AN INVESTIGATION OF BIG DATA IN HEALTHCARE AMONG HEALTH CONSUMER

SOH RAY QIU

The thesis is submitted in partial fulfilment of the requirements for the award of
Bachelor of Technopreneurship
With Honours

Faculty of Technology Management and Technopreneurship
Universiti Teknikal Malaysia Melaka

JUNE 2019

APPROVAL

I/We, hereby declared that I/We had read through this thesis and in my/our opinion that
this thesis is adequate in terms of scope and quality which fulfil the requirements for the
award of Bachelor Degree of Technology Management
(Technopreneurship)

Signature : ……

Name of Supervisor : DR. HAZMILAH BINTI HASAN

Date : ……

Signature : ……

Name of Panel : MISS ATIKAH SAADAH BINTI SELAMAT

Date : ……

DECLARATION

"I admit that this report is the result of my own, except certain explanations and passages where every part of it is cited with sources clearly stated in References."

Signature : ……

Name : SOH RAY QIU

Date : .....

DEDICATION

This humble work is especially dedicated to:
Dr. Hazmilah Hasan, my project supervisor,
All the respondents,
And
To my family and all my loved ones,
Thank you for being my guidance and support.

ACKNOWLEDGEMENT

Taking this opportunity, I am appreciating sincere from the heart to all the parties who have facilitated me in this research project throughout the duration of the time until the completion of the research.

Firstly, I would like to express our gratitude to my supervisor, Dr Hazmilah Hasan, guided me to complete this research project. She had spent her precious time to help and guide me when I was in doubt through the development of the research. Her valuable guidance, support and suggestion have helped me a lot when I am facing difficulties in my research project. Moreover, I am truly thankful to my panel, Miss Atikah Saadah Binti Selamat which helping throughout the whole research project.

Besides that, I would also like to show appreciation towards the respondents for their time and effort in completing the interviews. Without their contribution from their responses, this research project would not have proceeded. Next, special appreciations to my dearest family and friends who given me their moral support in completing this research.

Last but not least, I would like to express my appreciaton to the Faculty of Technology Management and Technopreneurship for giving me the opportunity to work on this research. Through this research project, I strongly believe that all the precious experiences that gained would make me become a more competent person in the near future.

ABSTRACT

Malaysia is one of the few countries with a structured Big Data Analytics (BDA) roadmap to unleash the value of big data. At the turning point of digital revolution, the powers of big data can be used to describe a problem, assess a situation, forecast results, and prepare solutions in all sector. Therefore, big data in healthcare is important as it can be used in the prediction of outcome of diseases prevention of co-morbidities, mortality and saving the cost of medical treatment. In Malaysia, the focus on big data has started and some initiatives have been put in place to share information patient's medical records and knowledge among general public, private hospitals and clinics. However, there are many challenges in implementing big data in healthcare especially in relation to fraud activity (i.e. scam), privacy, security, and etc. Based on the challenge that had been mention, this research focused on the level of trust health consumers which later leads to intention of sharing their data to healthcare providers. TRA model was used in this research to study consumer's experiences, belief, and intention on sharing data. Qualitative research method with the approach of interview was used in this study. The research will be conduct in Melaka. Moreover, this research is used non-probability sampling techniques such as purposive sampling and snowball sampling. Sampling size for this research is 6. As a conclusion, the result finding achieved and supported by literature review on big data in healthcare and significance of study is benefit and help future researchers as their guidance and reference as well as open up more forum in this topic and broaden their views on this issues.

Keywords: *Big Data Analytics (BDA), Healthcare, Fraud Activities, Trust, Privacy, Interview, Health Consumers.*

ABSTRAK

Malaysia adalah salah satu daripada beberapa negara yang mempunyai pelan tindakan Big Data Analytics (BDA) yang berstruktur untuk melepaskan nilai data besar. Pada titik perubahan revolusi digital, kuasa data besar boleh digunakan untuk menggambarkan masalah, menilai keadaan, keputusan ramalan, dan menyediakan penyelesaian di semua sektor. Oleh itu, data besar dalam penjagaan kesihatan adalah penting kerana ia boleh digunakan dalam ramalan hasil pencegahan penyakit morbiditi, kematian dan penjimatan kos rawatan perubatan. Di Malaysia, tumpuan terhadap data besar telah bermula adalah untuk berkongsi rekod perubatan maklumat dan pengetahuan pesakit di kalangan orang awam, hospital dan klinik swasta. Walau bagaimanapun, terdapat banyak cabaran dalam melaksanakan data besar dalam penjagaan kesihatan terutamanya berkaitan dengan aktiviti penipuan (iaitu penipuan), privasi, keselamatan, dan sebagainya. Berdasarkan cabaran yang telah disebutkan, kajian ini memberi tumpuan kepada tahap pengguna kesihatan amanah yang kemudiannya membawa kepada niat untuk berkongsi data mereka kepada penyedia penjagaan kesihatan. Model TRA digunakan dalam kajian ini untuk mengkaji pengalaman, kepercayaan, dan niat pengguna untuk berkongsi data. Kaedah penyelidikan kualitatif dengan pendekatan wawancara digunakan dalam kajian ini. Penyelidikan akan dijalankan di Melaka dan kajian ini menggunakan teknik pensampelan purposive dan pensampelan bola salji. Saiz sampling untuk penyelidikan ini adalah 6. Sebagai kesimpulan, pencapaian hasil yang dicapai dan disokong oleh tinjauan literatur mengenai data besar dalam penjagaan kesihatan dan kepentingan kajian adalah manfaat dan membantu para penyelidik masa depan sebagai panduan dan rujukan mereka serta membuka lebih banyak forum dalam topik dan meluaskan pandangan mereka mengenai isu ini.

Kata kunci: *Big Data Analytics (BDA), Penjagaan Kesihatan, Kegiatan Penipuan, Amanah, Privasi, Wawancara, Pengguna Kesihatan.*

TABLE OF CONTENT

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

NO	ABBREVIATION	DESCRIPTION
1	BDA	Big Data Analytics
2	MCMC	Malaysian Communication and Multimedia Commission
3	HC	Health Consumers
4	MyHDW	Malaysian Heath Data Warehouse
5	MOH	Ministry of Health Malaysia
6	IPTK	Institute for Health Behavioral Research
7	TRA	Theory of Reasoned Action Model
8	DIKW	Data–Information–Knowledge–Wisdom Hierarchy

LIST OF APPENDIXES

CHAPTER I

INTRODUCTION

1.0 INTRODUCTION

In this chapter, background study will be discussed initially and then followed by the problem statement. Apart from that, the research objectives and research questions also will be developed in this chapter. Besides, this chapter also will cover the scope and limitations and significance of the study.

1.1 BACKGROUND OF STUDY

Data can create value for businesses, as insight gained from its analysis can inform decisions. Traditionally, companies have used business intelligence tools for data analysis. These are applied to structured data, which reside in fixed fields, as in a spreadsheet. By comparison, 'big data' analytics typically involve data from a wider variety of sources, which may be rapidly analyzed as they are collected (in 'real-time'). These data are often unstructured, such as images or videos. There is no single official definition of big data, but big data specialists often refer to its high volume, velocity and variety. These features make it difficult to readily manage and analyze using desktop computers and traditional data management tools (McKinsey Global Institute, 2016).

Big data is created every day by the interactions of billions of people using computers, GPS devices, cell phones, censors and medical devices, data-intensive areas such as atmospheric science, genome research and astronomical studies. Today big data opens huge opportunities to those who can use it effectively. Now realizing the great importance of big data, many analytical companies are engaged in finding hidden information in big data. According to internet experts the present technological advances to collect and analyze massive sets of data is likely to lead to revolutionary changes in business, and society. Many organizations are giving attention in big data analytics for development, education, disaster management, health care, and natural resource management for benefit of society (Ajit Kumar R, 2016).

Big data reveals trends and provides predictive analytics for retail, manufacturing, and has proven especially valuable for healthcare. According to Center for Disease Control is using big data to predict the latest flu epidemic or a hospital is trying to reduce post-operative infection, big data analytics improve patient care and practice operations. Pharmaceutical companies have been aggregating and organizing data into medical databases for decades. Payers and healthcare providers have adopted digital records; in fact HIPAA regulations require digitization. Government agencies have

been storing vast amounts of healthcare statistical data. Assimilating and analyzing that healthcare information is how big data isolates problems; by aggregating information and looking for patterns.

Big data in health informatics can be used to predict outcome of diseases and epidemics, improve treatment and quality of life, and prevent premature deaths and disease development. Big data also provide information about diseases and warning signs for treatment to be administered. This will help not only to prevent co-morbidities and mortality but also assists government to save the cost of medical treatment. It is very useful not only in clinical medicine for diagnosis/detection but also in epidemiological research as the big data will provide huge amount of data. The government, non-governmental organization and/or pharmaceutical companies can use the data to formulate policies, strategies, intervention or medical treatment such as drugs development (Fatt, 2018).

Big data has implications on healthcare on patients, providers, researchers, health professional. It's important to remember that the primary value from big data comes not from the data in its raw form, but from the processing and analysis of it and the insights, products, and services that emerge from analysis. The sweeping changes in big data technologies and management approaches need to be accompanied by similarly dramatic shifts in how data supports decisions and product/service innovation. With this much raw data at their fingertips, healthcare organizations can see how big data isolates problems that affect both patients and operations (Pires, 2014).

1.2 PROBLEM STATEMENT

Nowadays, there is an increasing demand for more information by the patients about their healthcare options or choices, and want participation in their health decision-making. The big data will help to provide patients with up-to-date information to assist them to make the best decision and to comply with the medical treatment. The Malaysia National Health and Morbidity Survey in 2015 has revealed that the number of obese Malaysians have risen to 17.7% compared to 4.4 % in 1996 and 17.5% of those aged 18 and above have diabetes compared to 11.6% in 2006. There is a need to capture and analyze this raw data information to provide better healthcare, accessibility, affordability and quality of healthcare from diagnosis, treatment and follow-up. In 2017, the Ministry of Health Malaysia (MoH) has launched the Malaysian Health Data Warehouse (MyHDW) to share information patient's medical records and knowledge among public, private hospitals and clinics.

MyHDW aims to synchronise patients' data from public hospitals (including university hospitals, armed forces hospitals), private hospital and clinics along with National Registration Department (NRD), National Department of Statistics, and other health related agencies where it will serve as one stop centre to provide healthcare providers to make decisive decision on treatments. As most of the time, medical data are collected in silos in their respective healthcare centres and is governed and controlled by hospitals or clinics administrative departments.

Consequently, if big data is successful implemented in Malaysia, it can reduce wasteful overheads and effective managed in healthcare sector. There are lots of benefits of big data bring toward all sector one of good example is e-commerce business which shown successfulness of application of big data. Therefore, big data should be implement on medical field or healthcare sector in Malaysia.

However, some key issues can become a threat of the usage of big data technologies such as fraud activity (i.e. scam), privacy issue, data security and lack of IT talents. One of biggest issue is fraud activity in big data whereby third party hacked

into system to take user's personal data. They can use these data to cheat people (i.e. scam) or just sell of the data that they have as some of these data are very valuable for example information in bank and military. For example, in October 2017, a massive data leakage was published for sale on a well-known Malaysian online forum, Lowyat.Net. Because the forum posted sample data leaks, it triggered thousands of users' anger as many were able to find their details leaked. The leakage dates back to 2012 to 2015 (Alta & Rahman & Mutalib, 2017). The data is a large list from Malaysian telcos, including Celcom, Maxis, Digi, Altel, Enabling Asia, FriendiMobile, MerchantTradeAsia, PLDT, RedTone, TuneTalk, UMobile and XOX. This could be the biggest data breach ever in Malaysian history. If this data falls into the wrong hands, it might enable other criminal activities, such as scam, phishing and targeted attacks to expand their targets. An estimated 50 million records were leaked and put up for sale.

Thus, due to fraud activity in data causes the awareness of people to be more concern on the data that they are sharing. Hence, when Malaysian Government launches MyHDW, one of the potential issues with MyHDW need to be ironed out now, and there is a need to take into account public trust, patient privacy issues and confidential information, as this is about handling the most sensitive of data — a person's health (Fatt, 2018). Nonetheless, the most frequent barrier to big data analytics is privacy concerns as consumer may afraid and feel fear to share their personal data. Thus, consumers' feelings for sharing their personal data are still an issue on today's society (Pires, 2014). To researcher knowledge, there are currently no findings on the trust on privacy in the application of big data by using Theory of Reasoned Action (TRA) model.

Hence, it needs for the researcher to further investigate or study on this topic. The researcher has prepared the research questions regarding with the challenge as mention above in order to proceed the study.

1.3 RESEARCH QUESTIONS

The following research questions had been determined to achieve the research objectives and to aims to study the level of trust Malaysian health consumers' which later leads to intention of sharing their data with healthcare providers.

RQ 1: What are fraud action or activity in medical health data among big data owner?

RQ 2: What are the level of trust among health consumers towards healthcare provider (i.e. Malaysian Health Data Warehouse (MyHDW))?

RQ 3: What are the appropriate solution to increase the level of trust among health consumers on big data analytics in healthcare?

1.4 RESEARCH OBJECTIVE

These are the research objectives that attempted to be achieved in the end of the study:

RO 1: To investigate fraud action or activity in medical health data among big data owner.

RO 2: To identify the level of trust among health consumers towards healthcare providers (i.e. Malaysian Health Data Warehouse (MyHDW)).

RO 3: To suggest appropriate solution to increase the level of trust among health consumers on big data analytics in healthcare.

1.5 SCOPE OF RESEARCH

This research focuses on study the intention of health consumers' about big data applications on healthcare. Moreover, this study aims to understand if health consumers' perception of value driven from big data applications on healthcare surpasses their miss feelings to share their personal data. Besides that, this research carried out with a sampling opinion representing people in Melaka Town Area (Melaka Tengah). The suggestions for solving the issue of privacy and security are also included in this research.

1.6 LIMITATION

There are several limitations the researcher have to face in order to complete this study. The following are the limitations through by the researcher:

- **LACK OF RESOURCES**

  As many journal, articles and published book is covered on the techniques or tools of big data analytics in e-commerce, supply chain management and healthcare. There are only few studies on the scope of awareness of big data analytics and their usage in healthcare. Hence, it could be more difficult for the researcher to gain suitable information and details that regarding with this topic.

- **BUDGET CONSTRAINT**

  Limited budget also acts as a limitation for this research. As it is difficult for the researcher because all the costs that involved in completing this study is by using personal savings. Thus, this research only conducted in Melaka as the budget of researcher is limited.

- **TIME CONSTRAINT**

  The duration of this research is limited as the research has to be complete in two semester which less than 6 months. Therefore, sample size of interview is small for this study as interview is very time consuming. Time management on the interview session is important to complete this study.

1.7 SIGNIFICANCE OF STUDY

As nowadays big data opens huge opportunities to those who can use it effectively. Now realizing the great importance of big data, many analytical companies are engaged in finding hidden information in big data. According to internet experts the present technological advances to collect and analyze massive sets of data is likely to lead to revolutionary changes in business, and society. Many organizations are giving attention in big data analytics for development, education, disaster management, healthcare, and natural resource management for benefit of society. However, there are also challenges on implementation of big data such as people are now more concern on privacy which involve sharing of personal data and data security. Thus, the research focused on the level of trust health consumers towards MyHDW. This research also covered the intention of health consumers for sharing their data to healthcare providers.

This research will benefit and help future researchers as their guidance and reference as well as open up more forum in this topic and broaden their views on this issues. As this research can be references for future researcher who aim to have future study on the implementing of big data analytics in medical field. This study is used TRA model as theoretical framework to study the behavior of health consumers on data sharing.

Moreover, this research can be benefit to policymakers (regulators) and government in term of strong the rules and regulation of personal acts (PDPA) whereby people are really concern on privacy and confidential information that share with third party. Besides, the results of research can be important to organization especially in healthcare sector as the finding may help organization to understand about the concern of consumers so that organization will not carry out any fraud activity.

1.8 SUMMARY

Basically, most of the study in this chapter explains on the basic of introduction of the research, element in this chapter consists of the research background, problem statements, research methodology, research question, research objective, scope of study, limitation of study and also significant of the study. Big data analytics is one of the important tool in health care sector in Malaysia as it provide real in time analysis on the diseases and ensure the health information in internet are reliable. Therefore, the next chapter in the literature review will be exposed in detail about definition and concept of big data and scenario of big data in Malaysia. Theoretical framework will be used to investigate the level of trust health consumers towards healthcare provider also covered in this study.

CHAPTER II

LITERATURE REVIEW

2.0 INTRODCUTION

This chapter will discuss in depth the literature review of Big Data which providing the reader with the academic status quo about big data's topic. Besides, this chapter also will discuss the scenario of Malaysia on the adoption of big data analytics in healthcare sector. Moreover, benefits and challenges on the implementation of big data in Malaysia will be explain. Other than, this chapter also include the model of theoretical framework based on the objectives that want to achieve. To conclude, this chapter presents current research about big data's field of study and the motivations for this dissertation to study big data's applicabilities on healthcare.