A Systematic Literature Review of Data Privacy Issues in Cloud BI

Jiyuan Zhang

A dissertation submitted to Auckland University of Technology in partial fulfilment of the requirements for the degree of Master of Business

Abstract

Cloud Business Intelligence (Cloud BI) can provide a competitive advantage to businesses, but it also has implications for data security and privacy. Leakage of sensitive data can cause serious impacts on the credibility of a company and result in financial as well as trust losses. Therefore, it is especially important to discuss the data security and privacy issues of Cloud BI. This study uses a systematic literature review to understand Cloud BI security. The results show that internal and external organisational factors have an impact on Cloud BI privacy. This paper also explored the security issues in cloud computing and found similarities and differences between cloud computing and Cloud BI. The findings of the research have applied and theoretical significance and will help organisations in successfully implementing Cloud BI security with a layered approach, that is discussed in the study.

Keywords: Cloud Business Intelligence, data security, data privacy issues, cloud computing, systematic literature review.

Table of Contents

Abstract 2

Table of Contents 3

List of Figures 6

List of Tables.. 7

Attestation of Authorship.. 8

Acknowledgements 9

Chapter 1: Introduction 10

Chapter 2: Background Literature 12

2.1. The Cloud Computing Context 12

2.1.1. Security Issues in Cloud Computing.. 13

2.2. Overview: Cloud Business Intelligence 14

2.2.1. The Conceptualisation of Cloud Business Intelligence 15

2.2.2. Challenges of Cloud Business Intelligence.. 17

2.3. Data Privacy in the Digital World.. 18

2.3.1. What Is Data Privacy?.. 18

2.3.2. Data Privacy in Cloud.. 21

2.3.3. Data privacy in Cloud Business Intelligence 24

Chapter 3: Research Method 27

3.1. Research Question 27

3.2. Systematic Review Steps 28

3.3. Thematic Analysis 33

Chapter 4: Results 35

4.1. Data privacy issues in Cloud BI 38

4.1.1. Data privacy issues in outsourcing 38

4.1.2. Lack of user training 39

4.1.3. Security risks in data migration 39

4.1.4. Privacy issues in SMEs 40

4.1.5. Similarities between Cloud security and Cloud BI security 41

4.2. Summary of security differences between Cloud BI and General Cloud Services. 42

4.2.1. Layers of Cloud BI security 43

4.2.2. Security cost 43

4.2.3. Management of security rules 44

4.2.4. OLAP security 44

4.2.5. Threat Management Approach 45

4.2.6. Security architecture 45

4.2.7. Security control 45

Chapter 5: Discussion & Conclusion 47

5.1. Discussion 47

5.2. Limitations 49

5.3. Future Research 49

5.4. Conclusion 50

References 51

Annexure A List of privacy issues in Cloud BI 56

Annexure B Evaluate the quality of reviewed papers 69

Annexure C Description of various types of delivery models 72

List of Figures

Figure 1 Business Intelligence Constituent Themes 15

Figure 2 The architecture of Cloud BI referenced by Willem and Jakobus (2010). 16

Figure 3 The different layers of the BI model (Juan-Verdejo, Surajbali, Baars, & Kemper, 2014) 24

Figure 4 The searched results in Scopus 31

Figure 5 The leader authors researched on Cloud BI 32

List of Tables

Table 1 Data Privacy Stakeholders (Venkataramanam & Shriram, 2017) 21

Table 2 Risks associated with cloud computing services (Clemons & Chen, 2011)...... 22

Table 3 The risk of the development of new technology (Clemons & Chen, 2011) 23

Table 4 Security assessment matrix, the evaluation level is divided into high, medium and low (the letters "H, M, L" are represented in the matrix). 26

Table 5 The steps of selection guidelines from Templier and Paré (2015). 29

Table 6 Summary of data privacy issues in Cloud BI.. 36

Table 7 Summary of Security differences between Cloud BI and Cloud services 43

Attestation of Authorship

I hereby declare that this submission is my own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person (except where explicitly defined in the acknowledgements), nor material which to a substantial extent has been submitted for the award of any other degree or diploma of a university or other institution of higher learning.

Signature：

Acknowledgements

Firstly, I would like to thank my supervisor, Dr. Ranjan Vaidya, for his assistance and conscientiousness throughout the research process. Due to the Covid-19 impact, I was unable to return to New Zealand to conduct research. He was still able to hold regular online meetings and guidance through each stage of the research. Besides, he was able to find mistakes in my dissertation and guide me onto the right track. I am deeply grateful for his extra contribution and guidance.

I also want to thank my mentor, Harminder Singh. Without his referral, I could not have met my supervisor. He provided correct guidance for my research topics and research methods and he has been following the progress of my thesis.

Finally, I would also like to thank my parents. Without their financial support, I would not have been able to get a quality education.

Thank you!

Chapter 1: Introduction

Big Data is identified as a strategic resource for organisations. It enables organisations to control information and improve efficiency and helps to inform business strategy (Balachandran & Prasad, 2017). Consequently, the adoption of big data technologies has shown a continuous increase since 2015 (Watson, 2019). To make better use of big data analysis, business intelligence (BI) applications are becoming more important. Many companies and organisations are looking to use business intelligence to help them make better and faster business decisions to gain competitive advantage (Llave, 2019). Moreover, BI helps organisations to not only optimise the operational aspects of the business but also to combine organisational strategy. Therefore, BI is important throughout the organisational hierarchy. However, the effective use of business intelligence requires data support and optimum storage space of the data. Implementing a BI system requires not only the purchase of a combination of software and hardware, but also the need to obtain the appropriate infrastructure and resources (Yeoh, Koronios, & Gao, 2008). To this end, most organisations and enterprises face many challenges in implementing business intelligence.

In order to solve these challenges, many companies are considering implementing BI on the cloud, so they can move a large amount of data to the cloud to reduce cost and increase flexibility and scalability. Literature shows that cloud computing provides not only the advantages of cost reduction but also a greater processing power and overall agility in data management (Thompson & Van der Walt, 2010).

However, there are some security challenges in implementing Cloud BI. Past studies have discussed the differences between the security challenges of cloud computing, and those of Cloud BI (Al-Aqrabi, Liu, Hill, Ding, & Antonopoulos, 2013). Cloud BI security is different from cloud security as Cloud BI consumes specific artefacts such as dashboards, Online Analytical Processing (OLAP) reports, and visualisation (Al-Aqrabi, 2016). Therefore, this paper puts forward the following two questions for research：

1. What are the data privacy issues in Cloud Business Intelligence?

2. How are data privacy issues in Cloud BI different from the data privacy issues in general cloud services?

When exploring the security issues of Cloud BI, the security issues of cloud computing provide a larger ambit for discussion. The following section will briefly describe the security issues of cloud computing, followed by a discussion of the challenges of Cloud BI.

This dissertation is structured as follows Chapter two presents the background literature by providing an overview of cloud computing and its security challenges. Further to this discussion, a conceptual overview of Cloud BI is provided. Chapter three describes the research method, which is the thematic analysis. Chapter four summarises the results of the systematic literature review, and Chapter five presents the discussion and conclusion.

Chapter 2: Background Literature

This chapter first explores the concepts and characteristics of cloud computing and Cloud BI, and summarises the challenges and security issues that they face. Also, since the research questions are related to the privacy and security issues of Cloud BI, this chapter also presents an overview of data privacy issues in the digital world. The terms 'data security' and 'data privacy' are both used in this paper largely to refer to data security. This usage is based on past studies that suggest privacy security is the core theme in security (Kang, Lee, Chun, & Song, 2007). The sub-sections below are discussing the cloud computing context, the concept of Cloud BI, and data privacy.

2.1. The Cloud Computing Context

The concept of cloud computing can be traced back to John McCarthy talk at the MIT Centennial in 1961, proposing the basic concept of providing computing power to the public as a public facility (Gillam & Antonopoulos, 2017). However, as the computer field and network technology were not very developed, his concept was not acceptable. It was not until the 20th century, that the term cloud computing resurfaced. As cloud computing has been studied by many scholars, different understandings and definitions have been developed. At present, the widely recognised definition of cloud computing comes from the National Institute of Standards and Technology (NIST) (Mell & Grance, 2011). According to the NIST definition, cloud computing is "a model for enabling ubiquitous, convenient, on-demand network access to a shared pool of configurable computing resources (e.g., networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort or service provider interaction" (Mell & Grance, 2011). Balachandran and Prasad (2017) have summarised important benefits of cloud computing: cost efficiency, continuous availability, scalability and elasticity, fast deployment and ease of integration, resiliency and redundancy, and increased storage capacity**.** Therefore, with the development of computer technology, cloud computing technology has been widely concerned by scholars in recent years and attempts to combine its technical characteristics with some existing technologies for innovation.

Given the characteristics and benefits of cloud computing above, many contemporary organisations try to combine cloud computing and business intelligence to solve the challenges of implementing business intelligence (Kasemsap, 2016). Organisations can leverage several types of cloud services using different types of delivery models to reduce storage costs and increase agility. For example, organisations can have their own private clouds, or they can share resources in the public clouds or hybrid cloud. Also, different types of resources and services can be hired in the cloud: infrastructure as a service (IAAS), Platform as a Service (PAAS) Software as a Service (SAAS) or Analytics as a Service (AAAS) (Clemons & Chen, 2011). Annexure C provides the definitions/ descriptions of each of these terms. Therefore, when business intelligence is combined with cloud computing, companies can not only move large amounts of data to cloud servers to reduce storage costs but also migrate business intelligence applications to the cloud to achieve high flexibility and scalability.

2.1.1. Security Issues in Cloud Computing

Cloud computing has many advantages but also many security issues. For example, Sun (2018) introduced the three dimensions of cloud computing security: computer security, network security and information security. Cloud-based business intelligence may also have some of the same types of security issues, and that is why it is relevant to briefly discuss the security issues of cloud computing. Similarly, some types of cybersecurity attacks are particularly relevant to Cloud BI, for example, DoS attacks at the software level, and eavesdropping and Spoofing at the network level can appear in Cloud BI because the general Cloud BI is based on the SaaS model.

Sun (2018) discusses the dimension of information security and explains the activity of identity management in organisations. It is a set of activities used to verify that users have access to a computed object, such as an application or system. Identity management is a core activity in security and is replete with security challenges. For example, password management is a challenge, especially in a distributed environment, which is a typical feature of computing. Also, when an authentication system is deployed in multiple geographies, such as in the cloud, the cost of identity management increases. Another challenging issue is the protection of identity information itself. In a distributed context, if an authentication system is connected to a remote server, an attacker has a greater opportunity to monitor the identity information. Thus, privacy issues are a major concern in cloud computing and often percolate down to the challenges faced by Cloud BI. The next section discusses Cloud BI and its security challenges.

2.2. Overview: Cloud Business Intelligence

In the earlier stages of conceptualisation, the term "intelligence" referred to automation, and a BI system was described as a system that automates information technology processes. For example, Luhn (1958) discussed the automated extraction and transfer of files. Dresner (2008) combines the concept of information technology with BI and describes it as a method and concept that helps businesses make decisions. Since then, more concepts about business intelligence have been defined by scholars. One comprehensive definition of BI comes from Watson (2009) who defines it, "BI is a broad category of applications, technologies, and processes for gathering, storing, accessing, and analysing data to help business users make better decisions".

Moreover, Grossmann and Rinderle-Ma (2015) have presented the purpose and characteristics of BI. The purpose of BI is to provide decision support for business objectives in different domains within the organisational and institutional frameworks. BI mainly relies on data to make decisions and requires different types of knowledge and theories to generate information. Furthermore, in making recommendations for BI implementation the authors mention that BI must be implemented with a systematic approach, utilising the practical capabilities of information and communication technology (ICT). Finally, the authors point out that BI systems must deliver information to the right people in the right form at the right time.

Figure 1 presents the architecture of BI, where the development of data warehouses, the use of analytics, and reporting in organisations are closely linked to BI. According to Kasemsap (2016), data warehouses, data sources, data marts and query and reporting tools are the major components of BI. These components are distributed in different layers, namely a presentation layer, a logic layer, and a data layer. In a cloud-based BI, data, the logic, the data access, and the presentation layers can be located across different service providers (Juan-Verdejo et al., 2014) and this can have implications for Cloud BI security.

In addition to the main components of BI mentioned above, other emerging technologies are also closely related to BI. These include big data analysis, text analysis, web analysis, network analysis, and mobile analysis. Chen, Chiang, and Storey (2012) explicitly analysed the characteristics and technical requirements of these emerging technologies. Therefore, this paper takes a holistic view of business intelligence, which also includes related technologies of data and knowledge mining, analytics, and reporting.

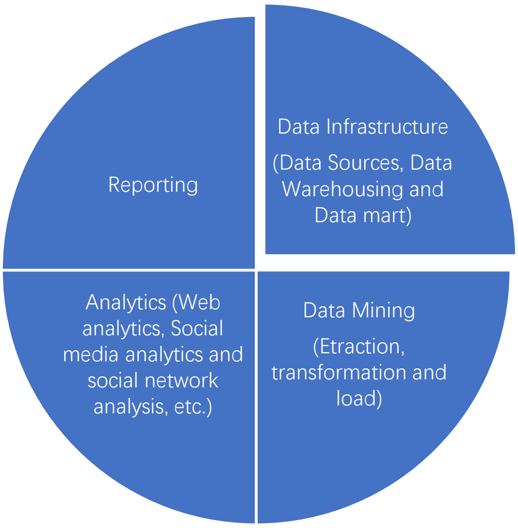

Reporting
Data Infrastructure
(Data Sources, Data Warehousing and Data mart)
Analytics (Web analytics, Social media analytics and social network analysis, etc.)
Data Mining
(Etraction, transformation and load)

Figure 1 Business Intelligence Constituent Themes

2.2.1. The Conceptualisation of Cloud Business Intelligence

Cloud BI refers to a set of business intelligence applications that are made available by the Cloud Service Providers (CSP) in the cloud. These applications can be used by the

Cloud Service Users (CSU) for meeting their strategic as well as day-to-day business intelligence requirements. Some examples of Cloud-based BI market leaders and their solutions are GoodData, SAP, MicroStrategy, Tableau, SAS, Oracle, IBM, Tibco, Qlik, Information Builders, Microsoft and Bime (Muntean, 2015). Herwig (2013) introduced that Cloud BI is a revolutionary concept of Business Intelligence, which is based on cloud architecture as a service with features of low cost, rapid deployment, and flexibility. Software as a service BI is also used by many small and medium-sized businesses that want to accelerate their business growth through BI and analytics tools. Cloud BI offers all the benefits of cloud computing in the context of business intelligence. These benefits primarily include cost-effectiveness, flexibility, scalability, reliability, and data sharing capabilities (Al-Aqrabi, Liu, Hill, & Antonopoulos, 2015).

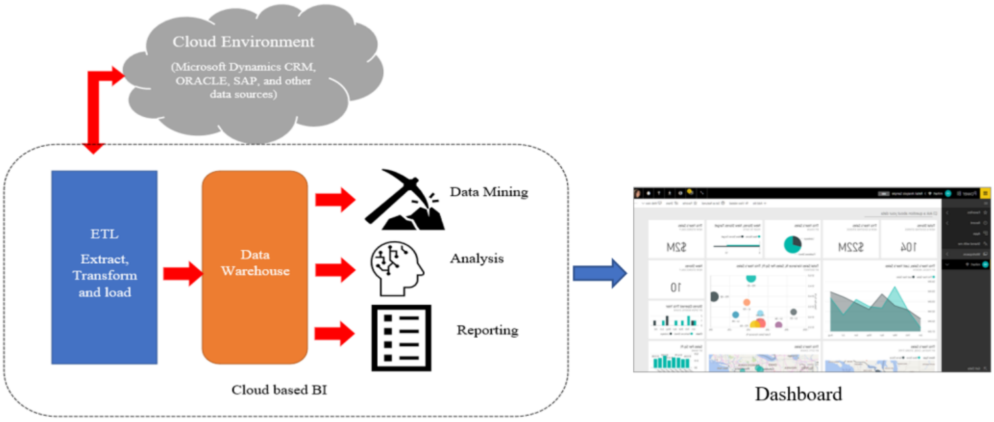

Cloud Environment
(Microsoft Dynamics CRM, ORACLE, SAP, and other data sources)
ETL
Extract, Transform and load
Data Warehouse
Data Mining
Analysis
Reporting
Cloud based BI
Dashboard

Figure 3 The architecture of Cloud BI referenced by Willem and Jakobus (2010).

Figure 2 shows the typical architecture of the Cloud BI. By using the Extract, Transform, Load (ETL) process, data from different source systems is loaded into the data warehouse of the cloud environment. The Cloud BI solution also allows users to execute data mining to obtain valuable data and access management reports. Moreover, a distinctive feature of Cloud BI is the dashboard, which supplies a visualisation of business execution information, making information easy to read, consult, and actively correct. The dashboard can run multiple reports at the same time to save time and quickly presents users with an overall visualisation of all aspects of the business. (Willem & Jakobus, 2010)

2.2.2. Challenges of Cloud Business Intelligence

Implementing a BI solution in the cloud not only brings benefits for businesses, but also brings new challenges that need to be addressed. One is security concerns at various levels of Cloud BI. According to Muntean's (2015) definition, cloud-based BI refers to any analytics effort that includes one or more elements in the cloud, whether public or private. The six elements are data sources, data models, processing applications, computing power, analytics models, and sharing or storing of results. However, most of the literature on Cloud BI conceptualises Cloud BI through a layered approach, which requires that Cloud BI security also adopts a layered approach. Juan-Verdejo et al. (2014) argue that various studies have taken a layered approach to Cloud BI, however, the Cloud BI security discussions fail to recognise the Cloud BI layers and security requirements at various layers. Juan-Verdejo et al. (2014) mention that the BI model generally follows a three-tier architecture, including a presentation layer, a logic layer, and a data layer. Therefore, when BI gets migrated to the cloud, all these layers can be distributed across the different cloud service providers, which may bring security challenges. These challenges are reflected in the following aspects: data security and trust, data ownership, the performance of the migrated BI system and the selection of cloud server products (Juan-Verdejo et al., 2014).

Apart from the above technical challenges, data privacy issues present a major challenge to the implementation of Cloud BI (Sheshasaayee & Margaret, 2015); The reason is that data interaction for Cloud BI uses Internet connectivity, potential Internet security threats will always exist. Subashini and Kavitha (2011) explained that a large amount of personal data and business data are currently placed in the cloud, people are paying more attention to the security of the environment. Likewise, Tamer, Kiley, Ashrafi, and Kuilboer (2013) also pointed out that data privacy issues should be a concern in Cloud BI because, as the data stored in the cloud database will be supervised by providers, there is a risk of personal privacy data leakage. Therefore, data privacy issues are the challenge of Cloud BI implementing.

In addition to the above concerns about data privacy in the cloud environment, on-premises sensitive data security also needs to be addressed. Because Cloud BI comprises not only cloud-based data warehouse, data marts, and reporting, but also, integration with on-premises data, sensitive on-premises data may need to be integrated with the data hosted on the cloud. Such integration is much needed for strategic decision making as strategic information is obtained by the integration of internal and external data (Turner & Lucas Jr, 1984). This suggests that not only the staff members working on Cloud BI applications but also those having access to on-premises sensitive data need to be made aware of the privacy issues in Cloud BI.

In summary, when BI migrates to the cloud, it faces many security challenges, among which data privacy is the main one. The next chapter will review the literature on data privacy and provides an overview.

2.3. Data Privacy in the Digital World

2.3.1. What Is Data Privacy?

The concept of data privacy was not defined until the late 20th century when it emerged with the development of the Internet and the popularity of computers and mobile communication devices. Sharma (2019) mentions that the concept of privacy comes from a school of thought in the West, with its original roots as a defence against state actions and privacy infringements. The traditional concept of privacy is to protect ourselves and our activities from outside interference, such as houses, documents, and personal lives. However, with the development of society and technology, we must share our personal information with authorised individuals or organisations, such as banks, security, health insurance companies, which hold and manage a large amount of our personal data. When sharing such sensitive data, identifying infringement activities is necessary. Kang et al. (2007) described six types of privacy infringement in their paper: information collection without consent, inappropriate monitoring by internet companies, inappropriate information analysis, inappropriate distribution of privacy information, the unauthorised push of advertising messages, and inappropriate storage of information. Therefore, in the face of privacy infringement, we must take effective measures to protect our legitimate rights and interests.

Regarding privacy issues, most countries in the world have laws or regulations to protect privacy issues, and they are consistent with global data privacy laws. Global Data Privacy (GDP) laws have been updated for 2019, bringing the number of countries with data privacy laws to 132 (Greenleaf, 2019). The data privacy laws of these jurisdictions cover both the private sector and the public sector in most cases and meet the minimum formal standards set out in international agreements (Greenleaf, 2019). In addition to General Data Protection Regulations (GDPR), the data privacy protection laws in various regions include the California Consumer Privacy Act (CCPA), the China Cyber Security Law (CCSL), Health Insurance Portability and Accountability Act (HIPAA) and the Gramm-Leach-Bliley Act (GLBA). These regulations limit the misuse of data by organisations (Greenleaf, 2019).

There are several types of data information involved in privacy:

- **Internet information** is the most important, it refers to all personal data left over in the online interaction process. For example, most websites collect, store, and track users' access trajectory and personal information through logs or cookies and push relevant content on customers. In another example, various search engines can collect and mine personal data, which leads to the risk of revealing personal information and data.

- **Financial information** is also affected by privacy because it stores copious amounts of sensitive personal information that can be used to commit fraud. Financial information includes information about a user's financial transactions, such as total assets, holdings of stocks or funds, which are sensitive. If criminals gain access to information such as someone's account or credit card number, the person could become a victim of fraud or identity theft.

- In the medical field, any **medical information** is protected by law and cannot be disclosed to third parties (Herring, 2014). Patients do not want their medical records to be available to unauthorised individuals and organisations, as such information can affect employment and insurance coverage (Appari & Johnson, 2010).

- With the popularity of mobile devices, **location information** is a privacy issue of great concern. According to research, four points of space-time (about where and when) are enough to uniquely identify 95 per cent of the 1.5 million people in the mobile database (De Montjoye, Hidalgo, Verleysen, & Blondel, 2013). In addition to location information, mobile devices store a large amount of sensitive personal data, such as photos, files, call records, and contacts. These data may be collected by a mobile phone application, which will involve data privacy issues (Kelley, Cranor, & Sadeh, 2013).

- **Political information** is also a concern such as when citizens are exercising their basic rights to vote by secret ballot, they want to ensure that no one knows their political views. If this information is leaked, there will be serious political consequences, such as political unrest.

When conducting data analysis in an organisation, there are many stakeholders involved. The different stakeholders that affect data privacy in organisations are listed in Table 1.

Stakeholder	Description
Company	Any organisation like a bank, an insurance company, or e-commerce, retail, healthcare, or social networking company that holds substantial amounts of customer-specific data. They are responsible for the data.
Customer/Record owner	An organisation’s customer could be an individual or another organisation that shares their data with the company.
Government	The government defines what data protection regulations the company should follow.
Data anonymizer	A person who anonymizes and provides data for analysis or as test data.
Data analyst	This person uses anonymized data to carry out data mining activities like prediction, knowledge discovery, etc.

Stakeholder	Description
Tester	Outsourcing of software testing is common among many companies.
Business operations employee	Data analysts and software testers use anonymized data that are at rest or static, whereas business operations employees access production data because they need to support customers' business requirements.
Adversary/data snooper	Data are precious and their theft is quite common. An adversary can be internal or external to the organisation.

Table 1 Data Privacy Stakeholders (Venkataramanam & Shriram, 2017)

2.3.2. Data Privacy in Cloud

To study the differences between cloud services, and on-premise services in the field of data privacy management, the literature on the risks and challenges of data privacy associated with cloud computing were reviewed. Bokhari, Shallal, and Tamandani (2016) mentioned that cloud computing has risks such as out of control, invalid storage, access control and data boundary, all of which are privacy issues that can arise in cloud computing. Ouahman (2014) also proposed that ensuring the security of cloud providers and networks needs to be taken seriously. A comprehensive understanding of the risks is provided by Clemons and Chen (2011) who describe many of the risks associated with cloud computing services, which are essentially the same as the risks associated with opportunistic behaviour, shirking, poaching, and opportunistic renegotiation of outsourcing contracts. A description of these risks is presented in Table 2.

Outsourcing opportunism risks	Description
Shirking and deliberate underperformance	This risk arises from suppliers deliberately underperforming and demanding full payment for services and is caused by information asymmetry. Such risks may manifest in cloud computing as: 1. Deliberately underinvesting in server capacity, resulting in slow performance, which may be attributed to the network rather than the service provider. 2. Deliberately underinvested standby, dynamic security system monitoring, and other data quality services can only detect when data loss or security violation occurs.

	3. Investment in excess or spare capacity is only detected when demand is highly correlated, because of market crises, major shopping holidays, and because these attacks are not detected until key events occur. The spike in demand for cloud computing was caused by events that led to a surge in transactions.
Poaching	This risk refers to the way that a supplier misuses data or programs existing on the vendor's site, such as stealing critical resources and supplying directly to a competitor to the detriment of customers. Users of e-government services will be greatly concerned at this type of risk.
Opportunistic renegotiation	This is probably the biggest risk in the cloud environment. Although SaaS can be made safe, other levels of cloud service provision (such as SaaS and PaaS) are very vulnerable to vendor lock-in and subsequent opportunistic repricing. This is because lack of interoperability at a higher level of SaaS and PaaS poses a significant risk to the client, as vendors are likely to exhibit opportunistic behaviour over time. Also, vendor lock-in is a problem widely studied by corporate buyers and government contractors, but it could become a bigger problem as e-government moves inexperienced government organisations into the procurement of online services.

Table 2 Risks associated with cloud computing services (Clemons & Chen, 2011).

The authors also discuss other risks: functionality risk, political risk, project risk, technical risk and financial risk. These risks are illustrated in Table 3.

Technology development risks	Description
Functionality Risk	Functional risk is raised because we are ignorant of the capabilities of cloud computing technologies. Such as, we do not know how dominant SaaS and PaaS will be in the future, or what applications developed in different clouds for enterprise consolidation or restructuring activities will require, or the laws governing the storage of cloud computing data in different countries.
Political Risk	Political risk arises when members of the group have reason to resist proposed new development efforts, often involving job security, status, or compensation. As cloud computing technology enables data to be stored in the cloud, this will reduce the number of staff within the organisation.

Project Risk	Project risk occurs when the combination of technologies or the scope of development efforts exceed the capabilities of the developer (the project is not completed on time and with high quality).
Technical Risk	Technical risk occurs when a project exceeds the capabilities of hardware or software technology or the skills of the best developers
Financial Risk	Financial risk is when a project fails to deliver the expected benefits. Although the benefits of cloud computing are many for some small and medium-sized enterprises, such as the huge operational cost reduction from the economies of scale of system management, due to their limited organisational size, they cannot receive benefit from load balancing opportunities without moving to the cloud. Similarly, for established companies, the benefits of IaaS and SaaS may be smaller than promised.

Table 3 The risk of the development of new technology (Clemons & Chen, 2011)

As mentioned at stakeholder in Table 1, there are many more stakeholders in data privacy of the cloud-based environment, such as cloud service providers and national governments. In cloud computing, the data processing chain involves multiple cloud service providers that serve the client. Besides, since most of the infrastructure involved in cloud computing services is located in different countries, national governments are also stakeholders (Maxim, 2015).

Some large global companies often outsource their business, operations, and processes to business processing outsourcing (BPO) companies and outsource their cloud servers. However, some BPO companies are related to financial services companies may be involved in conducting business operations with clients, such as securities trading, managing their financial transactions. However, access to customers' trading accounts in these processes can expose a large amount of sensitive information, which is unacceptable to customers, and national regulations do not allow such access (Venkataramanan &

Shriram, 2017). This data must be protected. In the next section, the data privacy issues in Cloud BI will be discussed.

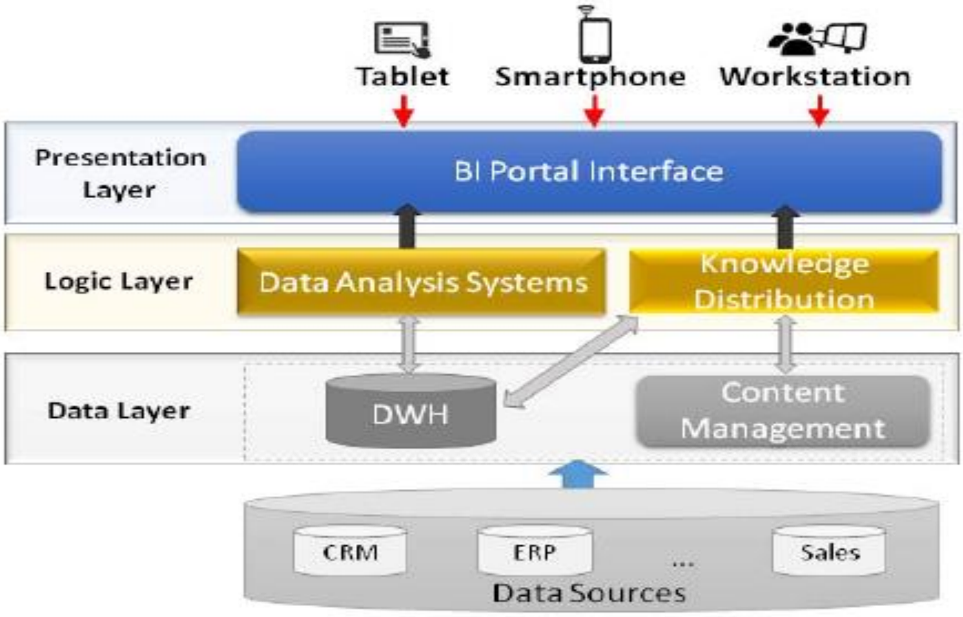


Figure 5 The different layers of the BI model (Juan-Verdejo, Surajbali, Baars, & Kemper, 2014)

2.3.3. Data privacy in Cloud Business Intelligence

Juan-Verdejo et al. (2014) introduced the management model of BI, as shown in Figure 3. This model conceptualises BI in form of a data layer, a logic layer and a presentation layer. The data layer stores structured and unstructured data from the operating system; the logic layer analyses these data differently; and the access layer allows BI users to access the advanced data they need. The author emphasises that when BI moves to the cloud environment, it is necessary to consider the security issues that occur at all layers of BI. These issues are the key factors that lead to data privacy and security.

At the data layer, due to the storage of a large amount of BI data, the reliability, security, and data integrity of the cloud vendors use for storage must be considered when migrating to the cloud environment. Furthermore, it is especially important to choose a cloud service model which involves data ownership. For example, as Juan-Verdejo et al. (2014) mentioned if you choose the SaaS or PaaS model, the data layer is managed by the software vendor, so there are uncontrollable risks in the data layer (shown as “M” in the Matrix). In the IaaS model, the organisation will be the owner of data rights, so the data will be directly controlled by the data owner (shown as “H” in the Matrix). Therefore, when an organisation chooses a cloud service model, it is important to encourage both parties to adopt transparent contracts and clarify applicable service level agreements (SLAs), laws, and regulatory policies. Moreover, organisations need to select a cloud environment based on the granularity level of cloud products and attributes such as scalability or service availability (Juan-Verdejo et al., 2014). Consequently, the organisation can clearly judge the future performance of the system, the security provided, and even the applicable laws after the migration.

The data privacy issue in Cloud BI is a subset of the data privacy issues of the cloud. For example, organisations may host their Cloud BI on the same hypervisor that is shared with other organisations in the same cloud. Within the cloud also, there are always risks related to poaching by the cloud service provider. These risks are also applicable to Cloud BI. However, apart from the common privacy risks, Cloud BI also has risks that are inherent to the architecture and structuring of Cloud BI. There are very few studies that have discussed these risks. This thesis explores the risks that befall Cloud BI and this section presents the unique aspects of Cloud BI from a security perspective.

The assessment of privacy in Cloud BI depends on the nature of cloud services availed as well as the BI layer. The levels of security in each situation can be different. This scenario is also explained in the matrix below (Table 4) that presents the security levels for the various intersections. The data layer can be further divided based on the sensitivity of the data. For example, an extremely sensitive data on the public cloud may be prone to higher risks. Moreover, apart from the basic three layers mentioned by Juan-Verdejo et al. (2014), there may be additional layers such as the network, transmission, session and authentication layers (Al-Aqrabi, 2016).

			Types of Cloud Services					
			Private Cloud	Public Cloud	Hybrid Cloud	IAAS	PAAS	SAAS
Cloud BI Layers	Data	Sensitive	H	L	M	H	M	M
		Public	H	L	M	H	M	M
	Logic		H	L	M	H	M	M
	Presentation		H	L	M	H	M	M

Table 4 Security assessment matrix, the evaluation level is divided into high, medium and low (the letters "H, M, L" are represented in the matrix).

Also, there are multiple physical objects in a Cloud BI that are embedded within these layers. For example, the typical objects may include a data warehouse, other data sources, data mart, query and reporting tools, and OLAP (Al-Aqrabi, 2016; Kasemsap, 2016). These objects may be embedded in the different layers described above. Thus, the privacy issues in Cloud BI have aspects that are unique to Cloud BI.

This paper focuses on data privacy issues with Cloud BI and explores how data privacy issues in Cloud BI are different from the data privacy issues in cloud services.

Chapter 3: Research Method

3.1. Research Question

The theme of this study is to explore the data privacy issues in Cloud BI. In the context of cloud computing, security issues have been widely discussed; however, very few studies have discussed the security issues of cloud-based business intelligence systems (Al-Aqrabi, 2016). For example, a layered approach to Cloud BI (discussed in the previous sections) will suggest that Cloud BI has a very specific structure and that security vulnerabilities can exist at the various layers of Cloud BI. Consequently, this research seeks answers to the following two research questions:

1. What are the data privacy issues in Cloud Business Intelligence?
2. How are data privacy issues in Cloud BI different from the data privacy issues in general cloud services?

In order to answer these questions, I have chosen a systematic literature review method. Compared to narrative review (Green et al, 2006) and literature review (Snyder, 2019), systematic review is more useful to identify, critically evaluate and integrate the findings of all relevant results from high-quality individual studies (Siddaway, 2014). Additionally, this method can explore and understand the existing research in a research field in a clear, rigorous and transparent way (Templier & Paré, 2015), so this approach is suitable for my research topic.

In this chapter, I explain the systematic review method. The last step of a systematic literature review is analysing the data. For data analysis, I have used thematic analysis and the constant comparison method for analysing the literature on Cloud BI. Both these methods are described in the sub-section on data analysis.

Since Cloud BI lies at the conjunction of both cloud computing and BI, I have decided to review the literature that includes both these themes and the theme of data privacy. Cloud

BI is closely related to terms such as analytics, big data, and cloud computing. I have thus also referred to related literature on big data analytics.

3.2. Systematic Review Steps

Templier and Paré (2015) have presented a 6-step model for conducting a systematic literature review. The implementation of this 6-step model is explained below.

Here are the 6 steps:

1. Formulating the research problem
2. Searching for the literature
3. Screening for inclusion
4. Assessing the quality
5. Extracting the data
6. Analysing the data

The detailed sub-steps are included in Table 6 below. In the section below, I describe the implementation of each of these steps.

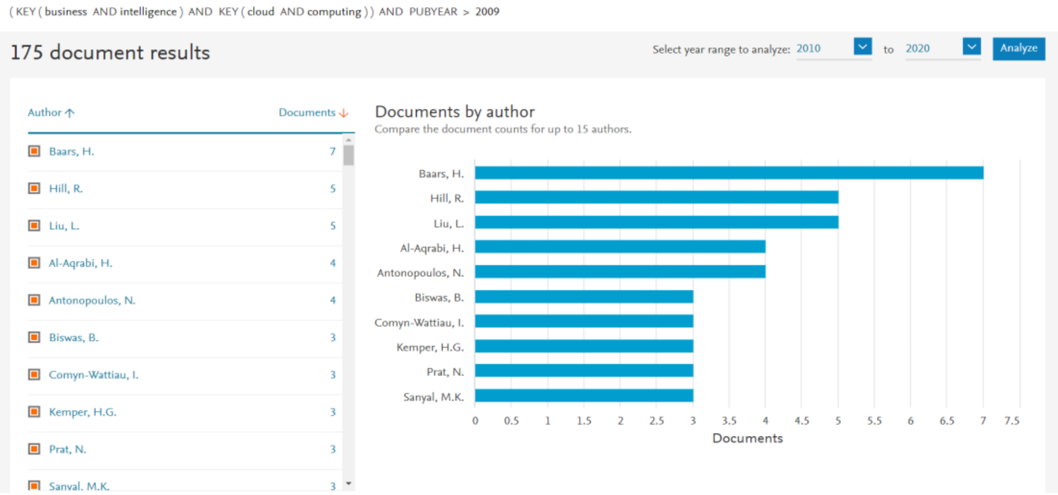
Steps	Recommended steps from Templier and Paré (2015)	Steps that were taken in this research
1	Formulating the problem 1.1 Specify the review's primary goal(s). 1.2 Clearly define the key concept(s) and establish the review's boundaries.	• Defined the research objectives and formulated the questions.
2	Searching for the literature 2.1 Specify the search procedures in sufficient detail. 2.2 Use a combination of data sources and search approaches.	• Identify potentially relevant research. • Used a combination of data sources and search approaches.

	2.3 Avoid restrictions that are not based on the research question(s). 2.4 Adopt strategies to minimise publication bias.	
3	Screening for inclusion 3.1 Specify the screening and selection procedures in sufficient detail. 3.2 Conduct parallel independent assessment of studies for inclusion. 3.3 Use inclusion criteria that reflect the research question(s). 3.4 Identify and be explicit about duplicate studies. 3.5 Include studies from reputable sources.	• Screened for inclusion to select the appropriate paper. • Conducted a parallel independent assessment of studies for inclusion. • Used inclusion criteria that reflected the research question(s). • Included studies from reputable sources.
4	Assessing quality 4.1 Use recognised quality assessment tools. 4.2 Consider the quality assessment in the selection of studies or the interpretation of the findings.	• Evaluated the quality of the research papers/literature.
5	Extracting the data 5.1 Specify the type of data to be extracted. 5.2 Use a structured procedure for data extraction. 5.3 Conduct parallel independent data extraction.	• Collected and extracted applicable information from each study. • Conducted parallel independent data extraction.
6	Analysing and synthesising data 6.1 Report the appropriate standards for the synthesis of the results. 6.2 Describe the logical reasoning and justifications behind the findings. 6.3 Provide a detailed summary of the included studies.	• Analysed and synthesised data .

Table 6 The steps of selection guidelines from Templier and Paré (2015).

In the first step of the process, the research objectives are defined, and the research questions are formulated. The research questions were formulated after an initial review of the selected set of studies on BI, data privacy, cloud computing. Through the initial review of the relevant papers I found that, while there was a lot of literature in the area of cloud computing, data privacy, and business intelligence, there was a dearth of literature in the area of Cloud BI related to data privacy.

The second step includes guidelines for identifying potentially relevant research. In this process, researchers determine the search scope, source, time span, and language (Levac, Colquhoun, & O'Brien, 2010). In this study, in order to ensure a comprehensive index, I combined multiple electronic data sources, namely Scopus, SpringerLink, Web of Science, Google Scholar, IGI Global, Wiley Online Library, IEEE Xplore and Applied Sciences Complete (EBSCO), to search related literature on the topic of my research. The searched results are shown in Figure 4. In this process, I used two search methods with keyword search terms: “cloud computing” and “privacy”, “Cloud BI” and “privacy”. Moreover, by applying the "filter function" to select the appropriate paper, the time searched was limited to between 2010 and 2020, and the language was “English” only. The purpose of this is able to locate the relevant literature on the topic. However, the search results were not very satisfactory. For example, when using “Cloud BI” and “privacy” as keywords, there were eight results in Scopus, and when using “cloud computing” and “privacy” as keywords there were 4660 results which seemed too large, and there were only a few articles that matched my topics, such as papers written by al-Al-Aqrabi (2016), Alsufyani and Chang (2015), and Balachandran and Prasad (2017). I carefully reviewed these articles’ Contents and References sections to find out the papers that matched my topic. Consequently, in order to explore more paper, I expanded my search keywords.



(KEY (business AND intelligence) AND KEY (cloud AND computing)) AND PUBYEAR > 2009

175 document results

Select year range to analyze: 2010 to 2020

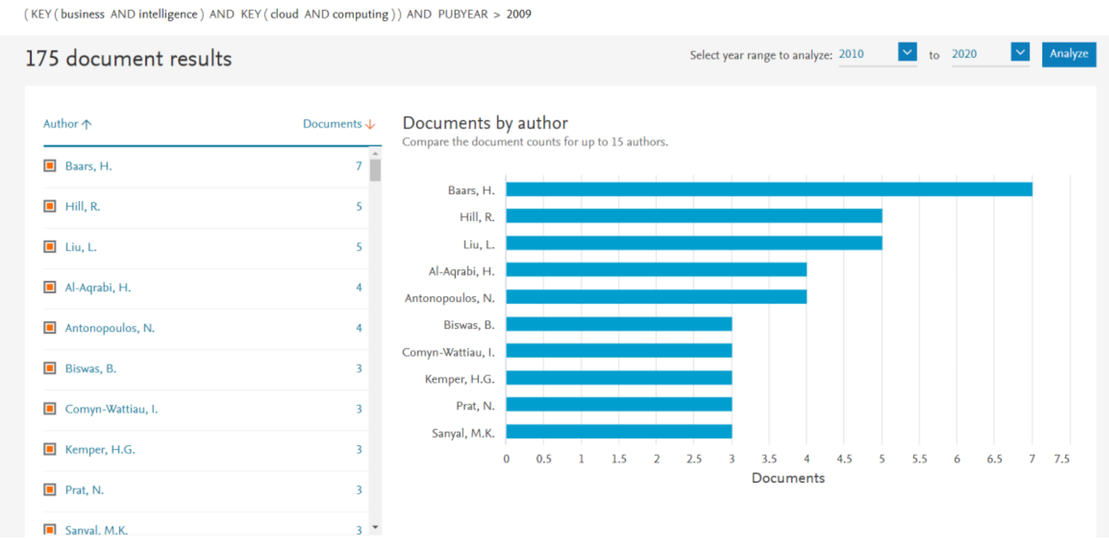
Author	Documents
Baars, H.	7
Hill, R.	5
Liu, L.	5
Al-Aqrabi, H.	4
Antonopoulos, N.	4
Biswas, B.	3
Comyn-Wattiau, I.	3
Kemper, H.G.	3
Prat, N.	3
Sanyal, M.K.	3

Documents by author

Compare the document counts for up to 15 authors.

Figure 8 The searched results in Scopus

Considering that business intelligence is linked to data analysis and big data, the search keywords also applied, such as “big data” and “data analytics” with “privacy”, which showed 294 results. Also, I used the keywords “security” instead of “data privacy” when I was trying to explore privacy issues in Cloud BI. Next, I used the Scopus analysis tool for the search results and present the graph shown in Figure 5. Scopus analysis tool is able to visualise, compare and export data to assess research trend, which can drive my decision making. The figure shows the names of several scholars who focus on the "Cloud Business intelligence", which enabled me to focus on the literature of these scholars to answer the research questions. According to the Scopus results, there were 25 relevant papers on my topic, which I shortlisted for further review.



(KEY (business AND intelligence) AND KEY (cloud AND computing)) AND PUBYEAR > 2009

175 document results

Select year range to analyze: 2010 to 2020 Analyze

Author	Documents
Baars, H.	7
Hill, R.	5
Liu, L.	5
Al-Aqrabi, H.	4
Antonopoulos, N.	4
Biswas, B.	3
Comyn-Wattiau, I.	3
Kemper, H.G.	3
Prat, N.	3
Sanyal, M.K.	3

Documents by author

Compare the document counts for up to 15 authors.

Figure 11 The leader authors researched on Cloud BI

An extensive and comprehensive search may yield many papers unrelated to the research question (Oxman & Guyatt, 1988). Therefore, it was necessary for me to select the appropriate papers in the initial search results in the third step. To demonstrate the effectiveness of the search results, all types of reviews should be combined with papers related to the topic of Cloud BI. To ensure that the quality of the selected literature, the supervisor will be involved to screen the quality and source. For cumulative and aggregated reviews, important efforts should be made in the selection of studies to enhance objectivity and avoid research bias. Duplicate viewpoints in the included literature were identified, and duplicated contents were evaluated and screened for deletion. Therefore, to ensure the quality of the literature, the following evaluation criteria were adopted.

1. The content of the paper specifically expounds the security of BI or cloud computing.
2. The content of the paper is instructive to help me explore my research topic.

The fourth step is to evaluate the quality of the papers. In the current scenario, I assessed the quality of the papers by looking at the source journals, and whether these journals were listed in the Australian Business Deans Council (ABDC) journal rankings. Given the dearth of studies on the topic of privacy in Cloud BI, there were very few studies, and hence some of these were recent without any citations. The citation counts of the shortlisted studies are included in 0. Although few papers did not have many citations, these papers were very useful and relevant to my research papers, so they were reviewed.

The fifth step is to collect and extract applicable information from each study. This step is important because the results of the extracted data are the main material for analysis. When extracting data, only the real data should be extracted, because if false or biased information is extracted, it will lead to the wrong description of the research in the following analysis, thus reducing the validity of the conclusion (Templier & Paré, 2015). Therefore, in the process of summarising the extracted information, I encoded and concept classified the topics to review and judge the validity of the information. In addition to categorising the extracted information, I also documented the authors who had aggregated conclusion and how they reached their conclusions, to identify the validity of their conclusions (Higgins & Green, 2008). To minimise deviations and errors of judgment in the process of data extraction, the supervisors also participated in the inspection of included literature, to check the process of information extraction and data accuracy, and discuss with me if there were different opinions.

The last step of a systematic review is to analyse the data. I chose to use the thematic analysis method for analysing the data. In the next sub-section, I explain the thematic analysis method and its applications in this research.

3.3. Thematic Analysis

Thematic analysis tends to identify the sub-themes and themes in a data corpus. This data corpus can be of various types including research papers. In this study, I use thematic analysis to identify the themes in the research papers. These themes will help me answer the research questions. There are many different versions of thematic analysis based on different philosophical and conceptual assumptions, and there are procedural differences, but in this paper, I am going to cite psychologist, Virginia Braun, whose concepts are widely recognised by scholars. According to statistics, her paper "Using thematic analysis in psychology", published in 2006, has been cited more than 70,000 times in Google scholar (Braun & Clarke, 2006). She advocates coding as the primary process of thematic analysis, which identifies items of interest in the data and tags them with coding tags. I will follow Braun and Clarke’s (2006) Six Phases of Thematic Analysis to review the results and generate reports.

Chapter 4: Results

Before presenting the findings of the literature review, I will briefly discuss the dearth of literature on Cloud BI and privacy that has constrained this study. While many studies discuss the privacy issue in the context of cloud computing, there are very few studies that discuss privacy issues in the context of Cloud BI. Thus, finding literature that exclusively discusses Cloud BI and the privacy issues has been a major constraint in this study. Nonetheless, there are 15 papers directly discuss privacy issues in cloud BI while 4 papers indirectly discuss the cloud BI. A summary of the reviewed papers is shown in Annexure A. In this section, I will discuss the answers to the two research questions presented in the "Research Question" section.

In order to explore privacy issues in Cloud BI, the data privacy issues of BI were first reviewed. Since the characterises of Cloud BI is to migrate the basic components of BI to the cloud, it is necessary to discuss the essential data privacy issues of BI then to compare and evaluate them with the findings of data privacy issues in Cloud BI. A summary of these results is shown below in Table 7 (the contents shown in the table correspond to the contents in Annexure A. Please review the Annexure A for details).

Code Number	Author (year)	Privacy issues	Threat Category (External/ Internal)
Privacy issues in BI			
1	Ali and Ouda (2017)	• Data breaches were caused by hacker attacks or mistakes by business partners. • User threat	External and Internal
2	Fernández-Manzano and González-Vasco (2018)	• User threat • Lack of training	Internal
3	Chen et al. (2012)	• Business Intelligence analysts lack knowledge of IT security	Internal
4	Gupta and Saxena (2010)	• Data privacy issues during the users sharing data	Internal
Privacy issues in Cloud BI			

5	Muntean (2015)	- Fraud, cyber-attacks, external risk, and legal issues. - Lack of security standards.	External and Internal
6	Qiu, Li, and Wu (2008)	- Lack of IT expertise and computer infrastructure. - Risk of outsourcing data mining tasks.	Internal and External
7	Mircea, Ghilic-Micu, and Stoica (2011)	- Data migration. - Lack of user training	External and Internal
8	Moyo and Loock (2019)	- Risk of public cloud service, cyber-attack. - Security breach.	External and Internal
9	Krishna Kagita (2019)	- The plug-in was destroyed by a network attack, botnet attack, identity theft, DoS, and man-in-the-middle attacks.	External
10	Elena Jaramillo, Munier, and Aniorte (2013)	- Data sharing and log file data).	Internal
11	Alsufyani and Chang (2015)	- Lack of training for employees, and lack of supporting resources. - Increased dependence on the third party, so it incurred the risk of losing control. - The protection of Cloud BI requires more budgetary cost.	Internal and External
12	Attasena, Harbi, and Darmont (2006)	- Cloud provider located in different countries raises complex legal issues.	External
13	Gurjar and Rathore (2013)	- Reliability of Cloud service providers, it may have risks of data security and lack of data control issues.	External
14	Tole (2014)	- Cyber attack，such as phishing sites or emails. - Cloud service providers have illegal or unauthorised access to users' data.	External

Table 7 Summary of data privacy issues in Cloud BI

Summarising the above findings, showed affecting factors of data privacy issues in BI, which are internal factors and external factors (Ali & Ouda, 2017). The internal factors are reflected in the leakage of the company's internal personnel data (Ali & Ouda, 2017;

Fernández-Manzano & González-Vasco, 2018), and the lack of IT security knowledge training for users (Chen et al., 2012; Fernández-Manzano & González-Vasco, 2018). According to Ali and Ouda (2017), more than 70 per cent of data breaches are caused by internal users (such as curious employees). When many internal users (such as analysts) have separate roles and responsibilities for accessing BI systems, they have the right to review some sensitive information for data analysis or data sharing. However, if they lack awareness of privacy, they can pose a serious threat to data security. Moreover, the participation of technologically illiterate users is another factor in data privacy breaches (Fernández-Manzano & González-Vasco, 2018). Therefore, it is necessary for companies to provide information security training to employees and develop data privacy protection rules. And for the BI analytics, they not only need the ability of business analysis, as well as, more importantly BI applications operating, but they should also have IT security knowledge (Chen et al., 2012).

The external factors that affect BI's data privacy issues are caused by hacking attacks and the mistakes of business partners (Ali & Ouda, 2017). Ali and Ouda (2017) described that external risks are caused by illegal intrusion/hacker attacks because BI has a large amount of valuable data which may become the target of hackers. They also stated that there is a risk that the data may not be controllable when shared with business partners.

In addition to the above factors, there are also data privacy issues in the data collection process. Many companies are now implementing business intelligence in their business processes to help them become more competitive in the marketplace. They collect vast amounts of data and integrate and analyse everything from social networks to financial transactions and shopping records. Data analysis is especially useful for forecasting and analysing, but it also leads to concerns and invasions of privacy. This is mainly because the BI extracting data from the data warehouse may automatically collect a large amount of users' confidential information (Te-Wei, Verbitskiy, & Yeoh, 2019). For example, the email provider may automatically scan an email from a user to infer the confidential information, and the user may agree to the provider's privacy policy without realising that their email is being analysed (Qiu et al., 2008). Therefore, data privacy issues may happen during the data collection process, which is worthy of our attention.

4.1. Data privacy issues in Cloud BI

As mentioned above, there are very few studies that discuss the data privacy issues in Cloud BI. However, there are some studies that have exclusively discussed this theme. Table 7 (Code Number from 5 to 14) presents the studies that discuss data privacy issues in Cloud BI. A summary of these results shows that they are also consistent with data privacy issues in BI as mentioned above. The following sub-sections discuss the finding of data privacy issues in Cloud BI, which will answer the research propositions.

4.1.1. Data privacy issues in outsourcing

Because Cloud BI has the characteristics of outsourcing, often in collaboration with other cloud server providers, there are risks of data security and data privacy. Through outsourcing, organisations can obtain specific human resources (such as skilled programmers) and technical resources (such as more powerful IT infrastructure) at a lower cost to meet their needs for IT services (such as data analysis) (Qiu et al., 2008). However, the practice of outsourcing data mining tasks involves extensive collaboration across different organisations (for example, exchanging or sharing data). Both the raw data and the information displayed after analysis contain the business intelligence (BI) and customer privacy of the organisation. Security issues with the potential risk of exposing private information in outsourcing activities are also mentioned in the article by Qiu et al. (2008). Furthermore, outsourcing increases the risk of losing control as it increases reliance on providers (Alsufyani & Chang, 2015). For example, when sensitive data is uploaded to cloud service providers in different countries or regions, complex legal issues arise and the security and control of the data will be threatened (Attasena et al., 2006; Muntean, 2015). Although companies can currently apply technologies such as transpositions, encryption, and data masking to ensure information security, these technologies are not suitable for storing data in the original format, especially when the report needs to be refined to a fine-grained situation (Gupta & Saxena, 2010). Besides, Tole (2014) explained that the BI solution provides security only at the UI level and that the data stored in the cloud database will be exposed to the provider, so cloud computing cannot guarantee security to cloud users. At the same time, Gupta and Saxena (2010) argued that to protect privacy and security, data cannot be copied and stored. Moreover, Gupta and Saxena (2010) also concluded that most algorithms used to protect privacy may not be suitable for data requirements related to data mining or analysis and processing. In summary, without proper security policies and technologies, privacy may be very vulnerable to security breaches.

4.1.2. Lack of user training

Given the above issues, since technical methods cannot solve these privacy issues well, we should try to find solutions at the management level. As described earlier, the lack of user training in Cloud BI is also an internal factor that makes data insecure (Mircea et al., 2011). Alsufyani and Chang (2015) mentioned that employee training is a way to improve employees' privacy awareness to deal with privacy issues. Employees engaged in Cloud BI not only need to receive training on the use of BI tools but more importantly, they also need to be made aware of potential privacy violations. It is necessary to add responsibility measures for data governance to its existing roles to ensure that the chain of responsibility for people is determined in the event of the data breach. Furthermore, analysts need to be aware of the sensitivity associated a data integration. Any data integration that uses sensitive data and stores it in the cloud can harm organisational data or expose it to potential competitors. Therefore, it is particularly important to improve the safety awareness of employees and to strengthen security management, which are important parts in the implementation of Cloud BI and cloud computing.

4.1.3. Security risks in data migration

It is precisely because of the characteristic of Cloud BI of flexibility and pursuing the definition of "everyone can access", Cloud BI requires cloud storage and sharing of data to achieve efficient office work. However, this characteristic makes Cloud BI less private (Alsufyani & Chang, 2015). Many security and privacy issues often arise in the process of data migration (Alsufyani & Chang, 2015). In order to work efficiently and flexibly, users of Cloud BI usually store substantial amounts of data in cloud servers. However, data security and privacy problems will arise during the process of data migration (Mircea et al., 2011). The issues include lack of governance, cyber-attacks, loss of control, and data insecurity or data incompleteness (Krishna Kagita, 2019; Mircea et al., 2011; Moyo & Loock, 2019; Muntean, 2015). Furthermore, Jaramillo et al. (2013) found that because the data of different users may be stored on the same cloud provider, cloud users are worried about the risk of their data being manipulated by other users on the cloud. This is consistent with the security issues in traditional cloud computing: availability, integrity, and confidentiality (Elena Jaramillo et al., 2013).

Furthermore, many companies will use SaaS business intelligence to meet efficient office anytime, anywhere. However, Mircea et al. (2011) stated that the migration of critical applications and infrastructure to the cloud involves more human resources to manage sensitive data and applications. Meanwhile, Mircea et al. (2011) also indicated that it is impossible to completely migrate BI applications to the cloud, so the migration process requires gradual and SaaS software and traditional software will co-exist in the organisation.

4.1.4. Privacy issues in SMEs

Cloud BI presents opportunities for small and medium enterprises, as well as security challenges (Moyo & Loock, 2019). Moyo and Loock (2019) stated that many SMEs are suffering from insufficient funds and lack of IT computing infrastructure and security technical support, which makes these enterprises often become the target of hacker attacks. This is because small and medium-sized companies are willing to adopt SaaS and the public cloud to achieve Cloud BI because of its low cost and IT technology resource requirements (Moyo & Loock, 2019). However, unlike private clouds, low-cost or free public clouds face serious security challenges when hosting BIs and sensitive data (Moyo & Loock, 2019). And by exploiting SaaS security vulnerabilities, hackers can compromise the security of any customer data by using simpler ways to bypass security controls (Moyo & Loock, 2019).

One type of attack on a company is cyber-attack, Krishna Kagita (2019) analysed the security and privacy issues of business intelligence in Internet of Things (IoT). In his paper, he introduced that Cloud BI mainly relies on plug-ins that require Internet connections. If the browser is not protected by security system, attackers can execute and destroy related plug-ins in the system's browser. The author also introduced that lack of security equipment will allow attackers to track the details of the system and carry out attacks or steal data. There are different types of IoT attacks in business intelligence: botnet attacks, identity theft, denial of service and man-in-the-middle attacks. Krishna Kagita (2019) stated that in the case of a cloud attack, the data that an organisation stores on the server for data analysis may be affected. In addition to the above types of network attacks, Sun (2018) also introduced the types of attacks currently faced by cloud computing.

The next section discusses the differences between data security issues in cloud computing and Cloud BI. Before, discussing the differences between the two, I briefly discuss the similarities between cloud computing security and Cloud BI security.

4.1.5. Similarities between Cloud security and Cloud BI security

E. Jaramillo, Munier, and Aniorté (2013) mentioned that Cloud BI also involves some security issues similar to cloud computing. The first similar issue is about responsibility (E. Jaramillo et al., 2013). Because business processes are orchestrated through multiple domains and/or cloud providers, it is difficult to detect failures and data leaks to establish chains of responsibility (E. Jaramillo et al., 2013). E. Jaramillo et al. (2013) mentioned the issue of cloud users not having direct control over processes and data. This problem leads to legal aspects and possible conflicts of data across borders. With regard to data availability concerns, Moyo and Loock (2019) also illustrated that Cloud BI has the risk of data availability, because when the network connection is damaged or the cloud provider suspends the service, the data in Cloud BI cannot be operated normally. Also, due to the different operating platforms and service providers of Cloud BI, this can lead to a lack of standards in the interaction process that can hinder interoperability (Gurjar & Rathore, 2013). E. Jaramillo et al. (2013) also discussed security concerns of cloud users who fear that data may be manipulated by another cloud user from a shared resource with the same vendor. Moreover, there is also the risk that monitoring activity logs will be disclosed in Cloud BI, leading to privacy issues. This is because the activity log contains a lot of workflow and user operational information (including accounts and passwords, etc.), and the leakage of this data can raise security issues (E. Jaramillo et al., 2013). The security issues of these Cloud BI are similar to the security issues of cloud computing.

4.2. Summary of security differences between Cloud BI and General Cloud Services.

Table 8 is a summary of the results of the different security between Cloud BI and Cloud services. There are major differences between the security of Cloud BI and cloud computing. These differences are in the areas of layers of Cloud BI security, security cost different, management of security rules, OLAP security, threat management approach, security architecture and security control.

Code number	Author (year)	Different security/ privacy issues between Cloud BI and general cloud services
15	Juan-Verdejo et al. (2014)	Since Cloud BI security should adopt a layered approach, it is important not only to recognise the Cloud BI layer but also to understand the security requirements at various layers when discussing Cloud BI security.
16	Al-Aqrabi et al. (2013)	• The Cloud BI security system should especially include the metadata of the integrated data. • Security cost is different. • Consider not only technical safety but also the safety of the personnel involved. • Uses of Mix models (UTM and DTM) for security control of Cloud BI. • OLAP security.
17	Witti, Guegan, and Benkhelifa (2018)	The security requirements in a multi-cloud environment, which may differ from Cloud BI security requirements.
18	Ouf and Nasr (2011)	The security requirements in a multi-cloud environment, which may differ from Cloud BI security requirements.
19	Al-Aqrabi (2016)	The BI framework will require access control of various BI processes and all authentication layers.

		Increased security verification may cause network overload, and BI security control is different from traditional database applications.

Table 8 Summary of Security differences between Cloud BI and Cloud services

4.2.1. Layers of Cloud BI security

One distinguishing aspect of conceptualising Cloud BI is through the layers that distinguish it from cloud computing. Cloud BI is conceptualised along with a data layer, a logic layer, and a presentation layer (Juan-Verdejo et al., 2014). Other studies have also followed similar conceptualisation and have suggested that in addition to the basic three layers, Cloud BI may have other layers, such as network, transport, session, and authentication layers (Al-Aqrabi, 2016). For example, Muntean (2015), proposed that Cloud BI comprises of six elements: data models, processing applications, computing power, analytic models, and sharing or storing of results such as in reports or dashboards. Since Cloud BI security should adopt a layered approach, it is important not only to recognise the Cloud BI layer but also to understand the security requirements at various layers when discussing Cloud BI security (Juan-Verdejo et al., 2014). Besides, the assessment of privacy in Cloud BI depends on the nature of the cloud service used and the BI layer, and the security level may be different in each case. Witti et al. (2018) and Ouf and Nasr (2011) have the same point of view; and they believed that the security requirements in a multi-cloud environment, which may differ in Cloud BI security requirements. As pointed out by Al-Aqrabi (2016), BI framework will need access controls at various BI process and all layers of authentication. Therefore, the privacy and security issues in Cloud BI are unique.

4.2.2. Security cost

The security cost is an important distinguishing characteristic of Cloud BI. Because Cloud BI requires multiple layers of protection and always includes backup cloud storage, effective protection will require more budgetary costs (Alsufyani & Chang, 2015). Al-Aqrabi et al. (2013) stated that in Cloud BI, all layers including sessions, presentations, applications, networks, and transmissions require security control. Mircea et al. (2011)

also explained that migrating BI's critical applications and infrastructure to cloud services involves more human resources to manage sensitive data and applications, which will affect the security costs of Cloud BI solutions. Therefore, the security cost of Cloud BI is different from cloud security.

4.2.3. Management of security rules

The security considerations in Cloud BI may transcend technical security to people who play a significant role in the management of business rules (Al-Aqrabi et al., 2013). Al-Aqrabi et al. (2013) explain that Cloud BI will require frequent human intervention and the need to follow business rules that may change over time. This is because Cloud BI security management is necessary to identify the authority of data access and control. As mentioned by Moyo and Loock (2019), managers must have more security assessment knowledge so that they can independently make the right judgment when choosing the cloud service provider that is appropriate for the need of Cloud BI (Moyo & Loock, 2019). Therefore, Cloud BI security needs to pay more attention to the security management of participants and requires higher security knowledge requirements for people at management level.

4.2.4. OLAP security

Security in the cloud may only focus on front-end application (browser) security and database security, but Cloud BI security includes OLAP application server security hosting the OLAP dashboard (Al-Aqrabi, 2016). Moreover, because some vendors, such as SaaS, PaaS, and IaaS, may use different frameworks, implementing Cloud BI requires coordination of architectural details to design and deploy services to support the various layers of BI and OLAP frameworks (Al-Aqrabi, 2016). Therefore, when designing and deploying a security framework for Cloud BI one should consider whether the security framework supports OLAP and the various layers of Cloud BI.

4.2.5. Threat Management Approach

The threat management approach in Cloud BI is different from cloud computing. Al-Aqrabi (2016) tested separately the “Distributed threat management approach (DTM)” and “Unified Threat Management Approach (UTM)” on the Cloud BI. The results have shown that using either approach alone in Cloud BI is not feasible, so Al-Aqrabi (2016) suggested that the BI security model on the cloud should apply a unified threat management approach for security control at layers of the network, transport, session, and presentation, and apply distributed security components for the application layer. Therefore, to ensure the security of Cloud BI, one should use both UTM and DTM approaches (Al-Aqrabi, 2016).

4.2.6. Security architecture

Because Cloud BI has a special security framework and contains a multi-dimensional security architecture in database objects, a large number of interactions may be generated in the BI framework during security control and data encryption at various levels. This makes the BI framework heavier in terms of computer overload and capacity. Besides, as security verification increases, BI frameworks may also overload the network (Al-Aqrabi, 2016). Tole (2014) also explained that in order to protect users from providers, Cloud BI allows users to encrypt data to protect sensitive data, which may result in slower response times for data loading. Therefore, Cloud BI security architectures are different from traditional cloud computing, and will require more network capacity and speed of the internet.

4.2.7. Security control

Another point of difference between security of Cloud BI and cloud security is that security control of Cloud BI system should especially include the metadata of the integrated data (Al-Aqrabi et al., 2013). BI involves integrated information, which is multidimensional, strategic, and extremely sensitive. If system security is compromised, unauthorised users can easily access integrated sensitive information (Al-Aqrabi et al., 2013). Thus, the security of integrated data needs to be handled differently in Cloud BI than the security of the imaged online data. Therefore, security controls are required at metadata of the integrated data, while that may not be necessarily the case in cloud computing data.

The above findings have answered the corresponding proposition, which showed the privacy issues related to Cloud BI, and the differences between Cloud BI and cloud computing in terms of data privacy and security. The findings are summarised and discussed in the next section.

Chapter 5: Discussion & Conclusion

In this chapter, a comprehensive discussion will be conducted based on the research results, limitations of the research will be presented, and the future directions for further research will be suggested.

5.1. Discussion

Moving important components of BI to the cloud is significant. The cost efficiency, flexibility, scalability, reliability, and data sharing capabilities of Cloud BI provide enterprises with competitive advantages. However, there are several security challenges in the migration of BI components to cloud servers, and data privacy is a particular concern for many enterprises because Cloud BI usually collects and analyses some users' sensitive information for business purposes, and stores and uploads information to the cloud server. However, once sensitive information is leaked, it will have a serious impact on the company. The previous chapter has summarised the findings of data privacy issues in Cloud BI, which contains some internal and external factors, and other data privacy issues. Furthermore, because Cloud BI has integrated cloud computing technology and BI technology, when exploring the data privacy issues of Cloud BI, data privacy issues in cloud computing are also involved. Through a systematic literature review, it is found that Cloud BI has some data privacy issues similar to cloud computing, but also has some unique security differences. These unique security differences include layers of Cloud BI security, security cost, management of security rules, OLAP security, threat management approach, security architecture and security control. The above research findings will provide some important contributions to companies when implementing Cloud BI strategies.

One contribution is that the study indicated Cloud BI is conceptualised as a layered phenomenon. In this phenomenon, there are the data layer, logic layer, and presentation layer. Therefore, privacy security management needs to be implemented through all three layers. The current study provides an outline for a framework that can be used for developing a Cloud BI security framework. In this framework, Cloud BI security is assessed at each of the three layers. Also, each of these layers can be outsourced to different cloud service providers. Therefore, the framework will also need to include not only the layers at which the security is assessed but also the distribution of the layers to different cloud service providers. Similarly, the cloud services that are availed can be of different types such as SaaS, IaaS, and PaaS. Therefore, a layered framework for Cloud BI security will include dimensions of Cloud BI Layers (data, logic, presentation), Cloud Services (IaaS, PaaS, SaaS), and Cloud Type Layers (private or public). This is an important contribution of this dissertation as existing security studies have neglected the security implications of the layered approach to Cloud BI. Past studies have also suggested that Cloud BI security discussions have failed to recognise the different layers of Cloud BI, and hence the security implications at each of these layers were hardly discussed (Juan-Verdejo et al., 2014).

The study also found that the data privacy issues of the BI software itself are a combination of internal and external factors, which also appear in Cloud BI. In addition to these influencing factors, Cloud BI data privacy issues are also reflected in infrastructure outsourcing, lack of user security training, and data migration. Besides, data privacy security is particularly important for small and medium-sized enterprises to implement Cloud BI, because their weak IT infrastructure is often the most vulnerable to hackers. These findings will help companies to successfully implement Cloud BI and provide valuable information on data security and privacy, avoiding losses.

However, the rate of return of Cloud BI for small and medium enterprises (SMEs) needs to be considered in advance. As explained in previous chapters, ensuring Cloud BI security requires more people to participate and more security costs. To ensure Cloud BI security may be difficult in SMEs, as they are weak in infrastructure, so that the risks posed by privacy and data security may be higher. Therefore, it is necessary for SMEs to evaluate the rate of return on Cloud BI and analyse whether they need Cloud BI.

5.2. Limitations

There are a few major limitations in this study that should be addressed in future research. First, because this article adopts a systematic review method to evaluate and analyse secondary data, the results of the literature review could not produce convincing evidence for the research topic.

Second, as this paper's research was on an emerging research topic, it was limited by privacy and security issues in Cloud BI. Moreover, due to the limitations of the sample, this study could not obtain more accurate results. It was difficult to identify the quality of the selected papers for review, as some papers were not widely cited because they were just published. Therefore, the small sample size is a limitation of this study.

Third, as the current 5G technology matures, the issues of network overload and slow loading speed caused by the protection of Cloud BI data security may be improved or resolved. Shorgin, Samouylov, Gudkova, Galinina, and Andreev (2014) introduced that 5G wireless technology brings more possibilities to cloud computing because it is safer and faster.

5.3. Future Research

Based on the limitations described above, I have summarized the following directions for future research. First, researchers could investigate in more detail data privacy issues in Cloud BI based on the findings presented in this study. Researchers could use interviews or surveys to obtain data for analysis and comparison to verify some of the findings in this study. For example, researchers could conduct surveys with Cloud BI software vendors or Cloud BI users to investigate the privacy issues of the data they provide, then compare and evaluate the data obtained from the survey with the data shown in the relevant literature. Research results obtained in this way would be more accurate and complete.

Another research direction would be to explore what impact the implementation of 5G technology will have on Cloud BI and what data security issues can be solved by 5G technology.

5.4. Conclusion

This research first reviewed the background of Cloud BI and cloud computing and observed some challenges that Cloud BI faces, among which data privacy issues are particularly important for organisations. To this end, this research focused on the data privacy issues of Cloud BI. The research method adopted was a systematic literature review, and relevant literature was analysed using the thematic analysis approach. The study found that internal and external organisational factors have an impact on Cloud BI privacy. The research also discussed security issues in cloud computing and discovered the similarities and differences between cloud computing and Cloud BI. The research results have provided valuable reference information for enterprises to successfully implement Cloud BI and ensure information security.

References

Al-Aqrabi, H. (2016). Cloud BI : a multi-party authentication framework for securing business intelligence on the Cloud: University of Derby.

Al-Aqrabi, H., Liu, L., Hill, R., & Antonopoulos, N. (2015). Cloud BI: Future of business intelligence in the Cloud. *Journal of Computer and System Sciences*, *81*(1), 85-96. https://doi.org/10.1016/j.jcss.2014.06.013

Al-Aqrabi, H., Liu, L., Hill, R., Ding, Z., & Antonopoulos, N. (2013). Business intelligence security on the clouds: Challenges, solutions and future directions*IEEE*. Symposium conducted at the meeting of the 2013 IEEE Seventh International Symposium on Service-Oriented System Engineering

Ali, O., & Ouda, A. (2017). A content-based data masking technique for a built-in framework in Business Intelligence platform (pp. 1-4): IEEE.

Alsufyani, R., & Chang, V. (2015). Risk analysis of business intelligence in cloud computing. *2015 IEEE 7th International Conference on Cloud Computing Technology & Science (CloudCom)*, 558.

Appari, A., & Johnson, M. E. (2010). Information security and privacy in healthcare: current state of research. *International journal of Internet and enterprise management*, *6*(4), 279-314.

Attasena, V., Harbi, N., & Darmont, J. (2006). Sharing-based privacy and availability of cloud data warehouses

Balachandran, B. M., & Prasad, S. (2017). Challenges and benefits of deploying big data analytics in the cloud for business intelligence. *Procedia Computer Science*, *112*, 1112.

Bokhari, M. U., Shallal, Q. M., & Tamandani, Y. K. (2016). Security and privacy issues in cloud computing. *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, 896.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Chen, H., Chiang, R. H., & Storey, V. C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, *36*(4), 1165-1188. https://doi.org/10.2307/41703503

Clemons, E. K., & Chen, Y. (2011). Making the decision to contract for cloud services: Managing the risk of an extreme form of IT outsourcing*IEEE*. Symposium conducted at the meeting of the 2011 44th Hawaii International Conference on System Sciences

De Montjoye, Y.-A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the crowd: The privacy bounds of human mobility. *Scientific reports, 3,* 1376. https://doi.org/10.1038/srep01376

Dresner, H. (2008). *The performance management revolution: Business results through insight and action*: John Wiley & Sons.

Fernández-Manzano, E.-P., & González-Vasco, M.-I. (2018). Analytic surveillance: Big data business models in the time of privacy awareness. *El profesional de la información (EPI), 27*(2), 402-409. https://doi.org/10.3145/epi.2018.mar.19

Gillam, L., & Antonopoulos, N. (2017). *Cloud computing: principles, systems and applications* (Vol. 2nd ed) [Book]: Springer.

Greenleaf, G. (2019). Global data privacy laws 2019: 132 national laws & many bills.

Green, B. N., Johnson, C. D., & Adams, A. (2006). Writing narrative literature reviews for peer-reviewed journals: secrets of the trade. Journal of chiropractic medicine, 5(3), 101-117.

Grossmann, W., & Rinderle-Ma, S. (2015). *Fundamentals of business intelligence*: Springer.

Gupta, V., & Saxena, A. (2010). Privacy Layer for Business Intelligence*Springer.* Symposium conducted at the meeting of the International Conference on Network Security and Applications

Gurjar, Y. S., & Rathore, V. S. (2013). Cloud business intelligence–is what business need today. *International Journal of Recent Technology and Engineering, 1*(6), 81-86.

Herring, J. (2014). *Medical law and ethics*: Oxford University Press, USA.

Herwig, V. (2013). Business intelligence as a service for Cloud-based applications*IEEE.* Symposium conducted at the meeting of the 2013 IEEE 7th International Conference on Intelligent Data Acquisition and Advanced Computing Systems (IDAACS)

Higgins, J. P., & Green, S. (2008). Defining the review question and developing criteria for including studies. In *Cochrane handbook for systematic reviews of interventions* (Vol. 1, pp. 83): Wiley Online Library.

Jaramillo, E., Munier, M., & Aniorté, P. (2013). Information security in business intelligence based on cloud: A survey of key issues and the premises of a proposal Symposium conducted at the meeting of the Proceedings of WOSIS 2013: 10th International Workshop on Security in Information Systems - In Conjunction with the 15th International Conference on Enterprise Information Systems, ICEIS 2013 Retrieved from https://www.scopus.com/inward/record.uri?eid=2-s2.0-84887568907&partnerID=40&md5=f2f06b5ec3601998752d8538840e49e0

Juan-Verdejo, A., Surajbali, B., Baars, H., & Kemper, H.-G. (2014). Moving business intelligence to cloud environments*IEEE.* Symposium conducted at the meeting of the 2014 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS)

Kang, Y., Lee, H., Chun, K., & Song, J. (2007). Classification of privacy enhancing technologies on life-cycle of information*IEEE.* Symposium conducted at the meeting of the The International Conference on Emerging Security Information, Systems, and Technologies (SECUREWARE 2007)

Kasemsap, K. (2016). Implementing business intelligence in contemporary organizations. In *Business Intelligence* (pp. 33-48). Hershey, PA, USA: IGI Global. Retrieved from http://services.igi-global.com/resolvedoi/resolve.aspx?doi=10.4018/978-1-4666-9562-7.ch002. https://doi.org/10.4018/978-1-4666-9562-7.ch002

Kelley, P. G., Cranor, L. F., & Sadeh, N. (2013). Privacy as part of the app decision-making process Symposium conducted at the meeting of the Proceedings of the SIGCHI conference on human factors in computing systems

Krishna Kagita, M. (2019). Security and Privacy Issues for Business Intelligence in lOT (pp. 206-212): IEEE.

Levac, D., Colquhoun, H., & O'Brien, K. K. (2010). Scoping studies: advancing the methodology. *Implementation science, 5*(1), 69.

Llave, M. R. (2019). A Review of business intelligence and analytics in small and medium-sized enterprises. *International Journal of Business Intelligence Research, 10*(1), 19-41. https://doi.org/10.4018/ijbir.2019010102

Luhn, H. P. (1958). A business intelligence system. *IBM Journal of research and development, 2*(4), 314-319.

Maxim, M. (2015). The rights and obligations of the main stakeholders in cloud computing services. *Perspectives of Business Law Journal*(1), 190.

Mell, P., & Grance, T. (2011). The NIST definition of cloud computing.

Mircea, M., Ghilic-Micu, B., & Stoica, M. (2011). Combining business intelligence with cloud computing to delivery agility in actual economy. *Economic Computation & Economic Cybernetics Studies & Research, 45*(1), 1-16.

Moyo, M., & Loock, M. (2019). *Small and Medium-Sized Enterprises' Understanding of Security Evaluation of Cloud-Based Business Intelligence Systems and Its Challenges* (Vol. 973) [Conference Paper]: Springer Verlag. Retrieved from http://ezproxy.aut.ac.nz/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-85066911081&site=eds-live. https://doi.org/10.1007/978-3-030-11407-7_10

Muntean, M. (2015). Considerations Regarding Business Intelligence in Cloud Context. *Informatică economică, 19*(4/2015), 55-67. https://doi.org/10.12948/issn14531305/19.4.2015.05

Ouahman, A. A. (2014). Security and privacy issues in cloud computing. *Journal of Defense Resources Management (JoDRM), 5*(2), 99-108.

Ouf, S., & Nasr, M. (2011). Business intelligence in the cloud*IEEE*. Symposium conducted at the meeting of the 2011 IEEE 3rd International Conference on Communication Software and Networks

Oxman, A. D., & Guyatt, G. H. (1988). Guidelines for reading literature reviews. *CMAJ, 138*(8), 697-703.

Qiu, L., Li, Y., & Wu, X. (2008). Protecting business intelligence and customer privacy while outsourcing data mining tasks. *Knowledge & Information Systems, 17*(1), 99.

Sharma, S. (2019). *Data Privacy and GDPR Handbook*: John Wiley & Sons, Incorporated.

Sheshasaayee, A., & Margaret, T. S. (2015). The challenges of business intelligence in cloud computing. *Indian Journal of Science and Technology, 8*(36), 1-6. https://doi.org/10.17485/ijst/2015/v8i36/88493

Shorgin, S., Samouylov, K., Gudkova, I., Galinina, O., & Andreev, S. (2014). On the benefits of 5G wireless technology for future mobile cloud computing (pp. 1-4): IEEE.

Siddaway, A. (2014). What is a systematic literature review and how do I do one. *University of Stirling*(I), 1.

Snyder, H. (2019). Literature review as a research methodology: An overview and guidelines. Journal of Business Research, 104, 333-339.

Subashini, S., & Kavitha, V. (2011). A survey on security issues in service delivery models of cloud computing (Vol. 34, pp. 1-11).

Sun, X. (2018). Critical security issues in cloud computing: a survey (pp. 216-221): IEEE.

Tamer, C., Kiley, M., Ashrafi, N., & Kuilboer, J.-P. (2013). Risk and benefits of business intelligence in the cloud. *Proceedings for the Northeast Region Decision Sciences Institute (NEDSI)*, 86-95.

Te-Wei, W., Verbitskiy, Y., & Yeoh, W. (2019). Depicting Data Quality Issues in Business Intelligence Environment through a Metadata Framework. In *Applying Business Intelligence Initiatives in Healthcare and Organizational Settings* (pp. 291-304): IGI Global.

Templier, M., & Paré, G. (2015). A framework for guiding and evaluating literature reviews. *Communications of the Association for Information Systems, 37*, 112-137. https://doi.org/10.17705/1cais.03706.

Thompson, W. J., & Van der Walt, J. S. (2010). Business intelligence in the cloud. *South African Journal of Information Management, 12*(1), 1-15.

Tole, A. A. (2014). Cloud computing and business intelligence. *Database Systems Journal, 5*(4), 49.

Turner, J. A., & Lucas Jr, H. C. (1984). Developing strategic information systems.

Venkataramanan, N., & Shriram, A. (2017). *Data privacy : principles and practice* [Electronic document]: CRC Press.

Watson, H. J. (2009). Tutorial: Business intelligence-Past, present, and future. *Communications of the Association for Information Systems, 25*(1), 39.

Watson, H. J. (2019). Update tutorial: Big Data analytics: Concepts, technology, and applications. *Communications of the Association for Information Systems, 44*(1), 21.

Willem, J. J. T., & Jakobus, S. v. d. W. (2010). Business intelligence in the cloud. *South African Journal of Information Management, 12*(1), e1-e5.

Witti, H., Guegan, C. G., & Benkhelifa, E. (2018). *A Conceptual Framework of Security Requirements in Multi-Cloud Environment* (Vol. 10975 LNCS) [Conference Paper]: Springer Verlag. Retrieved from https://ezproxy.aut.ac.nz/login?url=https://search.ebscohost.com/login.aspx?direct=true&db=edselc&AN=edselc.2-52.0-85049362947&site=eds-live. https://doi.org/10.1007/978-3-319-94472-2_1

Yeoh, W., Koronios, A., & Gao, J. (2008). Managing the implementation of business intelligence systems:a critical success factors framework. *International Journal of Enterprise Information Systems, 4*(3), 79.

Annexure A List of privacy issues in Cloud BI

	Title of the reviewed paper	Authors	Year of publication	Major themes covered in the paper	Findings
Privacy issues in Cloud BI					
1.	A content-based data masking technique for a built-in framework in Business Intelligence platform	Ali, Osama; Ouda, Abdelkader	2017	The data privacy issues are caused by: 1. Hacking 2. Mistakes caused by business partners 3. Internal users	• The development of the BI system has improved the ability to integrate, store and utilize sensitive operation data. However, this has led to an increase in both internal and external data security breaches. These data breaches were caused by hacker attacks or mistakes by business partners. • Research shows that more than 70% of data breaches are caused by internal users (such as curious employees). When many internal users have different roles and responsibilities when accessing BI systems, they pose a serious threat to privacy.
2.	Analytic surveillance: big data business models in the time of privacy awareness	Fernández-Manzano, Eva-Patricia; González-Vasco, María-Isabel	2018	The data privacy issues are caused by 1. Internal users (using customer data	• Many OTT (Over the Top) companies and social networks collect data from their users, which, while making them more competitive, can easily violate

				without permission) 2. Lack of training for users	customer privacy (using customer data without permission) - In this growing data-driven environment, users, especially those who may be defined as technically illiterate, are indeed the most vulnerable actors.
3.	Business intelligence and analytics: from big data to big impact.	Chen, Hsinchun Chiang, Roger HL Storey, Veda C	2012	BI analyst Lack of knowledge of IT security.	- BI&A professionals must be able to understand business problems and develop proper analytical solutions. For BI & A professionals, the required business knowledge includes the fields of accounting, finance, management, marketing, logistics, and operations management, as well as the required domain knowledge to apply to specific BI&A applications, including IT technical knowledge. - BI&A professionals not only need to know how to transform raw data and information (through analysis) into meaningful and actionable knowledge for the organisation but also how to properly interact and communicate with the organisation's business and domain experts.
4.	Privacy Layer for Business Intelligence	Gupta, Vishal Saxena, Ashutosh	2010	The data privacy issues are caused by 1. Internal users sharing data	- According to the survey, more than 50% of companies share their sensitive data for project purposes.

					• Transposition, encryption, and data masking are techniques to strengthen information security, but these techniques are not suitable for retaining data in the original format, especially when the report needs to be refined to a fine-grained situation. • In the process of data analysis, to protect privacy and security, data cannot be copied and stored, • Most algorithms used to protect privacy may not be suitable for data requirements related to data mining or analytical processing requirements. • The author proposes a new method that requires the addition of a new layer to the business intelligence architecture to protect the privacy of sensitive information without changing the data for merge, processing, and policy aggregation.
5.	Considerations Regarding Business Intelligence in the Cloud Context.	MUNTEAN, Mihaela	2015	The data privacy issues are caused by: 1. Fraud 2. Cyber-attacks 3. Security standards 4. External Risk 5. Legal issues	• Definition of Cloud BI • Analysed the strengths and weaknesses of Cloud BI, including data security, auditing, and legal issues. • Introduced and compared differences of Cloud BI vendors in the market, and observed that Microsoft, Oracle, IBM,

					SAP and MicroStrategy meet all the criteria that the author proposed.
6.	Protecting business intelligence and customer privacy while outsourcing data mining tasks.	Qiu, Ling; Li, Yingjiu; Wu, Xintao	2008	The data privacy issues are caused by: 1. Lack of IT expertise and strong computing. infrastructure. 2. The risk of outsourcing data mining tasks.	• Most companies lack IT expertise and strong computing infrastructure, and functional departments may have to delegate or outsource their data mining tasks to the IT department. • The practice of outsourcing data mining tasks involves extensive collaboration across different organisations (for example, exchanging or sharing data). The raw data and information displayed after analysis contain the organisation's business intelligence (BI) and customer privacy. • There are security issues in outsourcing activities, exposing private information to potential risks. Without proper security policies and technologies, security can easily compromise this kind of privacy.
7.	Combining business intelligence with cloud computing to delivery agility in actual economy.	Mircea, Marinela; Ghilic, Bogdan; Stoica, Marian	2011	The data privacy issues are caused by: 1. Data migration 2. Lack of user training	• The main security risks of cloud computing integration are loss of governance, lock-in, isolation failure, compliance risks, management interface compromise, data protection, insecure or incomplete data deletion, malicious insider.

					• The migration of critical applications and infrastructure to the cloud involves more human resources for managing sensitive data and applications. • Users should be trained to use the Cloud BI system. • Cloud computing experts believe that it is impossible to fully migrate BI's solution to the cloud. The migration process needs to be gradual, and both SaaS software and traditional software will coexist within organisations for a long time. • The stages of implementing Cloud BI are 1. Initiation; 2. Analysis and evaluation of the present stage of a business; 3. Designing the Cloud BI solution; 4. Evaluation and selection of a Cloud BI provider; 5. Implementation of the Cloud BI strategy; and 6. Management of Cloud BI solution.
8.	Small and Medium-Sized Enterprises' Understanding of Security Evaluation of Cloud-Based Business Intelligence Systems and Its Challenges	Moyo, Moses Loock, Marianne	2019	• The data privacy issues are caused by: 1. The use of public cloud service. 2. Cyber-attack	• Cloud BI can help SMEs discover new business opportunities and save costs. • Many small and medium-sized enterprises adopt SaaS for Cloud BI implementation because of their low cost and low demand for IT technical resources. However, unlike private clouds, low-cost or free public clouds

				3. Security breach • The use of cloud services by SMEs is a major target of cyber threats. • Implementing BI in public clouds and free clouds carries significant security risks.	face serious security challenges when hosting BIs and sensitive data. • By exploiting security holes in SaaS, hackers can use less complicated methods to bypass security controls, thereby jeopardizing the security of any customer data.
9.	Security and Privacy Issues for Business Intelligence in lOT	Krishna Kagita, Mohan	2019	The data privacy issues in I0T are caused by 1. The plug-in was destroyed by a network attack. 2. Botnet attacks 3. Identity theft, denial of service (DoS) 4. Man-in-the-middle attacks.	• Business intelligence mainly depends on plug-ins that require an Internet connection. If the browser is not protected by firewalls and a security system, the attacker can execute and destroy the plugins associated with it in the system's browser. • Devices that lack security will allow attackers to track the details of the system and perform attacks or steal data. • Several types of IoT attacks in business intelligence-botnet attacks, identity theft, denial of service, and man-in-the-middle attacks. • In the case of a cloud attack, the data stored by the organisation on the server for data analysis may be affected.

10.	Information Security in Business Intelligence based on Cloud: A Survey of Key Issues and the Premises of a Proposal	Jaramillo, Elena Munier, Manuel Aniorte, Philippe	2013	• The risks in a distributed environment are the same as traditional risks. • The data privacy issues are caused by • Data sharing • The log file data	• Jaramillo et al. (2013) found that while data from different users may be stored on the same cloud provider, cloud users are concerned about the risk of their data being manipulated by other users on the cloud. • The traditional risks in distributed environments also bring usability, integrity, and confidentiality. • The author points out that the trustworthiness (of data and partners) is to monitor the operation of each partner and monitor workflow or data flow itself (the centre with the data). However, there is a lot of workflow and data information in the logs recorded by the monitoring activities, and these records have privacy problems. Once this information is disclosed, the company will be exposed to the risk of network attacks.
11.	Risk analysis of business intelligence in cloud computing	Alsufyani, Raed Chang, Victor	2015	The difference between the security of Cloud BI and cloud security. • Security cost increase The data privacy issues are caused by	• Because cloud-based technology has so much flexibility, it is less private because it follows the principle of "everyone accesses everyone." • Because cloud-based technologies require multiple layers of protection and always include backup cloud storage, the risk of data damage is

				- Lack of training - Loss of data control - External risk	lower. However, effective protection will require more budgetary costs. - Lack of training for employees in cloud technology makes data insecure. - Because cloud technology is often outsourced, it increases reliance on external third parties, which increases the risk of losing control.
12.	Sharing-based Privacy and Availability of Cloud Data Warehouses	Varunya Attasena, Nouria Harbi, Jérôme Darmont	2006	In the context of Cloud BI, privacy is critical.	- Nowadays, cloud service providers have addressed most security issues, but as the number of cloud service providers and subcontractors in different countries increases, complex legal issues arise.
13.	Cloud Business Intelligence – Is What Business Need Today	Yuvraj Singh Gurjar, Vijay Singh Rathore	2013	Privacy issues in Cloud BI.	- Data is important to many organisations because it contains the core data of the organisation. There are several ways to protect data in the cloud today, both during storage and at rest, including standard security measures, such as encryption keys, SSL, and certificates. - Integration of local data with cloud components is a challenge because it still exists in silos and requires access data passing through firewalls. - The reliability of cloud service: there may be risks of data security and lack of data control. Service level agreements are difficult to obtain

					(SLA) from the cloud provider. Therefore, the organization should already have comprehensive IT governance and service delivery standards and models.
14.	Cloud Computing and Business Intelligence	Alexandru Adrian TOLE	2014	The security and privacy issues in Cloud BI and Cloud Computing.	• Cloud BI applications should be designed not only to consider process design but also to set up data processing that ensures privacy. • In cloud computing is difficult to ensure the security of cloud users, who are responsible for the security of the software they implement in the cloud environment. If a user's account password is hacked and logged in illegally, the security of the entire cloud infrastructure is compromised, and many clients' data may be at risk. • To protect the user attack from the provider, the provider allows the user to encrypt the data to protect sensitive data, which may slow the response time of the data load. • Implementing a security solution may impose more costs on local hardware implementations. • Regarding privacy, BI solutions supply security only at the UI (user interface) level, and the data stored in

					the cloud database is exposed to the provider.
How data privacy issues in Cloud Business Intelligence are different from the data privacy issues in general cloud services.					
15.	Moving business intelligence to cloud environments	Juan-Verdejo, Adrián Surajbali, Bholanathsingh Baars, Henning Kemper, Hans-Georg	2014	The Cloud BI security discussions fail to recognise the Cloud BI layers and security requirements at various layers.	The authors suggest that in a Cloud BI there are the following layers: 1. Data layer 2. Logic layer 3. BI Access layer (presentation layer) In a Cloud BI, all these layers can be distributed across the different cloud service providers and this may pose security challenges.
16.	Business intelligence security on the clouds: Challenges, solutions, and future directions	Al-Aqrabi, Hussain Liu, Lu Hill, Richard Ding, Zhijun Antonopoulos, Nick	2013	The difference between the security of Cloud BI and cloud security. • Cloud BI security system should especially include the metadata of the integrated data. • Security cost • Consider not only technical safety but also the safety of	• In the cloud computing sharing space, BI is facing increasing security and privacy threats because unauthorised users may have access to extremely sensitive and unified business information. • The business process contains collaborative services and users from multiple cloud systems in different security domains that require dynamic participation at run time. If heterogeneous cloud systems in different security domains do not have direct authentication relationships, it is technically difficult to achieve security collaboration. • In Cloud BI, the security controls are needed at all the layers including

				the personnel involved. • Uses of Mix models (UTM and DTM) for security control of Cloud BI. • OLAP security.	session, presentation, application, network, and transport. This affects the costs of security for a Cloud BI solution. Thus, the security of Cloud BI differs from that of the cloud for security costs. • Security considerations in Cloud BI may transcend beyond technical security but also to people who play a significant role in the management of business rules. • The security control of Cloud BI should apply both Distributed threat management (DTM) approach and Unified Threat Management (UTM) Approach. • Security in the cloud may only focus on front-end application (browser) security and database security, but Cloud BI security includes OLAP application server security hosting the OLAP dashboard.
17.	A conceptual framework of security requirements in a multi-cloud environment	Hamad Witti, Chirine Ghedira Guegan, and Elhadj Benkhelifa	2018	The security requirements in a multi-cloud environment, which may differ in Cloud BI security requirements.	Multi-cloud Security Challenges includes: • There is a lack of interoperability between vendors due to their differences in SLA and policy. • Data transmission is not secure between different clouds and cloud providers.

					• Protecting privacy and security exchanges, policy and conflict management between the clouds is an important challenge. • Authors have proposed 10 security requirements in a multi-cloud environment: Availability, Access Control, Attack/Harm Detection and Prevention, Integrity, Accountability, Privacy, Binding of duties, Separation of duties and Delegation.
18.	Business intelligence in the cloud	Shimaa Ouf, Mona Nasr	2011	The security requirement on Cloud BI may be different from Cloud Computing.	• The Cloud affects the SaaS vendors: The new requirements will be applied on APIs, reporting, security, and service level agreements (SLAs).
19.	Cloud BI: a multi-party authentication framework for securing business intelligence on the Cloud	Al-Aqrabi, Hussain	2016	The difference between the security of Cloud BI and cloud security • The BI framework will require access control at various BI processes and all authentication layers. • Added safety verification	• Implementation of BI may face technical controls challenge at various stages of the BI process. BI framework will need access controls at various BI process and all layers of authentication. • Because it has a special security framework and contains a multi-dimensional security architecture in database objects, a large number of interactions may be generated in the BI framework during security control and data encryption at various levels. This makes the BI framework heavier in terms of computer overload and

				overload, BI security controls are different from traditional database applications.	capacity. Besides, as security verification increases, BI frameworks may also overload the network. • BI security controls are different from a traditional database application.

Annexure B Evaluate the quality of reviewed papers

Code Number	Author (year)	Citation count
Privacy issues in BI		
1	Ali and Ouda (2017)	2373
2	Fernández-Manzano and González-Vasco (2018)	0
3	Chen et al. (2012)	5341
4	Gupta and Saxena (2010)	1
5	Muntean (2015)	14
6	Qiu et al. (2008)	73

7	Mircea et al. (2011)	108
8	Moyo and Loock (2019)	0
9	Krishna Kagita (2019)	3
10	Elena Jaramillo et al. (2013)	10
11	Alsufyani and Chang (2015)	8
12	Attasena et al. (2006)	10
13	Gurjar and Rathore (2013)	58
14	Tole (2014)	13

15	Juan-Verdejo et al. (2014)	9
16	Al-Aqrabi et al. (2013)	15
17	Witti et al. (2018)	1
18	Ouf and Nasr (2011)	46
19	Al-Aqrabi (2016)	1

Annexure C Description of various types of delivery models

Name of deployment models	Description (Clemons & Chen, 2011)
Public Cloud	The cloud infrastructure is provisioned for open use by the public. It may be owned, managed, and operated by a business, academic, or government organisation, or some combination of them. It exists on the premises of the cloud provider.
Private Cloud	The cloud infrastructure is provisioned for exclusive use by a single organisation comprising multiple consumers (e.g., business units). It may be owned, managed, and operated by the organisation, a third party, or some combination of them
Community Cloud	The cloud infrastructure is provisioned for exclusive use by a specific community of consumers from organisations that have shared concerns (e.g., mission, security requirements, policy, and compliance considerations). It may be owned, managed, and operated by one or more of the organisations in the community, a third party, or some combination of them.
Hybrid Cloud	The cloud infrastructure is a composition of two or more distinct cloud infrastructures (private, community, or public) that remain unique entities, but are bound together by standardised or proprietary technology that enables data and application portability (e.g., cloud bursting for load-balancing between clouds).

Name of service delivery models	Description
Software as a Service (SaaS)	The capability provided to the consumer is to use the provider’s applications running on a cloud infrastructure.
Platform as Service (PaaS)	In this model, a layer of software or development environment is encapsulated and offered as a service, upon which other higher levels of service are built. The customers have the freedom to build their applications, which run on the provider's infrastructure. Although the customers do not manage or control the underlying cloud infrastructure, network, servers, operating systems, or storage, they have control over the deployed applications.
Infrastructure as a Service (IaaS)	It also referred to as Resource Clouds supply resources which are managed and can easily be scaled up, as services to a variety of users. They essentially deliver basic storage and compute capabilities as standardised services over the network. Servers, storage systems, switches, routers, and other systems are pooled and made available to handle workloads that range from application components to high-performance computing applications.