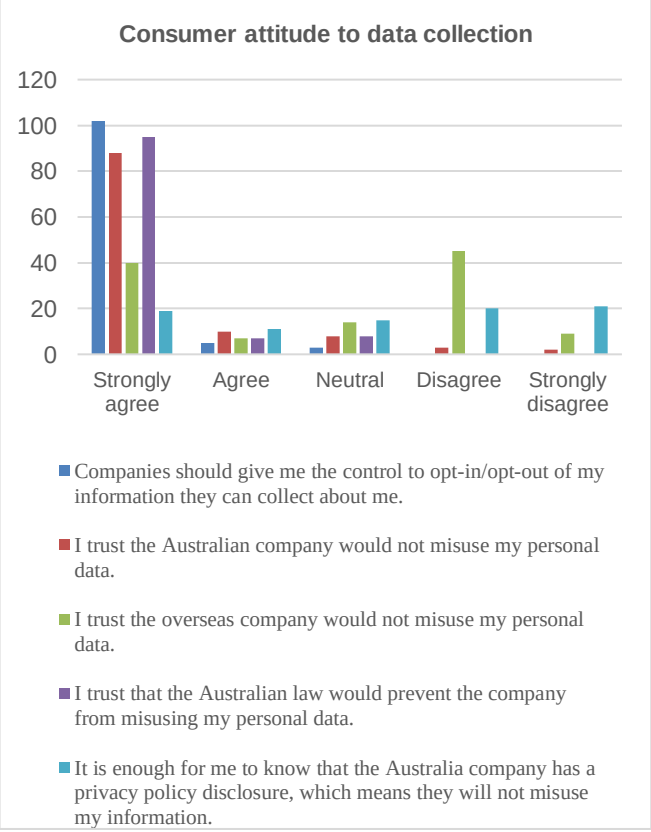

Figure 6. Consumer attitude to data collection

5.4 Consumer consent to data collection

Figure 7 shows the majority of participants consent to their data collection through online terms and conditions. We observed that participants (both male and female) will not usually consent when asked directly (such as via email or phone) from providers. Table 3 shows the results regarding the consumers (male and female) consent to their data collection.

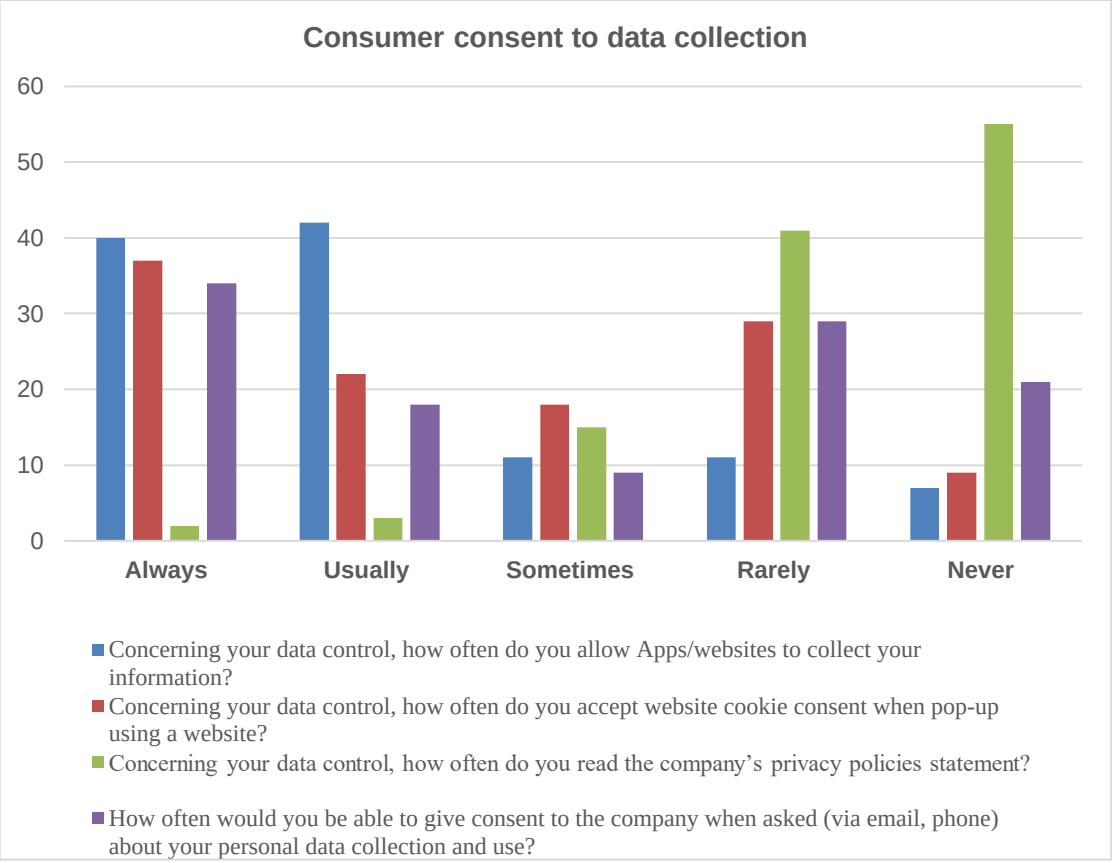

Figure 7. Consumer consent to data collection

Table 1. Consumer acceptability to data collection

F =Female M= Male	Completely acceptable		Somewhat acceptable		Neutral		Somewhat unacceptable		Not at all acceptable	
Questions	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**
How acceptable do you believe the company should monitor your online activities?	10%	12 %	15%	5%	7%	5%	7%	8%	60%	70%
How acceptable do you find for companies to store and use your data other than delivering the product or service you need?	5%	2%	5%	3%	8%	9%	27%	21%	55%	65%
How acceptable do you find for companies to share, use, or sell your data to third parties within Australia?	1%	0%	2%	0%	5%	0%	15%	7%	82%	93%
How acceptable do you find for companies to share, use, or sell your data to related companies or third parties outside Australia?	0%	0%	0%	0%	8%	5%	6%	9%	86%	96%

Table 2: Consumer attitude to data collection

F =Female M= Male	Strongly agree		Agree		Neutral		Disagree		Strongly disagree	
Questions	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**
Companies should give me the control to opt-in/opt-out of my information they can collect about me.	90%	95%	5%	3%	2%	2%	0%	0%	0%	0%
I trust the Australian company would not misuse my personal data.	85%	76%	10 %	8%	4%	9%	1%	4%	0%	2%
I trust the overseas company would not misuse my personal data.	35%	38%	7%	5%	9%	15%	45%	41 %	5%	11%
I trust that the Australian law would prevent the company from misusing my personal data.	89%	85%	7%	6%	4%	9%	0%	0%	0%	0%
It is enough for me to know that the Australia company has a privacy policy disclosure, which means they will not misuse my information.	46%	34%	14 %	7%	11 %	15%	16%	19 %	13 %	24%

Table 3: Consumer consent to data collection

F =Female M= Male	Always		Usually		Sometimes		Rarely		Never	
Questions	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**	**F**	**M**
Concerning your data control, how often do you allow Apps/websites to collect your information?	49%	25%	25%	49%	09%	11%	12%	8%	5%	7%
Concerning your data control, how often do you accept website cookie consent when pop-up using a website?	35%	32%	23%	17%	11%	21%	25%	27%	6%	9%
Concerning your data control, how often do you read the company's privacy policies statement?	2%	1%	3%	4%	15%	13%	34%	29%	46%	53 %
How often would you be able to give consent to the company when asked (via email, phone) about your personal data collection and use?	16%	22%	19%	14%	11%	5%	20%	31%	34%	28 %

6. Discussions and Implications

Nowadays, Big data is in every business. Big data is a core function of online companies and considered a vital factor for competitive advantages. As presented in Section 3, businesses use various data sources such as structured, unstructured, and transactional data and also store and control it. Companies need to access consumers' data continuously using multiple information resources to gain business value and satisfy customers [26]. One of the dynamic capabilities for value creation of any organization is information resources, which is intended at attaining, analyzing, and using Big data [27]. Big data resource exploration may be aimed at redesigning the business network, which involves new members in the business functions [27]. The study's findings show that businesses can also leverage consumers' data effectively to gain business value. The companies can derive value by creating transparency about consumer data [22].

According to Olszak and Zurada [27], businesses are well aware of the potential "value," resulting from Big data exploration and adoption. However, they are not continually to quantify these values completely [27]. One way to quantify the business value might be in giving consumers greater control over their data.

The findings show that most consumers find it unacceptable that businesses should store and use their data other than delivering products or services. Companies should give consumers greater control of data collection practices and reduce privacy concerns [22]. By doing so will also help to build consumer trust and confidence in the online environment.

Furthermore, according to Ghoshal et al. [29], businesses also perceive 'value' as social and economic values for adopting and using Big data. Social value suggests improving social wellbeing, including people's safety and security [30, 31]. This study's findings show that most participants (34%) are willing to give consent if asked about their personal data collection. This means businesses and Governments can use the data collection to enhance transparency. According to Kim et al. [32], the Government uses big data to improve transparency and support wellbeing and increase citizen engagement in public affairs. Therefore, the social value includes benefits for individual users, together with society benefits [33].

Economic values refer to monetary benefits that Big data help businesses to perform better financially [34]. The economic benefits of Big Data are well known.

Businesses use big data for competitive advantages using increased insights of their consumers. However, this study's findings also created questions related to an individual's autonomy and control, including the right to privacy regarding their data collection and use. Such as,

- How can consumers identify and trust the level of control over their data?
- What is the impact of consumer willingness to share personal data?
- How can companies ensure consumer privacy and security concerns?
- How can consumers get the power of data ownership?
- Can businesses use consumer data without consent?

In conclusion, this study increases the understanding of Big Data from a consumer perspective. The study identifies a gap in research of Big data collection and consumer autonomy and control. Online managers need to understand Big Data's full value potential to gain consumer trust. The findings also show how online businesses should accommodate a different kind of persona when collecting consumer data. Also, the results will help online businesses in the development of better-targeted gender-dependent and different age groups engagement strategies to increase their participation level in online business.

By discovering how consumers in Australia perceive Big Data, this study produces several opportunities for managerial implications and awareness of consumer autonomy and choice behavior. The impact this research in practice includes recommendations for policy development of Australian consumer law. Practical implications will be drawn regarding strategies and policy for the Australian online business industry and to the Australian Government of data availability and use.

6.1 Limitations and Future work

Further research is required to provide awareness, understanding, and choice to consumers, to build greater control over their collected data. This study presented the initial descriptive findings. Given the limited or lack of academic research exploring consumer awareness of Big Data, in future work, a multi-method approach will be used that includes both qualitative and quantitative methods, including consumer focus groups, in-depth interviews, and surveys.

In the first stage, a focus group with online consumers will determine the interpretation of Big Data from individual consumers' perspective over how their data is used and shared. The first stage will also enable critical themes to arise, which will then be considered to develop a conceptual model in the second stage. In the second stage, an online survey will collect a more significant data set from the Australian consumers' online community. Data will be analyzed, and statistical

tests such as structural equation modeling will be conducted to test the hypotheses. Finally, recommendations will be made.

7. References

[1] Lichy, J., Kachour, M., and Khvatova, T.: 'Big Data is watching YOU: opportunities and challenges from the perspective of young adult consumers in Russia', Journal of Marketing Management, 2017, 33, (9-10), pp. 719-741

[2] Helveston, M.N.: 'Consumer Protection in the Age of Big Data', Law Review, 2016, 93, (4), pp. 859 - 917

[3] Crawford, K.a.S., Jason: 'Big Data and Due Process: Toward a Framework to Redress Predictive Privacy Harms', Boston College Law Review, , 2014, 55, (93), pp. 13-64

[4] Erevelles, S., Fukawa, N., and Swayne, L.: 'Big Data consumer analytics and the transformation of marketing', Journal of Business Research, 2016, 69, (2), pp. 897-904

[5] Nguyen, P., and Solomon, L. 2019. "Consumer Data and the Digital Economy - Emerging Issues in Data Collection, Use and Sharing." Retrieved 15 Feb, 2019, from http://cprc.org.au/wp-content/uploads/Full_Data_Report_A4_FIN.pdf

[6] Berntzen, L., and Krumova, M.: 'Big Data from a Business Perspective', in Editor (Ed.)^(Eds.): 'Book Big Data from a Business Perspective' (Springer International Publishing, 2017, edn.), pp. 119-127

[7] Hofacker, C.F., Malthouse, E.C., and Sultan, F.: 'Big Data and consumer behavior: imminent opportunities', 2016, 33, (2), pp. 89-97

[8] André, Q., Carmon, Z., Wertenbroch, K., Crum, A., Frank, D., Goldstein, W., Huber, J., van Boven, L., Weber, B., Yang, H.J.C.N., and Solutions: 'Consumer Choice and Autonomy in the Age of Artificial Intelligence and Big Data', 2018, 5, (1), pp. 28-37

[9] George, G., Haas, M.R., and Pentland, A.: 'Big Data and Management', 2014, 57, (2), pp. 321-326

[10] Aguirre, E., Mahr, D., Grewal, D., de Ruyter, K., and Wetzels, M.: 'Unraveling the Personalization Paradox: The Effect of Information Collection and Trust-Building Strategies on Online Advertisement Effectiveness', Journal of Retailing, 2015, 91, (1), pp. 34-49

[11] Lycett, M.: "Datafication': making sense of (big) data in a complex world', European Journal of Information Systems, 2013, 22, (4), pp. 381-386

[12] Romanosky, S., Hoffman, D., and Acquisti, A.: 'Empirical Analysis of Data Breach Litigation', 2014, 11, (1), pp. 74-104

[13] Armstrong, K.: 'Big data: a revolution that will transform how we live, work, and think', Information, Communication & Society, 2014, 17, (10), pp. 1300-1302

[14] Mantelero, A.: 'Social control, transparency, and participation in the big data world', Journal of Internet Law, 2014, 17, (10), pp. 23-29

[15] Swan, M.: 'Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services', in Editor (Ed.)^(Eds.): 'Book Philosophy of Big Data: Expanding the Human-Data Relation with Big Data Science Services' (2015, edn.), pp. 468-477

[16] Sivarajah, U., Kamal, M.M., Irani, Z., and Weerakkody, V.: 'Critical analysis of Big Data challenges and analytical methods', Journal of Business Research, 2017, 70, pp. 263-286

[17] Boyd, d., and Crawford, K.: 'Critical questions for big data', Information, Communication & Society, 2012, 15, (5), pp. 662-679

[18] Reed, D.A., and Dongarra, J.: 'Exascale computing and big data %J Commun. ACM', 2015, 58, (7), pp. 56-68

[19] Zhang, J., Yang, X., and Appelbaum, D.: 'Toward Effective Big Data Analysis in Continuous Auditing', 2015, 29, (2), pp. 469-476, Available Online: http://raw.rutgers.edu/Accounting_articles/Toward%20Effective%20Big%20Data%20Analysis%20in%20Continuous%20Auditing.pdf

[20] Akerkar, R.: 'Big data computing' (C.R.C. Press, Taylor & Francis Group, 2014. 2014)

[21] Statista: 'Total retail e-commerce revenue in Australia 2017-2023', in Editor (Ed.)^(Eds.): 'Book Total retail e-commerce revenue in Australia 2017-2023' (2019, edn.), pp.

[22] Perreault, L. 2015. "Big Data and Privacy: Control and Awareness Aspects," International Conference on Information Resources Management (CONF-IRM): A.I.S., p. 15.

[23] Prince, C. 2018. "Do Consumers Want to Control Their Personal Data? Empirical Evidence," International Journal of Human-Computer Studies (110), pp. 21-32.

[24] Chen, Y.-H., and Barnes, S. 2007. "Initial Trust and Online Buyer Behaviour," Industrial Management & Data Systems (107:1), pp. 21-36.

[25] Cyr, D. 2013. "Website Design, Trust and Culture: An Eight Country Investigation," Electronic Commerce Research and Applications:0), pp. 1-12.

[26] Olszak, C.M., Bartuś, T., & Lorek, P. 2018. A Comprehensive Framework of Information System Design to Provide Organizational Creativity Support. Information & Management, 55, pp. 94-108.

[27] Wang, Y., Kung, L., Wang, W.Y.C., & Cegielski, C.C. 2018. An integrated big data analytics-enabled transformation model: application to health care. Information & Management, 55, 64-79.

[28] Celina M. Olszak & Jozef Zurada. 2020. Big Data in Capturing Business Value, Information Systems Management, 37:3, 240-254, DOI: 10.1080/10580530.2020.1696551

[29] Ghoshal, A., Larson, E.C., Subramanyam, R., & Shaw, M.J. 2014. The impact of business analytics strategy on social, mobile, and cloud computing adoption. Proceedings of the Thirty Fifth International International Conference on Information Systems, Auckland, New Zealand, December 14–17.

[30] Newell, S., & Marabelli, M. 2015. Strategic opportunities (and challenges) of algorithmic decision-making: a call for action on the long-term societal effects of' datafication. Journal of Strategic Information Systems, 24 (1), 3–14. Retrieved from http://dx.doi.org/10.1016/j.jsis.2015.02.001.

[31] Cech, T.G., Spaulding, T.K., & Cazier, J.A. 2015. Applying business analytic methods to improve organizational performance in the public school system.

Proceedings of the Twenty-First Americas Conference on Information Systems, Puerto Rico, August 13–15.
[32] Kim, G., Trimi, S., & Chung, J. 2014. Big-data applications in the government sector. Communications of A.C.M., 57 (3), 78–85. doi: http://dx.doi.org/10.1145/2500873.
[33] Loebbecke, C., & Picot, A. 2015. Reflections on societal and business model transformation and risking from digitalization and big data analytics: a research agenda. Strategic Information Systems, 24 (3), 149–157. Retrieved from http://dx.doi.org/10.1016/j.jsis.2015.08.002.
[34] McAfee, A., & Brynjolfsson, E. 2012. Big data: The management revolution. Harvard Business Review, October 2012, 59-69.