Otto Buure

CHALLENGES IN MOVING TO CLOUD COMPUTING ENVIRONMENT: CASE FINNISH TELEOPERATOR

UNIVERSITY OF JYVÄSKYLÄ
DEPARTMENT OF COMPUTER SCIENCE AND INFORMATION SYSTEMS
2020

ABSTRACT

Buure, Otto
Challenges in Moving to Cloud Computing Environment: Case Finnish Teleoperator
Jyväskylä: University of Jyväskylä, 2020, 67pp.
Information Systems Science, Master’s Thesis
Supervisor: Semenov, Alexander

Cloud based services are extremely popular among organization today. Cloud brings many benefits and opportunities for companies, but it also brings uncertainty and challenges of data security and privacy. The privacy of personal data has become more precise and regulations and legislations like EU General Data Protection Regulation (GDPR) has intervened how companies must process customers personal data so that the privacy remains. This Master’s Thesis explores what changes moving to a cloud computing environment causes compared to a traditional information system. The research focuses on the change of control cloud causes and how organizations can preserve the control in the cloud. This research also aims to clarify the goals of the GDPR and what it means to a companies that are using or intending to adopt a cloud. Cloud differs from a traditional on-premise information systems (IS) in many ways, but the existing practical security mechanisms can be utilized to ensure security and privacy in the cloud if organizations know what they are doing. The amount of control over the system decreases when moving to a cloud but this can be mitigated by contracts and agreements and proper security mechanisms. The official guidelines organizations get need to be updated to cover the tangible actions organizations need to take to ensure that following the regulations does not become too complex. Cloud is open to the internet and it requires a new kind of thinking when it comes to security. As a precaution, organizations need to invest in improving the general awareness of cloud computing among the employees that will simplify the designing of the security mechanisms that are utilized with the cloud. The awareness among organization can mitigate the security and privacy risk of sensitive data being stored and processed in cloud service or systems with insufficient security level.

Keywords: cloud, cloud computing, cloud environment, computer security, cloud security, GDPR

TIIVISTELMÄ

Buure, Otto
Challenges in Moving to Cloud Computing Environment: Case Finnish Teleoperator
Jyväskylä: Jyväskylän yliopisto, 2020, 67s.
Tietojärjestelmätiede, pro gradu -tutkielma
Ohjaaja: Semenov, Alexander

Pilvipohjaisista palveluista on tullut erittäin suosittuja yritysten kesken. Pilvi tuo mukanaan monia hyötyjä ja mahdollisuuksia yrityksille, mutta se tuo myös epävarmuutta ja haasteita datan turvallisuuteen ja yksityisyyteen. Henkilötietojen yksityisyydestä on tullut tarkempaa ja määräykset ja lainsäädännöt kuten EU:n yleinen tietosuoja-asetus (eng. EU General Data Protection Regulation, GDPR) ovat puuttuneet siihen miten yritysten täytyy käsitellä asiakkaidensa henkilötietoja siten, että yksityisyys säilyy. Tämä pro gradu -tutkielma tutkii mitä muutoksia pilviympäristöön siirtyminen aiheuttaa verrattuna perinteisiin tietojärjestelmiin. Tutkimus keskittyy kontrollin muutokseen, jonka pilvi aiheuttaa ja kuinka organisaatiot voivat säilyttää kontrollia pilvessä. Tämä tutkimus myös pyrkii selventämään EU:n yleisen tietosuoja-asetuksen tavoitteita ja mitä ne tarkoittavat yrityksille, jotka käyttävät tai aikovat ottaa pilven käyttöön. Pilvi eroaa perinteisistä tietojärjestelmistä monin tavoin, mutta jo olemassa olevia käytännöllisiä tietoturvamekanismeja voidaan hyödyntää tietoturvan ja yksityisyyden turvaamiseen pilvessä jos organisaatiot tietävät mitä ovat tekemässä. Kontrollin määrä järjestelmiin vähenee, kun siirrytään pilveen, mutta kontrollin vähenemistä voidaan pitää kurissa yritysten välisillä sopimuksilla ja oikeanlaisilla tietoturvamekanismeilla. Viranomaisten ohjeistukset yrityksille tarvitsevat päivitystä, jotta ne kattaisivat myös tarvittavat toimet, joita organisaatioiden tulee tehdä, jotta varmistutaan siitä ettei lainsäädännön noudattamisesta tule liian monimutkaista. Pilvi on avoinna internetiin, joten sen tietoturva vaatii uudenlaista ajattelua. Varotoimena organisaatioiden tulee panostaa yleiseen tietoisuuteen pilviympäristöistä kaikille työntekijöille, joka voi yksinkertaistaa pilvessä hyödynnettävien tietoturvamekanismien suunnittelua. Tietoisuus organisaation sisällä voi myös pienentää tietoturvan ja yksityisyyden riskiä, jossa arkaluonteista dataa tallennetaan ja käsitellään pilvipalvelussa tai pilvijärjestelmässä, jonka tietoturva ei ole riittävällä tasolla.

Asiasanat: pilvi, pilvilaskenta, pilviympäristö, tietoturva, pilvitietoturva, GDPR

FIGURES

TABLES

TABLE OF CONTENT

1 INTRODUCTION

Utilizing cloud computing has become extremely common among organization. Computing environment has changed dramatically in last decade and now computing is seen as an utility (Buyya, Yeo, Venugopal, Broberg, & Brandic, 2009; Varghese & Buyya, 2018). The changes in the business field require changes in the systems and infrastructure of many organizations (Coppolino, D'Antonio, Mazzeo & Romano, 2017). The key feature that the cloud computing brings forth is that the consumers do not need to acquire computing infrastructure or resources, but they can acquire them as a service with less cost (Singh & Chatterjee, 2017). Cloud is seen as a new norm for many functionalities and business processes. Clouds can create many benefits and opportunities for organizations, but it also brings forth uncertainty and challenges in data security and privacy. The privacy of personal data has become more precise and regulations and legislations like EU General Data Protection Regulation (GDPR) has intervened how companies must process customers personal data so that the privacy remains.

The concept of security and privacy in cloud environments is similar to a traditional concept of security and privacy in any traditional information systems (Chen & Zhao, 2012). Mogull, Arlen, Gilbert, Lane, Mortman, Peterson and Rothman (2017) stated that traditional security domains of normal information systems remains in the cloud, but there is a dramatical change in the nature of risks, roles, responsibilities and implementation. This leads to the view that maintaining the security in cloud environment is shared between the actors just like any other features in cloud (Mogull et al., 2017). Data security and privacy are always closely related. Privacy in cloud environments is more complicated than privacy in traditional information systems. Information in cloud environments is normally shared between geologically decentralized data centers which makes the physical location of the data complicated. Privacy issues are also the greatest factor which has slowed down the cloud adaptation. The development of cloud environments and the expectations of its potential benefits of cloud computing have caused businesses and organizations to see it in more positive light (Soares, Gonçalves, Parreira, Tavares, Carapinha, Barraca, Aguiar & Sargento, 2015). Soares et al. (2015) notes telco sector being one of the most active fields to explore the possibilities cloud environments have to offer. Although the seemly limitless benefits of cloud adoption, there are still many issues the organizations face when considering moving some functions to the cloud. Sensitive data is part of many business processes today, which raises the concern of security. Although cloud computing could speed some business processes up there are still doubt about their security when it comes to processing and storing sensitive Personally identifiable information (PII).

This Masters' Thesis explores what precautions and actions organizations need to take before moving to a cloud computing environment especially when the cloud is provided by a third party. This study aims to combine the answers from the literature with the results from empirical research to create a theory or guidelines for organizations that are intending to move their business processes or systems to operate in cloud. To answers this a research question was defined:

- What are the needed actions and precautions an organization must take when storing and processing personally identifiable information in a cloud computing environment provided by a third party?

To help to define the research problem and to get more profound understanding of this multidimensional problem two focusing questions were defined:

- How does cloud as an environment differ from traditional IS in control and responsibility and how can organizations preserve the control in cloud environment?
- What are the needed actions to ensure privacy and security in cloud computing environment?

Cloud related projects are current for many organizations. Although there has been a lot of research related to the cloud, there are still many issues and unsolved challenges that need to be considered before the decision to transition to cloud. These challenges are commonly related to security, privacy and the regulations around the cloud. These challenges require more profound reviewing. Cloud development has been extremely fast and more an more organizations are adapting it to their normal processes. The need for the research in this topic was identified when the GDPR became active. GDPR regulates how personal data must be controlled and processed by organization and it renders rights for the personal data back to the data subjects. GDPR does not only regulate the cloud but with all the challenges with security and privacy in the cloud, it became clear that there are a lot of things to consider. Because the GDPR is so recent there are not much literature where it is reviewed with the cloud. Also, the case company's interests is to maintain a high level of security and privacy that they have in their traditional information systems also in the cloud environments. Combining the interests of the case company and the amount of earlier research of the research topic, the research topic became very attractive and interesting to review more profoundly.

The the literature review of the research adapted Okoli & Schabram's (2010) methodology for conducting a systematic literature review. The literature that was used conducting the theory for the study was searched using three academic online libraries: Google Scholar, IEEE Explore and AIS electric library. Search words that were used and combined to search relevant literature were: cloud, cloud computing, cloud environment, service models, deployment models, computer security, cloud security, data security, privacy, GDPR, GDPR sanctions, GPDR compliance.

The empirical research of this study was conducted by qualitative methods. The research data was collected using semi-structured interviews (Hirsjärvi & Hurme, 2014). This research was carried out as a commission for a teleoperator operating in Finland. The interviewees that participated in this research were employees of the case company that work closely with cloud and cloud related topics. After the interviews were conducted they were transcribed to a text verbatim. After this the data was coded to three themes transitioning to cloud, the change in control and the needed tangible actions. The data was then analyzed by using qualitative methods.

After the introduction of the study, the literature review is presented in chapters 2-5. Chapter 2 defines the cloud computing as a term and technology. Chapter 3 reviews the security and privacy in cloud computing environments. Chapter 4 clarifies the goals of the GDPR. After the literature review, the empirical research is presented in chapters 5-7. Chapter 5 presents the research methodology. Chapter 6 presents the results of the study. Chapter 7 presents the discussion, which addresses the theoretical contributions of the study, limitations of the study and suggestions for interesting topics for the future researches. The final chapter, chapter 8, is the conclusion.

2 OVERVIEW TO CLOUD COMPUTING

This chapter gives background for cloud computing and how it has become one of the dominating technologies in whole IT field. In defining cloud computing as a term or a model, the most cited definitions from literature are presented and discussed. Origin of the cloud computing as model is also defined. In more detail this chapter focuses in cloud environment as a definition, gives slight background for cloud computing, cloud computing characteristic, cloud computing deployment models and cloud computing service models.

2.1 Cloud computing as a definition

But what is cloud computing? Cloud computing is a way to dynamically increase or decrease capacity and resources without the need to invest in new infrastructure, personnel or software licenses (Subashini & Kavitha, 2010). Cloud computing is term used in concept of referring both the shared software and shared hardware that can be found in cloud computing environments. Shared software means applications that are delivered as service through the internet and hardware means all the infrastructure (systems software, storage servers, compute servers and such) that are placed in a data centers which creates the groundings for cloud computing and cloud environments (Armbrust, Fox, Griffith, Joseph, Katz, Konwinski, Lee, Patterson, Rabkin, Stoica and Zaharia, 2010). Armbrust et al. (2010) stated that cloud computing can be seen as a combination of software as a service (SaaS) and utility computing without including small and medium data centers. This statement can be seen as definition only for a part of cloud computing, because it leaves out the other service models. While cloud computing means the activity that is happening in the cloud, cloud environment can be seen as the whole ecosystem where this happens and as factors that make it possible, including components from infrastructure all the way to user interface. Cloud computing is a way or a business model to reach computation resources without the need for upfront IT investment (Al Morsy, Grundy,

& Müller, 2016). Cloud computing creates a new channel for products and services combining technical and innovative opportunities with pricing models (Ramachandra, Iftikhar & Khan, 2017). There is a variety of definitions for cloud computing available in academic literature. Most of the definitions that can be found in literature define cloud computing as a flexible and economical way to share computing resources on demand and through internet. Ruan, Carthy, Kechadi, & Baggili (2013) conducted a survey on cloud forensic definitions where they also surveyed the definition of cloud computing. As a result of the survey on National Institute of Science and Technology's (NIST), Gartner's and Cloud Security Alliance's (CSA) definitions were as following: "83% of the respondents agree or strongly agree with the NIST definition of cloud computing version 15 and the Gartner definition. 68% of the respondents agree or strongly agree with the CSA definition". From these results Ruan et al. (2013) we were able to draw a conclusion that the cloud computing definition by the leading international organizations is strongly agreed. These three most commonly cited definitions can be found below in the table 1.

TABLE 1 Cloud computing definitions (Ruan et al. 2013)

AUTHOR (S)	CLOUD COMPUTING DEFINITION
NIST (2011) (Mell & Grance)	"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"
Gartner (2009)	"Gartner defines cloud computing as a style of computing in which scalable and elastic IT-enabled capabilities are delivered as a service to external customers using Internet technologies."
CSA (2011)	"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). Cloud computing is a disruptive technology that has the potential to enhance collaboration, agility, scaling, and availability, and provides the opportunities for cost reduction through optimized and efficient computing. The cloud model envisages a world where components can be rapidly orchestrated, provisioned, implemented and decommissioned, and scaled up or down to provide an on-demand utility-like model of allocation and consumption."

Ruan et al. (2013) also found out from the results of the survey that cloud computing is not believed to be neither entirely new technology nor a mere combination of already existing technologies. Delivery of the computing resources is something new and it can be seen as a consistent evolutionary step of IT evolution (Ruan et al., 2013).

2.2 Cloud environment

Cloud environment can be defined as a combination of deployment models, service models and exhibiting characteristics of cloud computing. Deployment models contain the infrastructure and its deployment which varies in different models. Service model layer explain how and what kind of cloud services are provided. According to Subashini and Kavitha (2010) cloud computing service models are the core of the cloud. The layer above the service models contains the cloud computing characteristics which exhibit in service models (Subashini & Kavitha, 2010). The cloud computing characteristics were defined a little diffrently by Subashini and Kavitha (2010) and by Mell and Grance (2011) which we go throuh more accurately in chapter 2.3. The cloud computing characteristics were decided to included in this cloud environment definiton are on-demand self-service, broad network access, multi-tenancy, rapid elasticity, measured service and resource pooling. The model of layers that cloud environment consists of and their components are shown in figure 1 (Mell & Grance, 2011; Subashini & Kavitha, 2010).

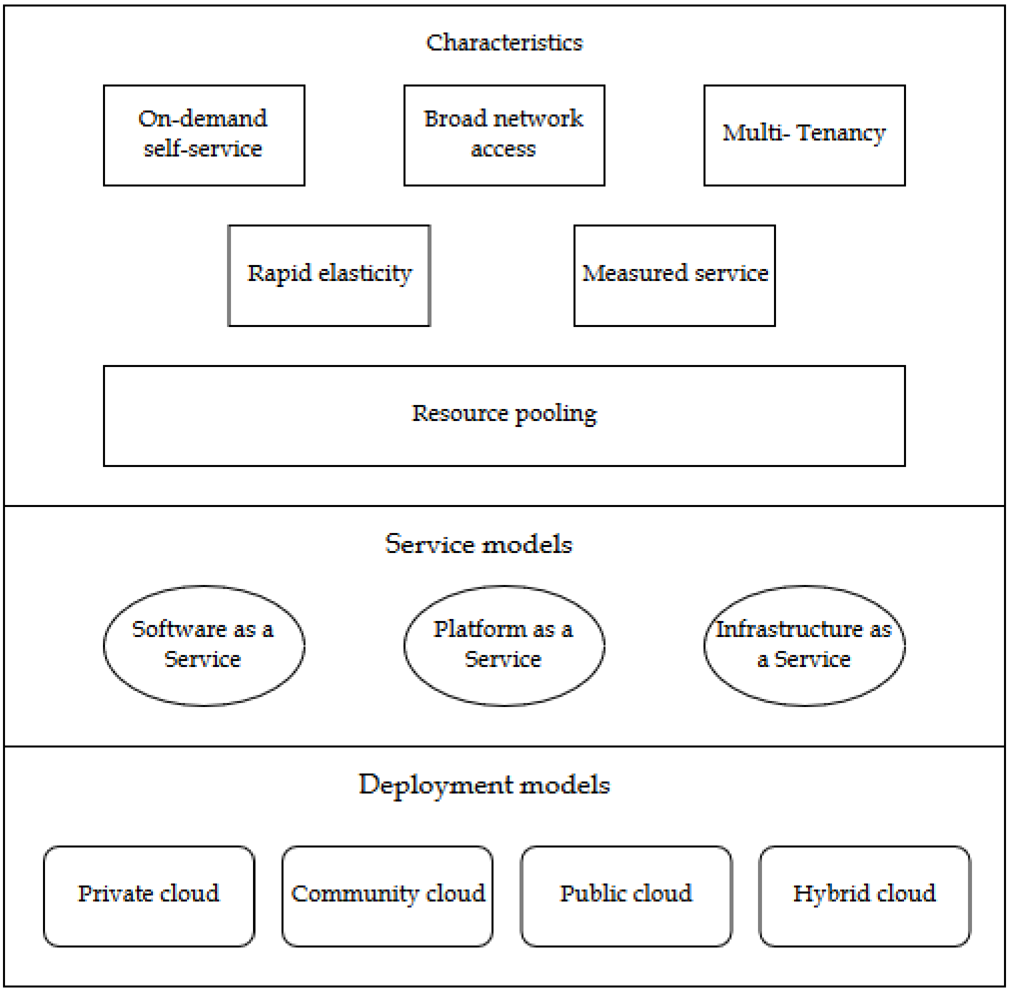


FIGURE 1 Cloud environment (Mell & Grance, 2011; Subashini & Kavitha, 2010)

2.3 Background for cloud computing technology

When defining the evolution of cloud computing Mather, Kumaraswamy and Latif (2009) cited *The Big Switch* (2009) where Nicholas Carr gives a great example of what kind of an effect cloud computing might have on IT. Carr (2009) argues that cloud computing will have similar effect on IT than electrification had in industrial age. Before electrification industrial companies had to produce the needed power by them self, but electrification changed that to just plugging in to the electrical grid. Carr (2009) saw similar change in IT with cloud computing as electrification. Earlier companies had to produce their own computation resources, but after cloud technology emerged, computation resources became also available through network by plugging in the network cable (Mather et al., 2008). Cloud computing has many similarities with grid computing. Cloud computing and grid computing both have the same vision to reduce the operation costs of computing and increase the flexibility and reliability by using shared hardware, through a network, often operated by a third party (Vaquero, Rodero-Merino, Caceres, & Lindner, 2008). Foster, Zhao, Raicu, & Lu (2008) compared grid computing with cloud computing and tried to clarify their differences. They stated that the idea behind cloud computing is not completely new. John McCarthy's prediction from 1961: "*computation may someday be organized as a public utility*" (Foster et al., 2008) is now a day quite close to what cloud computing consists of. It can be said that cloud computing is a result of the development of grid computing but is not entirely the same thing with newer technology. The grid computing as a term is from the mid-1990s, which meant the process or technology of obtaining computing power on demand. Foster et al. (2008) stated that cloud computing is not just a new name for grid computing but there are many similarities with these two concepts. Grid computing could be seen as an equivalent term for cloud computing today, but for the 1990s technology. Vaquero et al. (2008) stated that high state of virtualization and focus in usability of the Clouds are the key differences with these two computation paradigms. They also noted that there are many overlapping technologies and designs but as Foster et al. (2008) noted that Clouds and Grids are similar technologies for similar purposes but from different decades.

2.4 Cloud computing characteristics

NIST definition of cloud computing Mell and Grance (2011) define five essential characteristics for cloud computing:

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

These essential characteristics defines that information system is a cloud. If a system misses any of these characteristics it most likely is something else than a cloud (Mogull et al., 2017). On-demand self-service includes automation for distribution of computer resource capabilities customer needs without requiring communication with service provider (Mell & Grance, 2011). Broad network access means service availability for various devices with different platforms through internet. Resource pooling means that the service providers computing resources are pooled together to cover the needs of multiple customers at ones by dynamically sharing the physical and virtual resources according to customers' needs and demand (Mell & Grance, 2011). Rapid elasticity means often automated provision of capabilities quickly scaling up or down with demand.

This might create an impression for the customer of infinite capabilities that can be accessed any time. Measured service includes the measuring of needed and used resources. By measuring the resource usage, the transparency of the service gets higher for both of the parties, user and provider (Mell & Grance, 2011). Subashini and Kavitha (2010) also defined cloud computing characteristics which are similar to what can be found in the NIST definition of cloud computing. On-demand self-service, resource pooling, rapid elasticity and measured service can be found from both definitions. Also we can assume that broad network access by Mell and grance (2011) means the same concept as Subashini and Kavithas' (2010) defined ubiquitous network. Subashini and Kavitha (2010) also included multi-tenancy as a cloud computing characteristic which is not included in the NIST definition for cloud computing.

2.5 Cloud computing deployment models

Mell and Grance (2011) divided cloud computing deployment models in four different categories, which can be seen as different kind of cloud computing environments. These deployment models create a founding for service models to function. Deployment models of cloud computing that Mell and Grance (2011) defined are listed and explained as follows:

- Private cloud
- Community cloud
- Public cloud
- Hybrid cloud

Private cloud means a cloud computing environment, which is in private use of a single organization. Even though the cloud is provisioned for a single organization and is sometimes maintained and executed internally, the execution and maintenance of a private cloud can also be outsourced to a third party or be a combination of internal and external responsibility. (Mell & Grance, 2011). Basically private cloud is used in defining the internal datacenters of a company that are not publicly shared (Dawoud, Takouna, & Meinel, 2010). In private clouds the security level is easier to guarantee when compared to other cloud computing deployment models, but the economic cost with private clouds are higher (Pearson & Benameur, 2010) due to a lack of parties that are dynamically sharing resources.

Community cloud means a cloud environment that is meant for a use of certain community of people or organizations with shared regulations, policies or concerns about security issues. Community cloud can be owned, managed or operated by some organization or organizations inside the community, it can be purely operated by external third party or it can be some sort of a combination of these both (Mell & Grance, 2011).

Public cloud is a cloud environment, which is provisioned for public to utilize as utility computing (Dawoud et al., 2010). The owner, manager or operator of the public cloud can be almost any organization such as commercial enterprise and academic or governmental organization. Public cloud exists completely inside its providers facilities. (Mell & Grance, 2011). But what makes a cloud a public cloud is when it is made for anyone to utilize by pay-as-you-go manner (Armbrust et al., 2010). According to Pearson and Benameur (2010) public cloud is the most effective deployment model when considering cost reduction that is achieved by centralization of services.

Hybrid cloud means a cloud environment, which is some sort of a combinations of two or more earlier mentioned cloud environments or their unique infrastructure models (Mell & Grance, 2011). Although hybrid cloud might seem complex, which they more often are, hybrid cloud is said to be able to combine benefits of the public cloud such as efficiency with private clouds security controls (Linthicum, 2016).

2.6 Cloud computing service models

Cloud computing service models are the core of the cloud and they create a founding for cloud computing characteristics to operate (Subashini & Kavitha, 2010). In the NIST definition for cloud computing Mell & Grance (2011) provided three representational service models for cloud computing. These service models are Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). Cloud computing can be seen as a stack of layers where SaaS is built on top of PaaS which is operating on top of IaaS as seen as in figure 2. This definition of cloud environment does not include the major part of cloud deployments, but it clarifies the layer of architecture where service models operate (Mogull et al., 2017).

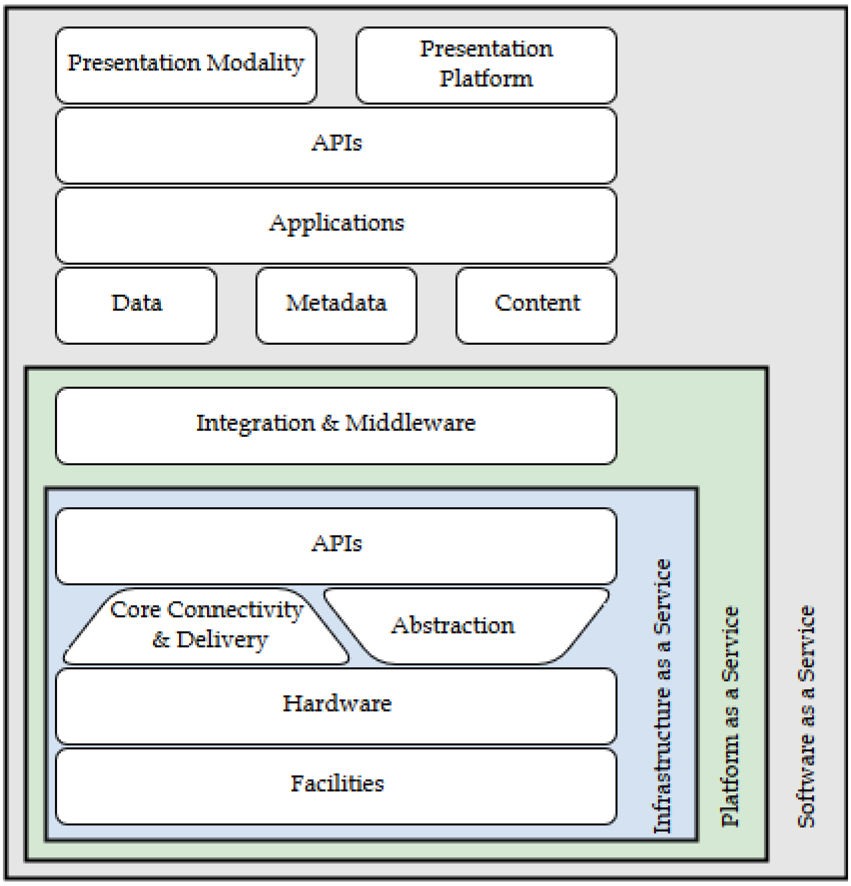


FIGURE 2 Service model layers (Mogul et al., 2017)

2.6.1 Infrastructure as a Service

IaaS is a way to abstract the needed physical infrastructure and infrastructure hardware through virtualization(Mell & Grance, 2011; Mogull et al., 2017). Mogull et al. (2017) defined that "IaaS consists of a facility, hardware, an abstraction layer, an orchestration (core connectivity and delivery) layer to tie together the abstracted resources, and APIs to remotely manage the resources and deliver them to consumers" which is presented in FIGURE 2 above. In IaaS the users buy abstracted and pooled together resources such as servers, storage, networks, processing and other essential computation resources from a service provider (Mell & Grance, 2011; Mogull et al., 2017). Through IaaS these resources can be hastily and accurately managed and scaled up or down to reach the optimal resource usage. In practice IaaS works as follows. In IaaS Physical servers execute two separate components at the same time: a hypervisor that enables virtualization and management software that controls the servers and connects them to controller of computing resources (Mogull et al., 2017). When the customer requests a certain sized virtual server, cloud controller determines which server has the ideal capacity for customers request. After finding a suitable server for customer the cloud controller creates a virtual hard drive for the requested capacity from storage controller, which is in charge of allocating storage resources, and then connects it to suitable host server via a network. (Mogull et al., 2017). Networking is also being allocated in this process. After this cloud controller send a server image copy to virtual machine and manages its configuration. This process creates a working virtual machine, virtual hard drive and virtual network which is ready to be used. After completing this process the metadata and connectivity information is sent to the customer by cloud controller when customer can log in and utilize the IaaS (Mogull et al., 2017). From customer point of a view IaaS has completely changed the deployment of applications by enabling the abstraction of hardware and people needed to run and maintain them (Subashini & Kavitha, 2010). One example of a popular IaaS product is Google Compute Engine.

2.6.2 Platform as a Service

Mogull et al. (2017) noted that PaaS is harder to accurately define or characterize than SaaS or IaaS due to its' many various implementation methods. PaaS is a cloud platform for applications and software where systems run on (Vaquero et al., 2008). Software developers are able to run various applications in various languages without worrying about underlying infrastructure or resources, which release time to focus in development itself (Mogull et al., 2017). These applications are created with programming languages, libraries, services and tools most often provided and supported by the service provider, but it does not automatically exclude other compatible methods that might not be directly supported by service provider (Mell & Grance, 2011). Compared to IaaS, PaaS adds an additional layer on top of IaaS which consists of integration with middleware capabilities, application development frameworks, and messaging, queuing, databases and such functions (Mogull et al., 2017). PaaS can be built directly on top of IaaS, like in FIGURE 2 where the integration and middleware layers are added on top of IaaS layers. In this case integration and middleware layer and IaaS layers are pooled together and exposed to customer using Application programming interfaces (API) as PaaS (Mogull et al., 2017). When utilizing PaaS, the cloud users do not see the infrastructure behind it. In cloud users interface only the platform is visible and cloud controller takes care of managing networking, servers, patches, etc. (Mogull et al., 2017) which simplifies the user interface of the cloud. Because the various implementations of PaaS, it does not require it to be built on top of IaaS. For example, PaaS can be customized like a stand-alone architecture as well. The most important definition for PaaS is that users can access the platform without accessing the underlying infrastructure (Mogull et al., 2017). One example of widely known and utilized PaaS service is Heroku platform by Salesforce.com.

2.6.3 Software as a Service

Services that are categorized in SaaS are normally multitenant applications that have complex underlying architecture like other larger software platforms. Like shown in FIGURE 2, many SaaS products are built on top of PaaS and IaaS or a combination of them to increase their resilience and other features (Mogull et al., 2017). SaaS is the most utilized cloud computing service model. It includes many everyday applications consumers use on computer maybe even without realizing it is a SaaS product or connected to a cloud. One example of widely popular SaaS product is Microsoft's Office 365. SaaS can be seen as the model for software deployment or business model for software where consumers buy license for application that is provided by the service provider without the need to buy the software itself (Mell & Grance, 2011; Safonov, 2016). In SaaS model consumers can use the applications with various devices through the internet or as Safonov (2016) defined they can: *"access to commercial software via the network"*. In SaaS model users do not or cannot normally control the cloud infrastructure and its components (Mell & Grance, 2011). Many SaaS applications utilize APIs for functionalities. APIs are needed to support different kind of clients where SaaS products run like mobile applications and web browsers. APIs are normally placed on top of application/logic layer and data storage (Mogull et al., 2017).

3 Privacy and security in cloud environments

This chapter gives background for security and privacy in cloud computing environments. Security in cloud computing environments is also examined in more detail in different cloud computing service models and service level agreements are defined and their purpose and significance are examined. There is also preview into privacy as a definition and how it exhibits in cloud computing environments. After this cloud forensics and logging is examined.

3.1 Security

Enforced security guarantees that have been assessed are an increasing priority for cloud users and data owners for the wide adoption of cloud. These security guarantees include data integrity, data confidentiality, access control and availability (Samarati, di Vimercati, Murugesan & Bojanova, 2016). According to Chow, Golle, Jakobsson, Shi, Staddon, Masuoka and Molina (2009) most of the concerns of privacy and security in cloud environments are not completely new problems at all. They picture the problems with regulations and trust issues as a same kind of problems organizations faced with offshoring and outsourcing. Security is in a significant role in the foundation of sense and trust between the cloud consumer and cloud provider (Arora, Khanna, Rastogi & Agarwal, 2017). It is fundamentally important for the cloud provider to mitigate all kind of security risks that may affect the user's data when all of it is managed and stored in the cloud (Arora, Khanna, Rastogi & Agarwal, 2017). In the earlier states of cloud computing organizations were already utilizing some cloud-based services, but because of the uncertainty of the cloud security, the consumers would not store their most sensitive data in the cloud (Chow et al., 2009). But now when cloud computing has spread wider and more and more business transactions are being done in the cloud, organizations in certain situations need to store and process sensitive data in the cloud. Some applications that are obligatory for organizations business processes might be executed entirely in the cloud such as certain SaaS applications. Thus, the possibility to store and process sensitive data in cloud is a mandatory for some organizations and it requires cloud providers to maintain and develop their security to keep the cloud security in high level. Chow et al. (2009) stated that already in 2009 many of the security problems that clouds face have already been there before the adoption of cloud. They also noted that these security problems that have been known earlier might play a positive role in cloud adoption, even though being problems with cloud security, because there are already existing solutions for them which can be implemented in cloud environments.

Traditional security models normally create a security boundary within stored sensitive data and self-control of computing resources. In many cases this boundary is firewall (Pearson & Benameur, 2010). According to Pearson and Benameur (2010) this model does not work in the case of public and hybrid clouds where the security boundaries become blurred, because sensitive information might be processed outside of known security boundaries. This is due to indistinct boundaries of data storage and processing. This creates the trust issue, which has been featured in the academic discussion around the cloud technology since its discovery. To ease this trust issue there needs to be more transparency in cloud environments to ease the concern of possible data breaches and to comply with regulatory aspects (Chow et al., 2009). Transparency helps to create trust around cloud environments and eases the doubt created by certain issues that may not be as severe as they seem like.

All cloud environments are different when it comes to privacy, security and trust requirements (Takabi, Joshi, & Ahn, 2011). One concern in privacy and security in cloud environments is the lack of control. The amount of control a cloud consumer has varies with the service model, just like the security responsibility. The responsibility of security in cloud environments also varies a lot depending on the deployment models and service models. The clearest variation in responsibility can be seen in service models. The more control the cloud user has, the more security responsibility is placed on user (Mogull et al., 2017). This variety of responsibility of security in different service models is shown in figure 3 below. Variety of responsibility of security in different service models creates a linear model for growing responsibility for security when moving from SaaS to PaaS to IaaS. The responsibility for security grows linearly with grown freedom of the user inside the environment.

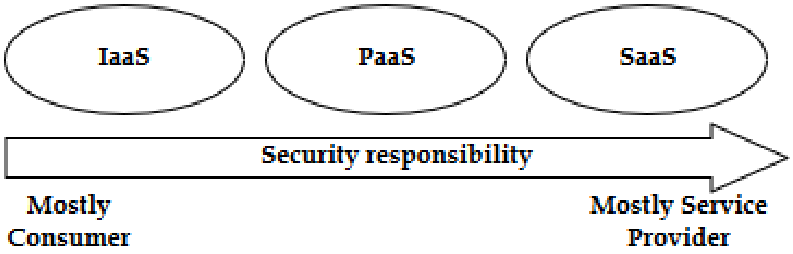


FIGURE 3 Shared security responsibility of service models (Mogull et al., 2017)

3.1.1 SaaS security

When utilizing a SaaS model, user does not have much control on security nor underlying architecture and infrastructure. Normally in SaaS, user can only access and manage the application they have license for, and cannot alter how the application is implemented or how it works (Mogull et al., 2017). Mogull et al. (2017) clarified this by an example where SaaS user is responsible for only managing the authorization and entitlements and SaaS provider carries the responsibility for application security, perimeter security and auditing and monitoring the use of the environments and keeping logs of transactions and sign-ins. In SaaS environment the service provider is responsible for the stored data because the cloud users cannot affect or view the underlying infrastructure as stated earlier. Pearson and Benameur (2010) defined this problem as the lack of user control. They also stated that the lack of control might force the users to move to a different service provider. According to European Data Protection Supervisor (EDPS) (2018) the specific security issues service model SaaS faces are:

- Procuring or acquiring SaaS without sufficient security consultation may lead decision makers to underestimate the risks or lead them to choose unfitting safeguard
- Lack of control and transparency over the technical infrastructure, organizational and technical safeguards and over the application code
- Basically noexistence control over the security measures if user authorisation and authentication is not counted
- Low implementation of auditability

Cloud user in SaaS has access to software application, but can only control the data that is processed and configuration of the application (EDPS, 2018). Overall cloud user has very low control over anything else than data that is processed and configuration of the application and tools to accommodate the rights of the data subject may be lacking. SaaS also faces lack of portability, but it could be increased by specific formats. Also, specific workflows, application business rules, settings and dependencies from other applications are possible constraints to increase portability. (EDPS, 2018).

3.1.2 PaaS security

The security responsibility between user and provider in PaaS differs from the security responsible in SaaS. In PaaS the user has more freedom to decide what to do in the cloud when they are paying only for the platform where they can develop and implement different solutions. When it comes to security, the PaaS provider is responsible for the security of the platform, not the applications cloud user has implemented on it. (Mogull et al., 2017). PaaS gives more freedom to user, but with this freedom comes wider security responsible. Compared to SaaS the responsibility for security in PaaS is shared more evenly with the provider and user (Mogull et al., 2017). EDPS (2018) listed some specific security issues service model PaaS faces as:

- Lack of transparency over the technical infrastructure and technical safeguards
- Lack of full control over network security and total lack of control over physical security of the data centers
- Nonexistence or limited implementation in network level auditability and total lack of control in physical security auditability

Cloud user in PaaS can control only some of the configuration aspects of provided platform but cannot control the underlying infrastructure and physical security of the data centers (EDPS, 2018). However according to EDPS (2018) cloud users are able to control applications that are developed on the platform and processed data. Tools to accommodate the rights of the data subject can be developed in PaaS environment. Due to possible variety of software platform implementations and variety of performance issues PaaS may face some portability challenges (EDPS, 2018).

3.1.3 IaaS security

While the responsibility of security in PaaS is quite evenly split between the provider and consumer, in IaaS the consumer carries the greater part of the responsibility for security. IaaS provider is only responsible for the security of the underlying infrastructure and the user has to configure the security for everything they have built on it (Mogull et al., 2017). According to EDPS (2018) the specific security issues service model IaaS faces are:

- Lack of transparency over the technical infrastructure and technical safeguards
- Lack of control in low level machine software security and total lack of control in physical security of the data centers.
- Lack of implementation in network level auditability and total lack of control in physical security auditability.

According to EDPS (2018) the service provider allocates the virtual machines from pooled resources in IaaS service model. Although the cloud user is able to control the configuration of IT infrastructure over the applications that are developed over the software platform, but cloud user still has no control in physical security of the data center (EDPS, 2018). Tools to accommodate the rights of the data subject can be developed in IaaS environment and IaaS also has lower risks related to portability (EDPS, 2018).

3.1.4 Service level agreement

There are many important security considerations in cloud security. Mogull et al. (2017) defined the most important security consideration in cloud environments as the up to date knowledge of who is responsible for what. Consumer needs to know what the provider is providing and how it all works. When consumers have up to date knowledge of this they are able to notice the vulnerabilities and create or acquire the necessary means to fill or control the gaps or in some occasions move to a different service provider with wider responsibility of security (Mogull et al., 2017). This all and the responsibilities need to be addresses in Service Level Agreements (SLA). SLAs are used in multiple different business processes, not only in security (SLA Management Team, 2004). SLA is a document which defines the relationship between the cloud provider and consumer (Kandukuri, Ramakrishna, & Rakshit, 2009). SLA is used to guarantee the quality of service that is agreed (Dawoud et al., 2010). According to Kandukuri et al. (2009) SLA is exceedingly important document which defines cloud user's needs, provides a framework for mutual understanding, simplifies the relationship, reduces the area of possible misunderstanding, encourages dialogue and eliminates the unrealistic expectations. It also sets proper boundaries for security responsibilities. When done correctly both the provider and consumer know whom is responsible for what and what is the required level of service. SLA does not solely improve the trust issues, but with enough transparency it eases the uncertainty.

3.2 Privacy

There is no single definition for privacy. Privacy rights include collection, use, disclosure, storage and destruction of personally identifiable information (Mather et al., 2009) and the means to affect them. The concern for privacy issues in online environments is getting more attention after the EU General Data Protection Regulation (GDPR) became effective. GDPR regulates "the protection of natural persons with regard to the processing of personal data and on the free movement of such data" (European Comission, 2019). GDPR is an extremely important step for strengthening the fundamental rights of individuals in digital environments. GDPR is also an important factor clarifying the rules for public bodies and organizations in digital single market, which facilitates business (European Comission, 2019). Pearson and Benameur (2010) categorized privacy as a fundamental human right, especially in European standpoint. Privacy can be seen as Mather et al. (2009) defined it, accountability of organizations to its data subjects and the transparency to organizations practices regarding personally identifiable information. Privacy in cloud environments can be examined and defined from two different perspectives, from consumers and organizations perspective. These perspectives of privacy and their focus vary with different cloud environments. Pearson and Benameur (2010) also stated that context should be considered when defining privacy issues in cloud environments because of the variety of them. For example, the privacy issues a private cloud faces differ from the ones that public cloud faces, and the same goes for the different service models as well. The character of the information also affects the privacy risk cloud faces, if information is meant as public and planned to be soon published, the privacy risk can be very low (Pearson & Benameur, 2010). The privacy risks and the need for privacy require close attention when the information that is handled in cloud is sensitive. If the information, that is collected, transferred, processed, shared and stored in dynamic cloud environment, contains personally identifiable information the privacy risk is significant (Pearson & Benameur, 2010). Pearson and Benameur (2010) listed several privacy concerns that public clouds especially face. According to Pearson and Benameur (2010) these issues include: ”lack of user control, potential unauthorized secondary usage, data proliferation, transborder data flow and dynamic provisioning”. In addition to these issues the retentation and disposal of data, and who controls it, is a key concern in cloud environments. In case of privacy breaches the faulty party needs be concludable and repair measures need to be known and ready in such cases. According to Gartner (2008) cloud service providers and their need to test, verify and ask the right questions from service developers to identify vulnerabilities (Heiser & Nicolett, 2008). According to Pearson and Benameur (2010) public cloud might not be suitable for treating sensitive data, at least in its state of privacy and security level of 2010.

Unauthorized secondary usage is also a security issue that needs to be taken into account (Pearson & Benameur, 2010). This issue needs to be adressed in user agreements before registration. According to Pearson and Benameur (2010) autharized secondary use of user data has been a standard business model for cloud providers. This authorized secondary use of user data is normally addressed in advertisements. Pearson and Benameur (2010) also mentioned that in case of bankcruptcy of the cloud provider or if the cloud provider is acquired by other company, it might not be stated in the contracts that what would happen to the data that is stored in said cloud environment. Thus cloud consumers need to be aware of what is stated in contracts such as SLAs.

Data that is stored in cloud environments is often replicated to reach higher availability. Required availability levels are often stated in SLAs. This process increases the amount of data that cloud provider is responsible for.

Pearson and Benameur (2010) defined this increase of data as data proliferation and listed it as one of the main privacy issues of cloud environments. Data proliferation causes difficulties when determining where the exact data is stored, especially in case of deletion of said data. Data proliferation is also connected to transborder data flow because most cloud providers have decentralized their data centers over the national borders. Like Chow et al. (2009) stated the problems with with cloud environments being multinationally decentralized, which is also causing the transborder data flow, these problems are quite similar as traditional outsourcing. According to Varghese and Buyya (2018) centralized data centers create plausiblible single point failures. Thus data centers are often geograhical decentralized which means that even the sensitive data that is in the cloud need to be transferred from its source to a different location. Transborder data flow is an issue even with sensitive data, because that sensitive data might be stored in a different country (Varghese & Buyya, 2018). When sensitive data is moved over and between national borders it might also cross the borders of legal jurisdiction (Pearson & Benameur, 2010). Transborder data flow is an issue, especially with the legistlation that changes while data is being transferred to a different country to be stored or processed.

Data security is one of the most troublesome issues regarding the cloud computing security. There are many proposed solutions to it, but these solutions happen to focus on only single stages of data life cycle (Yu & Wen, 2010). Data life cycle consists of 7 phases (see figure 4 below). According to Mather et al. (2009) these data life cycle phases are generation of information, use, transfer, transformation, storage, archival and destruction. Yu and Wen (2010) stated that focusing in only one phase of data cycle is not enough to reach sufficient level of data security because most issues affect data in its whole life cycle.

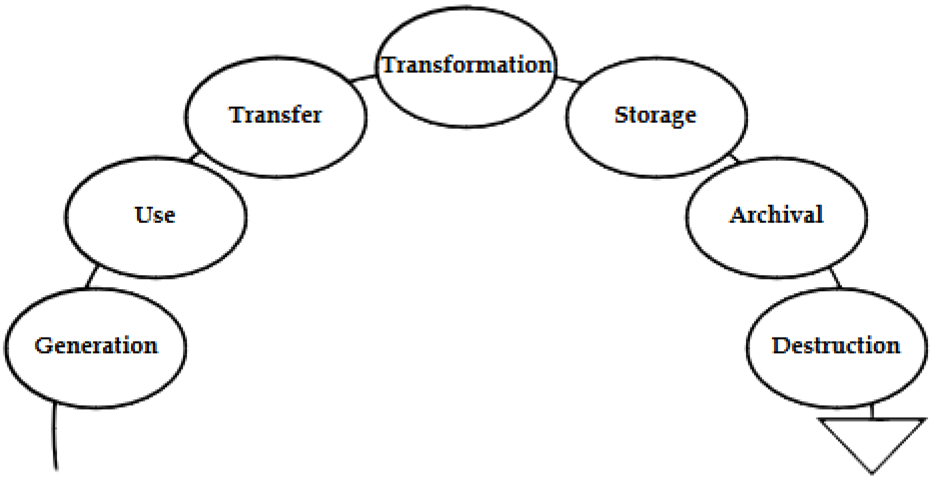


FIGURE 4 Data life cycle (Mather et al. 2009)

3.3 Cloud forensics and logging

In cloud environments malicious parties can exploit weaknesses by either attacking applications that run inside the cloud or launching attacks from machines that run inside the cloud. These kinds of issues are the concern cloud forensic is meant to solve. (Zawoad, Dutta, & Hasan, 2013). Cloud forensics is a new branch of digital forensic for cloud environments which can be defined as "applying computer forensics procedures in a cloud computing environment" (Zawoad et al., 2013). Ruan et al. (2013) proposed a definition for cloud forensics based on their survey results. They defined cloud forensic as a "application of digital forensic science in cloud computing environments". In more detail cloud forensics consists of a hybrid forensic approach, which includes at least virtual, remote, live, network, large-scale, thick-client, thin-client forensics, to generate digital evidence of different kind of events and actions in cloud environments. (Ruan et al., 2013). The definition varies between legal and organizational viewpoints. In legal viewpoint it commonly implies "multi- jurisdictional and multi-tenant situations" and in organizational viewpoint cloud forensics involves interaction with different cloud actors for internal and external investigations and auditions. (Ruan et al., 2013). According to Zawoad et al. (2013) many cloud computer architectures do not have suitable support for forensic investigations. Collecting and analyzing logs is important part of computer forensics, but when collecting logs from a cloud is more complicated matter. When collecting logs from cloud environments where computation and storage resources are shared, log API or cloud manager console is needed to collect and categorize the logs correctly. (Zawoad et al., 2013). Collecting logs from cloud environments is quite complicated because investigators or parties that require log information normally have very little control over the underlying infrastructure that supports the cloud. If users cannot collect the logs by their own means, high level of trust between the user and provider is required because it is extremely hard or impossible to verify that is the provided log information is valid or not. (Zawoad et al., 2013). According to Zawoad et al. (2013) shutting down a virtual machine where log information is wanted from, it is impossible to collect log information from terminated virtual machine. Zawoad et al. (2013) also raised their concern for means of preserving users' privacy and integrity when providing logs and highly sensitive information for investigation.

Like many other things in cloud environments, the cloud forensics procedures also vary in different deployment and service models. In SaaS and PaaS users have limited control over the network and process monitoring and they are more dependent on the logs provided by cloud service provider. But in IaaS users have more control and implementation of forensic friendly logging procedures or mechanism is possible. (Zawoad et al., 2013). The procedures for private and public deployment models vary as well. In public cloud the physical access to digital evidence is most likely impossible, whereas in private cloud physical access is easily provided. (Zawoad et al., 2013)

According to Marty (2011) log information should be collected from all infrastructure, not just from the user interface, and transported to a central log collector for analysis. Marty (2011) also proposed guidelines of where to focus in logging which varies with the environments and use cases. But he proposed the logs should include that at least the following information:

- Timestamp
- Application
- User
- Session ID
- Severity
- Reason
- Categorization

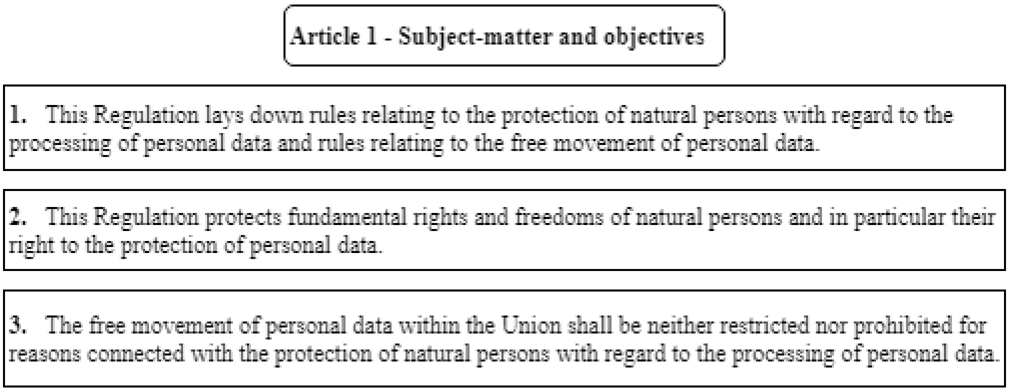
According to Marty (2011) these sections are needed to answer when, what, who and why question. Timestamp provides the information when the recorded event happened. Application field provides the information what application the log is from. (Marty, 2011). User field identifies the exact user through unique ID or user name. A session ID field is used to track single requests through varying tiers and applications. (Marty, 2011). Severity field categorizes the log information based on their significance or importance. The reason field aims to identify why something happened. (Marty, 2011). Categorization field categorizes the similar events through some identifier such as failed logins (Marty, 2011). According to Marty (2011) this field is highly important when analyzing logs or trying to find certain type of logs, which would be difficult without a simple category field that addresses all the certain type of log records.

4 General Data Protection Regulation

This chapter clarifies how GDPR compliance can be achieved in cloud computing. This chapter examines and gives background for GDPR and how it works especially with cloud computing environments. First there is an overview to GDPR. After this the goals of the GDPR are streamlined and enforcement mechanisms are examined. There is also a look into possible sanctions an organization may face incase a breach that has compromised PII that the organization is responsible for.

4.1 Overview to GDPR

EU General Data Protection Regulation came into effect in EU on 25th May 2018 (GDPR, 2016). The objectives and the subject-matter of the GDPR are to protect natural persons, their rights and the freedom of the movement of their personal data (GDPR Article 1, 2016) which can be seen in figure 5.



Article 1 - Subject-matter and objectives

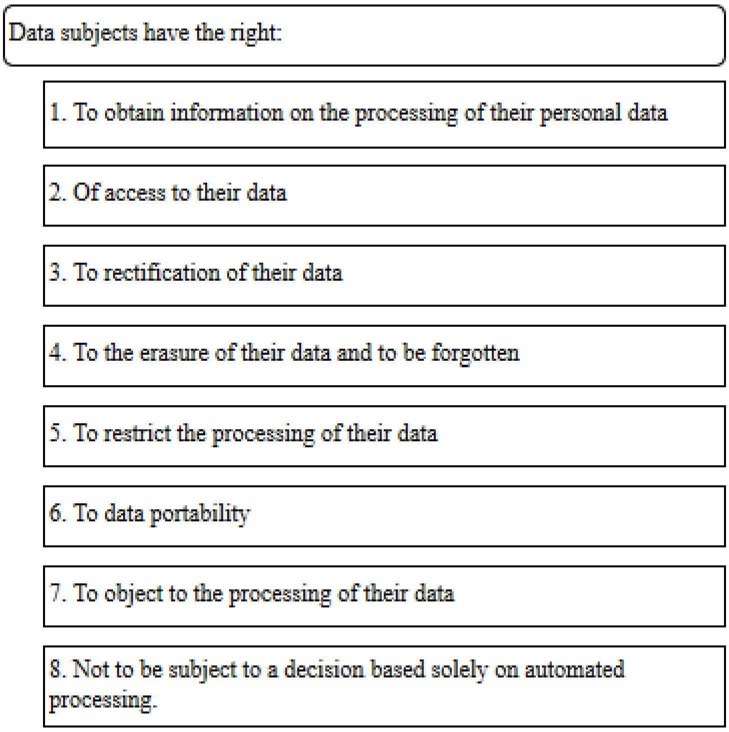
1. This Regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data.

2. This Regulation protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data.

3. The free movement of personal data within the Union shall be neither restricted nor prohibited for reasons connected with the protection of natural persons with regard to the processing of personal data.

FIGURE 5 Subject-matter and objectives (GDPR Article 1, 2016)

According to the Office of the Data Protection Ombudsman (2019) controllers needs to take appropriate measures ensuring that the data subjects' data protection rights are fulfilled when ever processing personal data. The Office of the Data Protection Ombudsman (2019) listed the rights of the data subject according to the GDPR, which can be seen below in figure 6. Facilitating the data subject's rights is also required from the controllers.

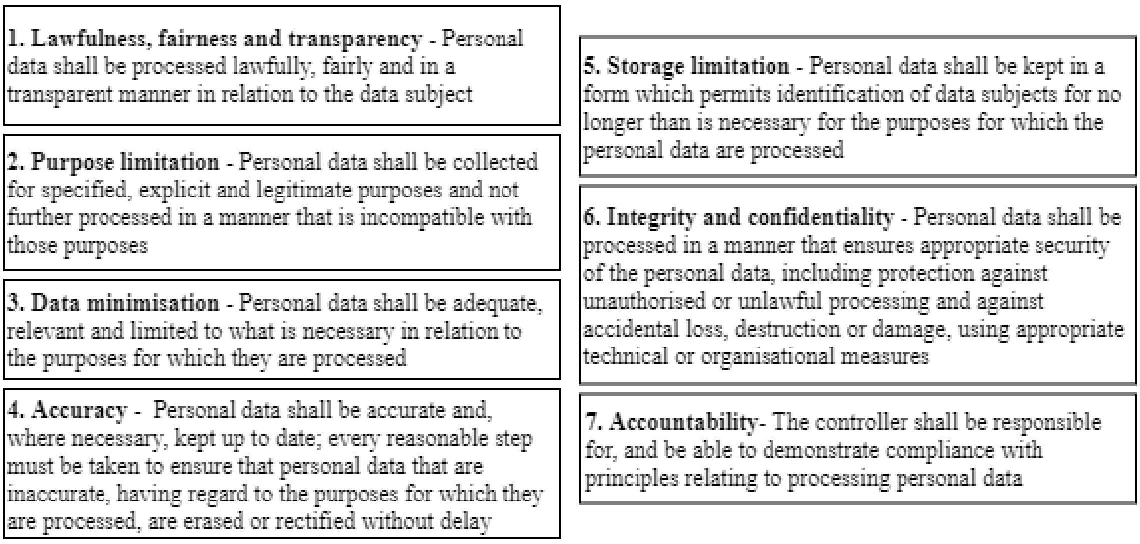


Data subjects have the right:

1. To obtain information on the processing of their personal data
2. Of access to their data
3. To rectification of their data
4. To the erasure of their data and to be forgotten
5. To restrict the processing of their data
6. To data portability
7. To object to the processing of their data
8. Not to be subject to a decision based solely on automated processing.

FIGURE 6 Rights of the data subject (The Office of the Data Protection Ombudsman, 2019; GDPR, 2016)

According to the Office of the Data Protection Ombudsman (2019) compliance with the data protection principles is required when ever processing personal data. Data protection principles from GDPR Article 5 can be seen below in figure 7.



1. Lawfulness, fairness and transparency - Personal data shall be processed lawfully, fairly and in a transparent manner in relation to the data subject

2. Purpose limitation - Personal data shall be collected for specified, explicit and legitimate purposes and not further processed in a manner that is incompatible with those purposes

3. Data minimisation - Personal data shall be adequate, relevant and limited to what is necessary in relation to the purposes for which they are processed

4. Accuracy - Personal data shall be accurate and, where necessary, kept up to date; every reasonable step must be taken to ensure that personal data that are inaccurate, having regard to the purposes for which they are processed, are erased or rectified without delay

5. Storage limitation - Personal data shall be kept in a form which permits identification of data subjects for no longer than is necessary for the purposes for which the personal data are processed

6. Integrity and confidentiality - Personal data shall be processed in a manner that ensures appropriate security of the personal data, including protection against unauthorised or unlawful processing and against accidental loss, destruction or damage, using appropriate technical or organisational measures

7. Accountability- The controller shall be responsible for, and be able to demonstrate compliance with principles relating to processing personal data

FIGURE 7 Data processing principles 1 (GDPR Article 5, 2016; Office of the Data Protection Ombudsman, 2019)

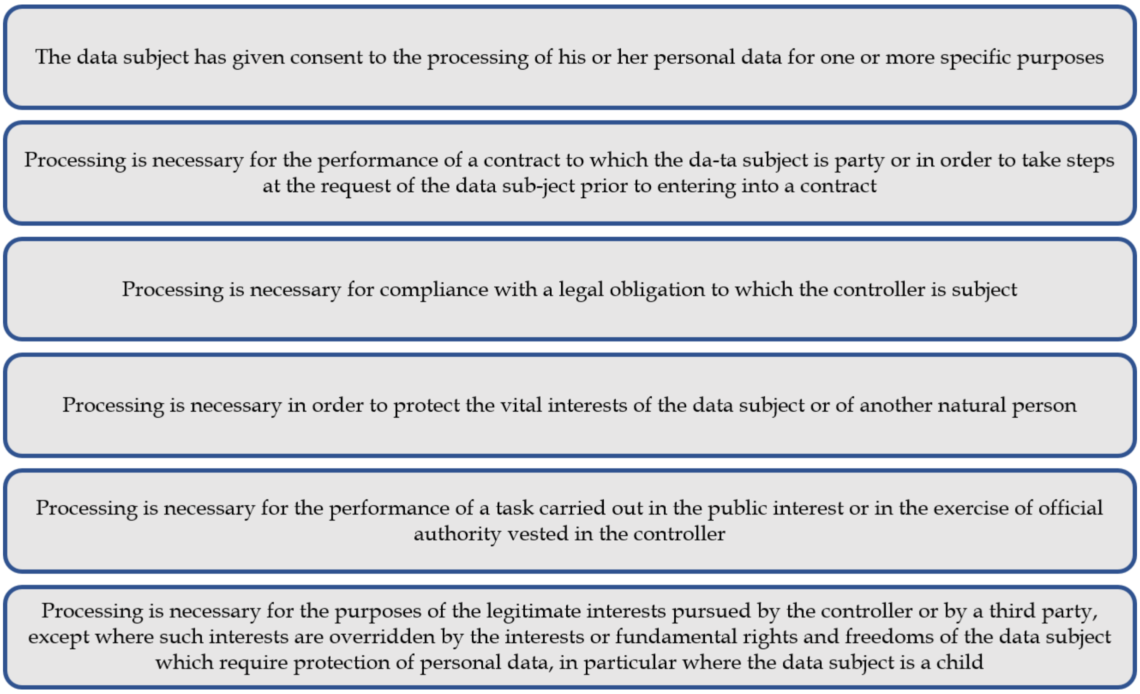
According to Duncan (2018) many organizations were inadequately prepared for new legislation. Information security causes challenges to all organizations who use traditional distributed network systems, but the challenges increase exponentially when cloud environments are utilized (Duncan, 2018). Many organizations that only utilize conventional information systems are having issues in complying with new regulations. But organizations that utilize any kind of cloud computing environments are having more complicated issues with it. (Duncan, 2018). According to Duncan (2018) the Cloud Forensic Problem is especially challenging. Duncan (2018) stated that even without the cloud forensic problem cloud computing environments are more complicated security environments, but this problem presents even more challenging barrier to compliance. According to Duncan (2018) the cloud forensic problem is especially challenging because all information systems are constantly attacked, but in case of cloud environments it is harder to prevent the intruder from getting their hands on a data that is covered by GDPR. Also, intruder is able to delete traces of the incursion which makes it harder to follow the traces in cloud forensic, and they might also delete other records in the process (Duncan, 2018).

Typically, a cloud service provider would qualify as a processor when your enterprise uses their services. The cloud service provider will process personal data, which are stored within their databases or servers, on your behalf: the controller. The cloud service provider cannot do anything with your data, unless you instruct them to do so and the data remain within your controllership. (Tolsma, 2019). GDPR affects all existing organizations that deals with even a single resident from EU, the organizations need to ensure that they are compliant with GDPR. If the company that is dealing with data of EU citizens suffers from a security breach that compromises the records of any EU resident, the GDPR is extended globally. (Duncan, 2018). Duncan (2018) stated that if the cloud forensic problem is not resolved in companies that are utilizing cloud environments, it will be very hard or even impossible to comply with GDPR.

The Data protection Working Party was founded under the terms of Article 29 of the Data Protection Directive in 1996 by the European Commission (Data Protection Working Party, 2012; Duncan, 2018). According to Duncan (2018) the article 29 Working party has been overseeing the development of GDPR and has been giving proposals for amendments. One of these proposals was to require organizations to report all breaches within the 72 hours of occurring, but which was later changed to requirement of reporting breaches within 72 hours of discovering the breach. According to Article 33 the processor must notify the controller without a delay incase of personal data has been breached. Duncan (2018) clarified the goals of the GDPR, its enforcement mechanisms and what will happen in case of data breach. The next three subsections will investigate these aspects of GDPR in cloud computing environments.

4.2 Clarified goals of the GDPR

Organizations need to streamline compliance by providing rules that would be same for everyone and would apply anywhere in EU using a One Stop Shop approach, which is covered in the GDPR in Articles 46 to 55 (Duncan, 2018; GDPR Articles 46 to 55, 2016). By this, creating a clarified approach for organizations inside and outside the EU is possible and preferred (Duncan, 2018). According to Article 6, processing of personal data must follow at least one of the principles from figure 8 below to be lawful (GDPR Article 6, 2016):



The data subject has given consent to the processing of his or her personal data for one or more specific purposes

Processing is necessary for the performance of a contract to which the da-ta subject is party or in order to take steps at the request of the data sub-ject prior to entering into a contract

Processing is necessary for compliance with a legal obligation to which the controller is subject

Processing is necessary in order to protect the vital interests of the data subject or of another natural person

Processing is necessary for the performance of a task carried out in the public interest or in the exercise of official authority vested in the controller

Processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child

FIGURE 8 Data processing principles 2 (Duncan, 2018)

Data subjects have a right to access personal data that is in possession of any organization that is compliance with the GDPR as described in the Article 15. (Duncan, 2018; GDPR Article 15, 2016). Right to Erasure in Article 17 provides right for the data subject to have certain data erased that is held by an organization that is compliance with the GDPR. The freedom of the data subject may overrule the legitimate interest of the controller in this kind of a case, which means that the controller needs to erase the data that data subject wants to be erased. (Duncan, 2018; GDPR Article 17, 2016). According to Duncan (2018) data subjects have rights in data portability, which is under the Article 20 of the GDPR. In data portability, data subjects are able to transfer personal data between electronic processing systems without data controller prevention (Duncan, 2018). Article 25 of the GDPR handles the data protection by design and by default (Duncan, 2018; GDPR Article 25, 2016). This article aims to ensure that privacy may be expected by the design, which is included in development of business processes (Duncan, 2018). When defining privacy and data protection by design it is especially important to highlight that encryption and decryption operations needs to be carried out fully locally and not by remote services (Danezis, Domingo-Ferrer, Hansen, Hoepman, Métayer, Tirtea & Schiffner, 2015). According to Duncan (2018) this means that privacy requirements by default should be at a high level. Duncan (2018) also clarified that technical and procedural measures are better to leave for controller to take care of to make sure processing in whole processing lifecycle follows the regulation.

According to Duncan (2018) the consent for the processing of the data subject's personal data for one or more specific processing purposes, needs to be necessary for:

- Taking steps at the request of the data subject before the contract is valid
- The performance of the contract data subject is accessory
- Compliance of controller with legal obligations as a subject
- To protect vital interests of data subject or other natural persons
- Carrying out a task in the public interest or by exercising official authority that is vested in controller
- Reasons of controllers or third parties' legitimate interests. These reasons cannot conflict with the fundamental rights, freedom or interests of the data subject which would require protection of personal data. Especially if the data subject is a child.

Article 7 and Article 4 of the GDPR defines that the consent from data subjects needs to be explicit about the data that is being collected and the purpose it is used for. Because of the nature of consent, data controlled need to be able to prove that they have the consent for the data which can also be withdrawn (Duncan, 2018; GDPR Article 7 & Article 4, 2016). According to Article 8 of the GDPR, If the data subject is a child, the verifiable consent needs to come from legal guardian of the child (Duncan, 2018; GDPR Article 8, 2016).

4.3 Enforcement mechanisms of the GDPR

Data protection officer needs to be appointed for all organizations that are processing data or regarded as data processor organization. Data protection officer needs vast experience and knowledge of data protection legislations and is appointed to assist organization in monitoring internal compliances with regulations. (Duncan, 2018). According to Duncan (2018) appointing the data protection officer may turn out to be challenging for the boards of large organizations because of human factor issues and myriad governance. In addition, the data protection officer needs to act independent inside the organization and will need to create a suitable support team. Duncan (2018) defined data protection officers' role as a "mini-regulator" within the company.

Data protection by default and data protection by design principles should be implemented by the data controller. This is mainly done by compliance demonstration to ensure compliance with the GDPR by ensuring that all required mechanisms are properly in place and defined correctly. (Duncan, 2018). The process of pseudonymizing, which is defined in Recital 78 of the GDPR, by encryption is one of these measures and it should be done as soon as it is possible (Duncan, 2018; GDPR Recital 78, 2016).

According to Duncan (2018) one goal of the GDPR is to provide accountability and responsibility for and by all parties that are involved in processing data. This needs to be done with wider notice requirements that cover the retention time for personal data and for data controllers and data protection officers contact information (Duncan, 2018). Automated decision-making for individuals, such as Article 22 of the GDPR's defined algorithmic means of profiling, is paid more attention (Duncan, 2018; GDPR Article 22, 2016). All actors who are included in any part of data processing processes are expected to be accountable for their actions and act responsibly (Duncan, 2018). According to Duncan (2018) high risks require risk assessment and risk mitigation, as well as prior approval from data protection authorities. Data protection impact assessment, like described in Article 35 of the GDPR, must be conducted is specific risks have occurred to data subject's freedoms and rights.

4.4 Data breach and sanctions

GDPR oblige data controllers to notify supervisory authority without unreasonable delays if data breach has occurred. According to Article 33 of the GDPR, data breaches must be reported within 72 hours of it's discovery to the supervisory authority. (Duncan, 2018; GDPR Article 33, 2016). Article 34 of the GDPR states that individuals must be informed incase of adverse impact, except if the data is encrypted. In addition Article 33 of the GDPR states that controller needs to be notified by data processor, incase of personal data breach, and it needs to be done without unreasonable delay. (Duncan, 2018; GDPR Article 33, 2016). A data breach happened in Salesforce.com Marketing Cloud in June 2018 which was caused by a rest API error (Schwartz & Ross, 2018; Esage, 2018; Salesforce.com, 2018). According to Salesforce.com (2018) the error was caused by a code change that allowed customers to view metadata of other customers. According to Schwartz and Ross (2018) Salesforce.com might still not be entirely sure that was customer data modified or not. Which leads to a question were there any proper logging mechanisms integrated to the Marketing Cloud to ensure its security. According to Salesforce.com they did not have any evidence that any malicious behavior happened, but they also added that they are unable to verify that certain customers data was not viewed or modified (Schwartz &

Ross, 2018). According to Schwartz and Ross (2018) Salesforce.com became aware of the issue they released an emergency release to resolve the issue, but they issued email alerts to their potentially affected customers 15 days later.

Duncan (2018) specified four possible sanctions in case of GDPR non-compliance (GDPR, 2016). First, a written warning can be imposed incase of non-intentional first non-compliance. Second, a regular periodic data protection audits can be imposed to make sure compliance with GDPR can be achieved. Third, a fine as high as 10 million euros or up to 2% of the whole annual turnover of the previous financial year. In this case this fine is imposed, it will be whichever is higher. (Duncan, 2018; GDPR Article 83, 2016). According to Duncan (2018) and Article 83, Paragraph 4 of the GDPR, this fine can be imposed if organization has had an infringement of the provisions that are found in table 2 below.

TABLE 2 Infringements leading to sanctions 1

"The obligations of the controller and the processor pursuant to Articles 8, 11, 25 to 39 and 42 and 43"
"The obligations of the certification body pursuant to Articles 42 and 43"
"The obligations of the monitoring body pursuant to Article 41(4)"

And in fourth, a fine as high as 20 million euros or up to 4% of the whole annual turnover of the previous financial year In this case this fine is imposed, it will be whichever is higher. (Duncan, 2018; GDPR Article 83, 2016). According to Duncan (2018) and Article 83, Paragraph 5 & 6 of the GDPR, this fine can be imposed if organization has had an infringement of the provisions that are found in table 3 below.

TABLE 3 Infringements leading to sanctions 2

"The basic principles for processing, including conditions for consent, pursuant to Articles 5, 6, 7 and 9"
"The data subjects' rights pursuant to Articles 12 to 22"
"The transfers of personal data to a recipient in a third country or an international organization pursuant to Articles 44 to 49"
"Any obligations pursuant to Member State law adopted under Chapter IX"
"Non-compliance with an order or a temporary or definitive limitation on processing or the suspension of data flows by the supervisory authority pursuant to Article 58(2) or failure to provide access in violation of Article 58(1)"

According to Duncan (2018) the information and details provided above are needed in order to understand what information is required and what actions need to be performed incase of a data breach for data processor to be compliant with the GDPR.

5 RESEARCH METHODOLOGY

The empirical research is introduced in this chapter. First the goal of the research is introduced which will be followed by case company introduction. After this the qualitative methods are introduced. Finally, the data collection analysis phase of the study is introduced.

5.1 The goal of the research

The goal of this Masters' Thesis is to find out what precautions and actions organizations need to take before moving to a cloud computing environment especially when the cloud is provided by a third party. This study aims to combine the answers from the literature with the results from empirical research to create a theory or guidelines for organizations that are planning to move their business processes or systems to operate in cloud. To answers this challenge a research question was defined:

- What are the needed actions and precautions an organization must take when storing and processing PII in a cloud computing environment provided by a third party?

To help to define the research problem and to get better understanding of this multidimensional problem two focusing questions were defined:

- How does cloud as an environment differ from traditional IS in control and responsibility and how can organizations preserve the control in cloud environment?
- What are the needed actions to ensure privacy and security in cloud computing environment?

Answering to the research questions aims to provide a theory or a guideline for case company and other organizations that are facing challenges with clouds. The next chapter gives a short introduction to the case company.

5.2 Theoretical background

Cloud computing environment is a combination of deployment models, service models and exhibiting characteristics of cloud computing. Cloud computing has had a huge impact on IT field creating massive amount of new solutions and completely new ways to do things via internet. Mell and Grance (2011) defined the essential characteristics that cloud computing must exhibit so it can be called cloud computing. These characteristics are on-demand self-service, broad network access, resource pooling, rapid elasticity and measured service. Mogull et al. (2017) stated that if information system is missing any of these characteristics it is most likely something else than cloud computing. There are four commonly accepted cloud deployment models. These are private cloud, community cloud, public cloud and hybrid cloud (Mell & Grance, 2011). Private cloud is used by only a single organization. Community cloud is for the use of a certain community of people or group of organization sharing the same regulation and policies. Public cloud is publicly available cloud commonly run by commercial organization. Hybrid cloud is a little more complex because it is combination of two or more deployment models that can be configured in many ways. (Mell & Grance, 2011). Cloud computing has three common deployment models. These are IaaS, PaaS and SaaS. IaaS is a deployment model where consumer buys an abstracted pool of infrastructure resources and build a platform and software on top of them (Mell & Grance, 2011; Mogull et al., 2017). PaaS is a deployment model where consumer buys the abstracted pool of infrastructure and the platform and then develop the software on top of them. SaaS consumer buy software as a service and therefore consumer does not have the need for development.

To understand the security and privacy in cloud computing environments the understanding of how cloud is constructed is essential. Although cloud is in many ways different compared to traditional on-premise information systems many security processes used in on-premise systems are still usable. The challenge in cloud computing security comes from the diversity of cloud environments and the inevitable loss of control. Mogull et al. (2017) stated that the responsibility for technical security of the cloud environment raises with the control consumer has. Mogull et al. (2017) also stated that when moving to a cloud the amount of responsibility is highest when utilizing IaaS service model, medium when utilizing PaaS and lowest when utilizing SaaS.

Privacy in the cloud environments does not differ from the privacy in the traditional on-premise information systems. But the concern for privacy is higher when operating in the cloud. Pearson and Benameur (2010) stated that when processing personally identifiable information in the cloud there is a serious need to pay attention to privacy and possible privacy risk. When processing PII in cloud computing environment processors and controllers need to pay extra attention to data life cycle and access control. Data that is stored and processed in cloud is commonly replicated to ensure the service levels which increases the amount of data in the cloud (Pearson & Benameur, 2010). Thus, it is important to have up to date knowledge where the data is stored. Data life cycle work the same way in the cloud as it would in on-premises systems. The difference is that cloud consumer might no longer have the same level of control over the data thus ensuring the privacy from generation of data to destruction of data becomes harder and less transparent.

To ensure the privacy and security in the cloud there is a need for security mechanisms like logging. Logging is also an obligation from the GDPR. Marty (2011) argued that logs should be collected from all infrastructure. Marty (2011) stated that logs must be able to answer question when, what, who and why. Thus, the collected logs should include at least the information of timestamp, application, user, session ID, severity, reason and categorization (Marty, 2011).

EU General Data Protection Regulation came into effect in EU in 2018. The GDPR aims to protect natural persons, their rights over their personal data and the freedom of the movement of said data. GDPR protects the privacy of EU citizens. The office of data protection ombudsman (2019) stated that controllers must ensure that data protection rights are fulfilled for the data subjects. The Office of Data Protection Ombudsman (2019) also stated that organization need to be compliant with the data protection principles from the from GDPR Article 5. These principles are:

- Lawfulness, fairness and transparency
- Purpose limitation
- Data minimization
- Accuracy
- Storage limitation
- Integrity and confidentiality
- Accountability

According to Duncan (2018) many organizations seemed to be inadequately prepared for the GDPR. Duncan (2018) also stated that the challenge with the GDPR is greater for the companies utilizing cloud computing environments. One of these problems is with cloud forensics. A good example of a problem with cloud forensics is the data breach with Salesforce.com Marketing Cloud in June 2018 where service provider was unable to ensure that customers data was not viewed or modified by malicious third party most likely because of insufficient security mechanisms. Schwartz and Ross (2018) stated that this led to a question that did Salesforce.com have a proper logging mechanism integrated in their service.

The control over the environments is shifting when moving to a cloud. When organization is using a traditional on-premise information system they have the most control over the environment. But when they are moving to a cloud they start losing control over some functionalities. There are way to preserve this control. These ways are the deployed security mechanisms and contracts and agreements that are made between the cloud provider and cloud customer. Although organizations lose controls over the environment they will not lose the legal responsibility. Thus, there is a need to ensure that the level of security and privacy is maintained in the cloud either by the mechanisms provided by the cloud operator or deployed by the cloud consumer. The research model for the empirical research is derived from the problem of understanding what happens in the process of moving from traditional on-premise information system to the cloud and how can organizations preserve the needed control to maintain their data privacy and security and maintain GDPR compliant. This challenge can be seen in research model below in the figure 9.

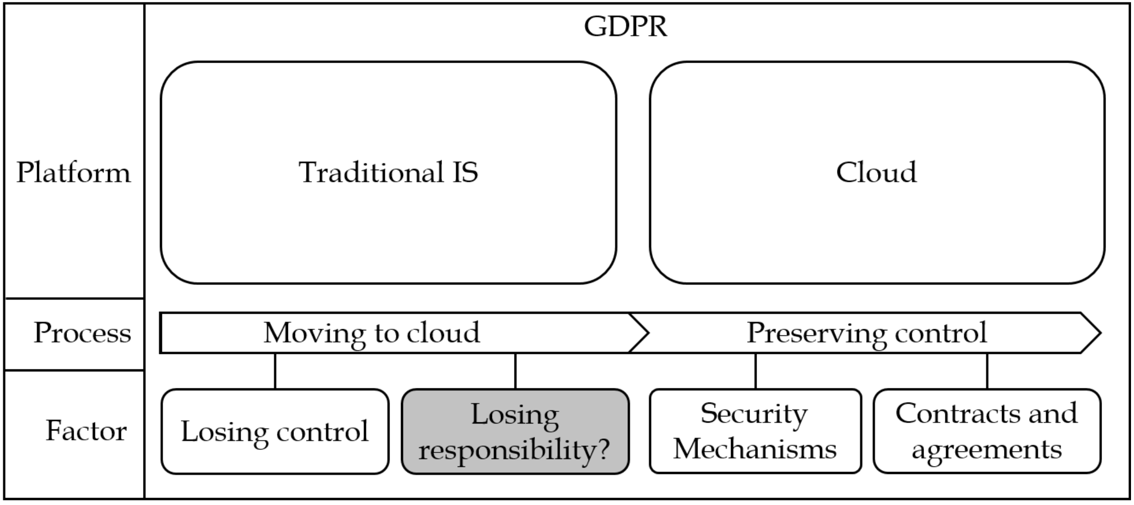


FIGURE 9 Research model

5.3 Case company introduction

The case partner for this case study is a teleoperator that is operating in Finland. All teleoperators in Finland have millions of customers whose personal data they are responsible for. There are three significant teleoperators in Finland with quite close market shares. Even without a tight competition teleoperators do not have much room for data breaches because they need to be compliant with the GDPR and Information Society Code (Information Society Code, 2014).

The contact with case partner was made when I was working in the field. After discussing the challenges that moving to cloud may cause within the organization the need for profound research that would suite the interests of company and my academic studies were raised. Although the case company is already using clouds in their processes there we still some challenges that needed more profound review. Case company also has ongoing cloud development projects and projects where business functionalities are being moved to a cloud. Thus, the timing for this research was beneficial for both parties and the aim was to provide knowledge and information how to operate in projects where data processing and other business functionalities are moved to the cloud. To gain a comprehensive view of the challenges the focus in the study was on the employees who work closely with clouds either in technical, legal or other professional positions.

5.4 Qualitative methods

This empirical research is conducted as qualitative research. Qualitative research is suitable for this research because the topic and research problems are relatively new and there are no completed theories to answer the problems. This qualitive research is a case study which is conducted using interviews. In this study the interview will be conducted as a semi-structured interview, also known as theme interview, which includes structured questions and open discussion (Hirsjärvi & Hurme, 2000). The objective of the study is to find out what are the precautions case organizations need to take when using cloud environments. This objective has both, legal and technical perspective and the goal of the study is to gather information from case organizations employees working with cloud computing, information security and with GDPR and other legislations, the qualitative research is optimal data gathering technique to reach the research objective.

According to Järvinen (2012) theory-creating studies, such as qualitative research, works well when the studied phenomenon is not well known beforehand. Interview as a research method can be seen as a conversation between the respondent and the interviewer with the purpose extracting information. The difference between an interview and a normal conversation is that interview has a concrete object (Hirsjärvi & Hurme, 2014). In semi-structured interview the conversation during the interview is kept in framework or the study by prepared questions. Semi-structured interview aims to value conversation of research themes over the detailed questions (Hirsjärvi & Hurme, 2014). This allows interviewees to speak more freely of the theme, rather than just answering the questions interviewer has designed beforehand. Semi-structured interview is more closely related to unstructured interview than structured interview. The reason it is still called semi-structured interview is because of the theme, topic and questions are the same for everyone (Hirsjärvi & Hurme, 2014). Semi-structured interview is suitable for this study’s’ research problem because we are aiming to find suitable solutions for practical challenges besides the theoretical challenges. Semi-structured interview also allows interviewees to bring forth their experience and knowledge of the theme without realizing even though it might not have been required to answer the question itself. The interviewees also work with different fields that are related to the study and they have different fields or expertise. The interview can not be done anonymously, thus we are able to ensure that only the essential interviewees answers are considered in each of the themes.

5.5 Data collection and analysis

According to Yin (2003), case studies may use several different data collection methods. Semi-structured interviews were chosen as a data collection method for this study. The aim of the research was to figure what are the needed actions organizations need to take in order to store and process PII in cloud computing environments provided by a third party. There are many different aspects that need to be considered in this kind of a research problem. This is a multi-dimensional problem that needs to be considered from technical, information security and legal points of a view. Therefore, the interviewees that have been chosen are working closely with different aspects of the problem and have experience and knowledge to answer the interview questions from their perspective. By doing this we were able to get a sampling that covers the multi-dimensional research problem comprehensively. Interviewees and their background information in presented in table 4.

TABLE 4 Background information of the interviewees

Interviewee	Field	Time in current position	Time in field
Person A	Security	5 years	21 years
Person B	Security	5 years	10 years
Person C	Security	1 month	31 years
Person D	Cloud dev	2 years	20 years
Person E	Legal	4 years	12 years
Person F	Privacy	3 years	5 years

The research was conducted individually for each interviewee. The research themes were derived from the literature and the research questions. To gain a comprehensive understanding of the topic and to answer the research question three research themes were defined:

- Moving to cloud
- The change in control
- The needed tangible actions

The interviews were conducted via skype and recorded using skype recorder –software. Although the interviews were conducted as a semi-structured interview, the interview followed questions that were prepared in advance, but additional focusing questions could be asked. After the interviews were recorded, interviews were transcribed accurately into a text form. Accurate transcription allows research data to be used in future researches and thus was transcribed verbatim. Transcribed interviews were then coded under the research themes for the analysis of the data. Coding is done by reading the research data and capturing its key thought and concepts that can be derived to codes (Hsieh & Shannon, 2005). Then the text is approached by making notes of the thoughts, impressions and initial analysis (Hsieh & Shannon, 2005). Then the codes are sorted into a different category by analyzing the codes relations and links (Hsieh & Shannon, 2005). The coding process does not return finished results, but it simplifies the analysis of the research data. The coded data that was categorized under the research themes is then used to derive results for the research questions. The results can be found from the chapter 6.

6 RESULTS

This chapter represents the results of empirical study that was conducted by semi-structured interviews. One of the main purposes of the study is to find out what needs to be done to successfully transition a system or business functionality to operate in cloud computing environment. This case study also focused to find out what are the things that organizations need to understand before moving to a cloud environment in security and legal point of a view. To answer these research problems, we first needed to find out what kind of changes transitioning to cloud causes compared to a traditional on-premises information system. Research model pictures in Figure 8, was utilized in conducting this case study to find out how cloud differs from traditional on-premises system, how does the control over the environment change and how can organizations prepare for the possible loss of control. The results are divided by three individual themes that were conducted from the literature. These themes were transitioning to cloud, the change in control and needed tangible actions.

The interviews were conducted in Finnish to prevent losing significant results if the interviewee is not comfortable with English. The quotations used in this chapter are translated in English from accurately transcribed interviews in a way that the original message remains. The next chapters represent the results extracted from qualitative interviews.

6.1 Transitioning to Cloud

Transitioning to cloud requires a lot of planning and understanding because the data is no longer in organizations on-premise servers. When data is stored to a cloud it is a big change from many points of a view compared to traditional information systems. Two of the interviewees mentioned that one of the biggest things that cloud changes is the physical and geological location of the data. There are many things that require careful planning because public clouds function in public internet. The location of the data and where it is accessible from is regulated precisely by GDPR. There is also lots of regulations for continuity of services for teleoperators where the geological location of the data might be essential factor in crisis situations. Cloud requires lot of new kind of planning that was not required with familiar on-premises information systems.

> "It is a big change that the data is no longer located and accessible from one place only. Thus, it needs to be thought little bit differently how it is managed and delimited, and there are lot of requirements from many directions"

> "For teleoperators there are lot of regulations for continuity of the services in Finland. And when we think about public cloud, how is the continuity secured if the borders are closed for example in war situations"

> "If cloud instance is planned poorly it is basically open from anywhere from the world"

Two interviewees also mentioned that there needs to be controls before moving to a cloud. One of these controls is a risk analysis that should be conducted at the start of the cloud integration projects. Before transitioning to the cloud, it is essential to carefully consider where the data is going to be processed. It is also important to keep situational awareness of the data locations because it might change in a short notice if cloud operator decides to make changes. Also, organizations should understand their subcontracting chain in cloud, because if something changes in that subcontracting chain that affects the location of the data it might have direct effects in organizations contracts and privacy statements. Organizations have privacy statements towards their customers that are required to name the locations where data is processed. As a controller, organizations define the countries where the data is going to be processed but it is essential to keep its privacy statements up to date. It is also important to know what happens when data is moved outside the on-premise systems, where organizations have direct control, to cloud under indirect control and how it affects the control of said data.

> "The smartest thing would be to conduct a profound risk analysis, but it's not always done now a days because business very often only sees the benefits and the functional side of the cloud"

> "It is fundamentally important to understand the division of responsibilities of what belongs to cloud operators' responsibilities and what belongs to our responsibilities"

> "It's not enough to have information security analysis say that everyone else is doing it like this"

The cloud changes the responsibilities over the system in many ways. But when it comes to GDPR the responsibility stays the same. The controller is always responsible for their processing even though the direct control of the data and processing might be shifted to cloud operator. So, in GDPR view it does not matter is the data being processed in on-premise information systems or if it is being processed in cloud operated by a third party. Three of the interviewees mentioned that the contract is in a key position when it comes to responsibilities and it should cover the how responsibility is allocated between the cloud operator and cloud customer. There is also a need to evaluate the maturity of the company when planning to adopt a cloud system. One interviewee stressed the importance of data life cycle management with cloud-based systems. When it come to teleoperator business there are different kind of data with different kind of requirements for its storing. Organizations need to be sure that the data that they are processing and storing in cloud meets the legal requirements that are concerning the data. Data life cycle management raises its importance when the processing and storing of the data is moved to a cloud, to organizations indirect control. Organizations need to ensure that the data is being processed by the contracts and legislations and its life cycle in managed appropriately and it is properly destroyed when it needs to be.

> "If we introduce some external cloud, we at least move the processing of personal data to that external actor, and it brings us legal responsibility of a controller"

> "When it comes to GDPR, the division of responsibility is unchanged. Controller is responsible for legal processing and that it goes by the privacy statement"

> "We need to evaluate how mature we are as an organization to move to a cloud in the view of this data life cycle where we command or handle things in the cloud that way that the data is there for the exactly the right amount of time, no less, no more"

The change in the division of responsibility does not only cause concerns. One interviewee mentioned that the change in the division of responsibility can also free employees time to focus in different things when some functionality is outsourced to the cloud.

> "The good thing about this is that we can get rid of these daily security patching and that we don't need to monitor does the logs fill our disc storage and these basic server maintenance tasks. That's the thing that changes, and in the view of division of responsibility we can raise the level of refinement of our jobs"

The cloud brings many good things compared to a traditional on-premise information system. We can assume that cloud is accessible at all times, and from anywhere in around the world. Cloud is also elastic and easily scalable shared resource. Cloud can also simplify the access control when it comes to external systems. One interviewee mentioned that it is quite a lot easier to get technical support for cloud system compared to a normal outsourced system that is in on-premise servers from system providers end because cloud is accessible from anywhere and anytime and temporary access is easy to grant. When thinking about IaaS and PaaS, cloud provides a fast starting-point for IT projects where basic infrastructure and possible the platform is already configured, and the development can be started right away. Cloud is also elastic and easily scalable so during a project it can be adjusted to fit the needs of the project. When thinking about SaaS the software is already done and accessible if access control is taken care of. Three interviewees mentioned that one of the clouds biggest benefit is scalability and the specialization of the provider to provide a good software. This can bring cost savings and higher level of innovation towards the product at the same time compared to a traditional inhouse software development.

> “Cloud brings a fast start if we think about a project, the infrastructure and capacity already exist, and it can be easily adjusted”

> “When these cloud companies are successful, for example this certain SaaS provider that provided customer relationship management software, customer relationship management is a big thing for the company as large as we are and it is quite similar to many other organizations customer relationship management thus the innovation ideas, needs, features that we are paying as a license for comes much more economically beneficial than that we would innovate these things by our self”

Cloud brings up many concerns when it comes to privacy and security. It might be harder to get cloud providers adjust their processes and operations to fit with legal requirements of customer organizations. What could be interpreted from the interviews was that it is better to be too cautious than take risks with cloud. There is also a legal concern with global clouds when data is transferred over country boarders and over legal jurisdiction. One interviewee also mentioned that the guidance organizations get from authorities is way too narrow which leaves all responsibility of trying to come up with sufficient policies and way of operation for organizations own consideration. All the interviewees agreed that security is a concern in cloud environments. When control over the system shifts towards the cloud operator also the control over security shifts. This needs to be considered when deciding a cloud provider and when drawing contracts. The responsibilities for information management, access control and logging mechanisms needs to be described accurately in contracts and the security policies and security safeguards needs to go through an audit.

> “Better to be cautious and understand what you’re doing. The worst excuse I’ve heard and still constantly hear is that: *everyone else is doing it like this*”

> “One concern is how the service is implemented, does the integrity last and does our information and data stay so that it is only our data. When we go to these basic information management questions, access control and such these needs to be delimited and described accurately and logging is also one thing that needs to be agreed on”

> “The control disappears. Your car is no longer in your own garage, but it is in neighbors’ garage where you might not even have access in. You neither cannot choose the leadership, employees or all the tools that are being developed. There is no longer same opportunity for control, so you just need to go by trust”

> “[…]and what makes this situation harder for commercial organizations is that the actual guidance that we get is like: *do a risk assessment, make good choices* and that’s all

> we get. There is no guidance paper, and there probably cannot even be, that says: *do not buy from that country*"

One challenge that three interviewees mentioned was getting sufficient understanding of what cloud changes. It takes time to get comprehensive understanding and keep up with all instructions and guidelines that is needed with cloud. Especially when moving to a global cloud operated by a third party it becomes challenging to get comprehensive understanding of cloud providers operations and processes. This challenge is even more concerning when organizations have legal regulations for privacy and security of personal data and obligation to indicate that the data processing is compliant with all legislations. It has been challenging to get needed information about the location of the data and where it is accessible which is extremely important with indicating that organization is GDPR compliant. There has also been challenges with integration of existing applications to cloud environment. Experience and knowledge how to do this can be bought from external consultants, but there is a need to get that sort of knowledge inhouse. developers and employees who oversee integration might not have sufficient understanding of this wholeness of the cloud which might end up in integration where every aspect of the system might not be configured and thought properly.

> "One simple challenge is that how to keep yourself up to date in this whole bustle. Where should you focus, there is cloud guidance overflowing the cloud but reading of all those guidance and the comprehensive understanding requires unbelievable amount of time"

> "There has been challenges about where the data is located where it is accessible from. Where it is physically located and where it is geologically accessible from and getting this kind of information can be challenging sometimes"

6.2 The Change in Control

The control over environment changes when moving from a traditional on-premise information system to a cloud. The change can be seen linear loss of a control over environment, where full control is in on-premise system and least control in SaaS based cloud service where the control from software to infrastructure is in cloud operators' hands. IaaS based systems have control over the platform and software but control over infrastructure is outsourced whereas in PaaS also the control over platform is outsourced to cloud operator.

The loss of control can be seen as one feature of the cloud because you cannot really outsource something while keeping full control over it and still expect benefit from it. One interviewee described this as feature that organizations needs to accept before moving some functionalities to operate in the cloud and it can be even a good thing. Two interviewees stated that when it comes to losing control in cloud it is important to understand how cloud environments works and how the monitoring in the cloud should be done. One interviewee stated that one of the first things that needs to be understood is that clouds exist in public internet thus the approach to the security of the cloud should be approached from the point off a view where it is as vulnerable as any system operating in the internet.

> "There are many different things going on, but sometimes I get a feeling that are we the ship Estonia that was sailing bow door open, because that's how it's always been done, and do not realize that a storm is coming that adjudicated. Because security cannot be 70% in place, just like the bow door of ship Estonia couldn't be less than 100% shut that it would have been secure. In this regard one must be very careful, and I feel like the message is received, but is it received in enough level, that's hard to say"

> "The vendor is then taking care of many basic things for us, and I think it's a good thing"

Three interviewees mentioned that contracts and mandatory contract attachments are the way to prepare for the loss of control. The contracts need to state who is responsible for what between the cloud operator and cloud customer. All the interviews revealed that careful planning is crucial part of preparing for the loss of control. Planning is closely related to the contracts. Two interviewees raised the continuity plans as a critical part of preparing for the loss of control. There needs to be plans for crisis situations and how to ensure the continuity of the services and how should the communication be managed in such situation but keeping a backup on-premise system does not seem reasonable. The risk analysis should always be done before going to the cloud. Business benefits should be measured and weighed against the risks the cloud brings in risk analysis. And after that the decision to move to a cloud could be made.

> "Contracts are a way to make cloud operator commit to certain principles what we are expected of and of course we do data privacy impact assessments for personal data"

> "I think it goes to that way of what is agreed on, what kind of control things are looked at, what kind of logs, what kind of reporting, what kind of security safeguards there over all is and how can we verify that the data integrity endures and that the data stays only as our data"

> "I don't believe we would give away some part of our functionalities of our infrastructure to the cloud service and then just to be sure maintain some sort of second system running if the cloud do not work, it does not work like that"

Although organizations lose control over the environment when they are moving processes and functionalities to the cloud, the opinion about that the responsibility still stays original was quite in-line. The legal responsibility of the controller stays within the organization even though they give control of processing the personal data to the cloud provider. Organizations need to maintain some control over the cloud provider if they decide to transfer the responsibility of personal data processing to the cloud provider. To maintain the control in these situations there needs to exist mechanisms for the control. If there are no such mechanisms, then organizations need build them before outsourcing the control. Organizations must be able to verify that they have done their part in deciding, testing and approval of the provider.

> "It is in the heart of our customer promise that customers information is safe and if we still do not have control of that information or if we do not have control mechanisms or even any idea, we only have a guess that maybe they are in a good safe over there, it's going completely wrong. Then you should not go. You can't say something and then do something else"
>
> "We are losing some basic functions, like command board kind of functions, which we sometimes have to build after wards, logging mechanisms and such. But it also frees our resources for more innovative things"
>
> "Our responsibility as a controller cannot be taken away"

Everyone strongly agreed that the loss in control does not mean loss in responsibility. The responsibility to organizations customers and legal responsibility stays the same when losing control over the environment when moving to a cloud. Although it was unanimous between the interviewees that organizations do not lose responsibility if they lose control in their systems, two of the interviewees discussed that there can be a slight change in responsibility. One interviewee specified that the change in responsibility can be seen shifting towards agreements and contracts.

> "In contracts it must be considered how to act in these security things, but it does not remove the fact that we are responsible for it. "
>
> "We can never stand that way that we are no longer the administrator of the information like we have been so far. In that sense it is just as important for us to make sure that all the things have been properly taken care of"

6.3 The Needed tangible Actions

Interviewees saw that cloud operator's credibility, reputation, reliability and references from other customers is important when choosing which cloud to start using. For all these features there needs to be evidence, documentation and standards for them to matter in choosing process. The size of the company was mentioned in a positive tone. It can be easier to choose a big and well-known cloud operator instead of a smaller one just to be safe. The bigger and more well-known organization that has been on the markets for a long time might have better support function and already well thought processes for security and privacy. The bigger provider may also be able to handle crisis situations better and will likely have more resources compared to smaller ones.

The geological location of the cloud company is also an important thing to consider because there can be some countries that organizations do not want their data to go. In addition to this there needs to be understanding in the whole subcontracting chain. When a cloud operator that is customer organizations cloud provider outsources some functionalities to subcontracting company and the customers data or access to said data moves to subcontractors' hands, it is important to know that especially when the access to the data moves over national or regulatory borders.

> "It shows that if someone has been on the markets for a long time, that they're doing at least something right"

> "It is behind a contract negotiation and a strict evaluation that what sort of services we decide. The challenge comes when we buy some services from a service provider that is behind another cloud service, which own solution includes another cloud service and then that our first subcontractor is responsible for their own solution, but it needs to be understood by us as well, that what kind of subcontracting chain or background machine there is behind the service provided by our subcontractor"

> "The risk classification that for some things there can be a little lighter provider if you think about that it is not that essential processing of data and this means about the size of the company as well because if something big happen and we get a compensation claim for example, so it does not help us at all that we drive some company down when they never had nothing to pay with"

Contract and contract negotiation is one the key parts in ensuring that cloud provider's security and privacy are in proper level. Contracts include many attachments where different thing about the service is agreed on. One of these attachments is Data Processing Agreement (DPA). DPA is required by the GDPR when controller and processor of the data agree on the processing of personal data. In addition to DPA the contracts need to include liability clauses where organizations agree on who is responsible for what.

Another key part is the verification of security and privacy that can be done by security auditing or monitoring and logging. One problem that came out in interviews was that most of the cloud providers do not let customers perform security audits, but cloud providers are audited at regular intervals by third party security companies and provide just the audit report to its' clients.

> "When our customer asks that how can you prove that your data is safe in GDPR point of a view, so we then have to prove that these are our subcontractors and here is our methods that we use to see that the subcontractors are and stays within the contracts and that they follow these our DPA's and that their security is up to date"

> "In big companies there are very sturdy contract practices and when we go to big actors then there are also big contract practices in the opposing side"

The level of control that organizations are used to in the traditional on-premises information systems is hardly possible or at least impractical in cloud environments. But organizations do not automatically lose all the control over the environment in the cloud. There are ways and mechanisms to preserve some control in the cloud. One important tool for preserving the control in the cloud is asset management systems. Asset management systems keep record of what systems and applications organization is using and who oversees them. One interviewee stated that asset management system is in the core of organizations technological heart where all needed information of all the systems can be found and it needs to be up to date.

> “This existence of asset management and its true accuracy and timeliness, it doesn’t matter is the application format in cloud or is it in on-premise system, it must always be kept up to date. This is the first control that needs to always exist”

Another important mechanism is Logging. Logging is also obligation from the GDPR and for teleoperators operating in Finland there is also obligations from the information society code. Logging in cloud may differ in solutions from how it is done in cloud environments. But the interviews revealed that it does not matter that logging is different in cloud, it must still be done as well and accurately than in on-premise systems. When discussing about the issue when the logging in the cloud is executed by the cloud operator, can organizations trust the log data provided by the cloud operator, all the interviewees agreed that there is no reason to doubt the reliability of the logs. When the logging is designed properly they cannot be tampered. One interviewee remarked that unlike in traditional on-premise information systems, logging need a throughout contemplating in cloud environments. It is important to design the logging mechanism so that it logs everything that is required by the law and other regulations, all that is needed for the monitoring of the service levels and such, but it should not log anything more than what is needed. One interviewee stated that when organizations use multiple cloud instances and cloud operators the possible amount of log data that they produce can overflow log management systems with useless data if the logging is poorly designed. Interviews also revealed that logs that cannot be processed by log management systems and security incident and even management systems (SIEM) become basically useless.

> “Now when our log mass is used for integrating event information, log, to our concentrated log management system and it is analyzed and some further actions, so this same should be possible when we it comes from the cloud”

> “If we ask from any cloud service or SaaS provides that we need GDPR loggin … The answer is always that everything is being logged. And when we go to check those logs there is just some cryptic references … But when we ask how does that when John looks Mary’s personal information can be seen from that hexa dump they just roll their eyes. People have also this kind of understanding that, when there is the

correct information among the log mass, so they think that now when we dump this to for example this [company’s SIEM system] an everything will be alright”

“The challenge is that how is the diversity (of the logs) modified as such that we see and we understand who did, and what, and whit what data, and when. It requires a change in our thinking”

7 DISCUSSION

This chapter represents the theoretical contributions that this research has to the field of information systems research. After this the limitations of the study is presented. Finally, the proposals and ideas for future studies of the topic are proposed.

7.1 Theoretical contributions

This chapter presents the contributions this study has to the research sections moving to cloud, preserving control and needed tangible actions and what is their significance to scientific community. The cloud adoption is ongoing process for many companies already and there is no indication that it will be slowing down any time soon. This study shows the challenges organizations face when they are moving to cloud computing environments and proposes some solutions or preliminary actions organizations should take before the adoption.

This study contributes to research field of cloud computing. Cloud computing is a widely studied during the last decade. Cloud computing was categorized as a disruptive technology with a profound effect to the whole IT sector (Dikaiakos, Katsaros, Mehra, Pallis, & Vakali, 2009; Botta, De Donato, Persico & Pescapé, 2016). Sultan (2012) categorized cloud computing as a disruptive technology for its' potential to destabilize existing information and communication technology markets. Cloud computing has indeed caused major changes to how organizations and individuals utilize information and communication technologies. Although cloud computing is a widely studied phenomenon among the academics, there are still many unsolved challenges with its' adoption and safe use. This research contributes to certain context of cloud computing. More specifically this study contributes to the changes moving to cloud causes for organizations, discovering what kind of changes happen to the control over the environment when moving to cloud and how can organizations preserve the control that is needed to ensure the security and privacy of the data they are responsible for. Because of relatively wide view of the study, privacy and security of the cloud is only previewed to support to get understanding of research topic and not examined more profoundly. This study was able to discover what kind of changes moving to cloud causes, what this means to control and responsibility over the environment, and what are the needed tangible actions to preserve the control over the security and privacy in order to stay compliant with the GDPR.

This research defined the concept of cloud computing, cloud computing deployment models and cloud computing service models. This study aimed to bring forth the security and privacy issues existing in cloud computing environments. The research thus contributed to the cloud computing research with combing cloud computing definitions from most cited academic journals from information systems research and information security research. Before this research there were only few studies focusing in cloud computing and the GDPR. This is due because the GDPR has only been active from May 2018 (GDPR, 2016). This study aimed to clarify the goals of the GDPR and what it requires from organizations operating in the EU or organizations that process, store or have access in the personal data of an EU citizen to stay compliant with the legislation. GDPR compliance is extremely important for all organizations but there are still many unanswered challenges with interpreting the legislation and how it functions with other regulations which this study tries to bring forth.

Cloud brings numerous benefits for organizations that can utilize it elaborately and carefully. To be able to utilize cloud properly there is a need for training and accurate guidelines within the company about how to do it. There are numerous professionals who have a profound understanding of how cloud works and how and for what it can be used, but in big organizations it is hard to ensure that every employee or team have the needed understanding of the risks cloud may bring along. Most cloud adoption cases case company has had, have been business oriented. In these kinds of projects, it is essential to ensure that there is understanding how to do it elaborately with paying attention to the required level of security. Although there are lots of guidelines and information available it seems to be too time consuming or too complex for all the employees to read through and understand. There is always the possibility to acquire this know-how from different consulting firms for individual projects, but it seems that there is a need to have it inhouse. Thus, case company should focus on having a mandatory training program for employees who are working in a position where moving to a cloud is considered. Training, understanding and comprehensive awareness of the cloud is essential. Awareness alone is still insufficient. There are already multiple security mechanisms and internal controls in disposition of case company that can be used in preserving the control over the system in the cloud. These controls refers to logging mechanisms, access control and asset management systems.

Logging was a familiar topic for the case company. The challenging part of it is when we move to a cloud environment. There is lots of regulations telling Finnish teleoperators what needs to be logged, but they do not not give the answer to how it should and can be done. In the literature of the research

Marty (2011) explained that logs need to answer the questions when, what, who and why. According to Marty (2011) the information needed to answer these questions are at least of timestamp, application, user, session ID, severity, reason and categorization. Logs should cover at least the mentioned information needs and depending on the situation some other entries as well. It seems to be quite common understanding that the more information is logged the better the logs are. But this is not always correct. Logs need to be moved to a concentrated log management system for analysis. The system case company is using is charging based in the amount of data that is transferred and processed in the system. Thus, the acquired licenses may be filled veritably fast if logging is not designed and planned properly. Organizations need to plan the logging carefully so that it is enough to answer the question when, what, who and why and other entries needed by the occasion and at the same time try to avoid logging useless information. Cloud bring another challenge for logging. This is the format of the log data. In the cases where case company has had challenges with logs in cloud environment, the cloud provider has provided the log data and the format of the data was not directly compatible with the log management system. The log data needs to be transferred to a format that can be uploaded into the log management system. This requires profound understanding of what information cloud provider provides and which log entry refers to which information. It seems like this sort of problem needs cooperation with the cloud provider to be solved. The responsibility for logging and formatting the log data should be positioned in contractual phase where the responsibilities are divided.

Access management and asset management are essential part of security in the cloud. Access management is essential in cloud especially because the GDPR classifies that access to the data cannot be given to actors operating outside the EU. In addition, access management in the cloud needs more attention because one of cloud characteristics is broad network access which means that the cloud is accessible from anywhere and anytime. When systems are accessible from outside to organization premises, organizations need to ensure that only the right people have access to the system. Asset management system is also an important tool to maintain control in in the cloud. Asset management systems are used to keep track of systems and interfaces organization is utilizing. It is important to keep central asset management system up to date when cloud enables rapid introduction of new services and software. It should be a requirement to apply the information of new services and software and the people responsible for them to the central asset management system before the introduction to maintain the up to date general view and control over organizations' IT.

Contract negotiations with the cloud operators were seen as quite strict and inflexible among the case company. Contracts and agreements are essential when dividing the responsibilities over the cloud. Contracts are used to ensure that cloud provider follows good practices which is a method to preserve the control in the cloud. Although if the contracts clearly divide the responsibilities the cloud consumer need to ensure that cloud operator follows good practices.

Guidelines that organizations get from official authorities seem too general. Guidelines explain what organizations must do to be compliant with regulation leaving out the explanation of how it should be done. This was seen as a challenge also in a big organization like the case company of the research, which means that the challenge is even harder to solve in smaller organizations with lesser resources. The official guidelines need to be updated. They need to cover the requirements of the regulations more comprehensible to avoid steering organization with scarcer resources to operate by decent or bad practices thus, endangering the privacy of data subject's personal data.

7.2 Limitations of the Study

To evaluate the credibility of this qualitative study, evaluating criteria proposed by Lincoln and Guba (1985) can be used. Lincoln and Guba proposed evaluating elements for qualitative study as transferability, credibility, conformability and dependability. Transferability of the study means how can the results of the study be utilized in other research objects (Lincoln & Guba, 1985). The results of the study could be used in a research of the same topic, although they cover only a small sector of the organizations that are facing challenges with the cloud and regulations. The credibility means the truthfulness of the study and research data (Lincoln & Guba, 1985). The interviews were conducted inside the company premises, but the anonymity of the interviewees and the trust between researcher and interviewees was emphasized. Thus, it is hardly plausible that the interviewees were motivated to answer untruthfully. Conformability means that the results of the study are indicted from the research data and not form the motivation or expectation of the researcher (Lincoln & Guba, 1985). This is a slight limitation for the accuracy of the research. There are not much earlier researches covering cloud and regulations with the same viewpoint and timeframe. Conformability of the research can be better evaluated when more research of the topic is published. Dependability of the research refers to the fidelity and consistency of data and data gathering. (Lincoln & Guba, 1985). The interviews followed predetermined questions and all the interviews were transcribed verbatim. Research data was then analyzed by coding it to three different themes. Research methods followed the instructions and good practices from earlier researches and research guidelines thus, the research could be repeated with similar results.

Some of limitations of the study became clear already when designing the research plan. Limitations that were observed during the research included research method, multidimensionality of the research problem, having only one case company, the number of interviewees, earlier studies of the same topic and the experience of the researcher. These limitations may affect the generalizability and the reliability of the study.

The number of the interviewees was relatively small. This is common for qualitative researches. Although the number of interviewees was small it is possible to gain useful and significant information even in small sampling when using qualitative methods (Hirsjärvi & Hurme, 2014). This study aimed to gain new information of the phenomenon by using semi-structured interviews in business organization. According to Hirsjärvi and Hurme (2014) the difference between groups and statistical generalization cannot be done when the number of interviewees is small. The interviewees were carefully selected among case company employees. The criteria in the selection was interviewee to work closely with cloud development, cloud related security, cloud privacy or legislations relating to cloud, privacy and security. This study tried to involve as many case company employees that fit the criteria to get best possible sampling for the purpose of the study. This study was able to generate new information even though the number of interviewees was small thus, we can assume that the small number of interviewees was not a significant limitation and we can assume that results of the study are generalizable and reliable with some limitation.

This study covered only one company that is operating in Finnish teleoperator field. This can be a limitation for the study if other teleoperators are doing cloud adoption differently. This study was done as an assignment for the case company and it could have been hard to get professionals from competing companies to participate the study. Although the study only covered one of three major teleoperators we can assume that the results would be similar enough if other companies were included. The challenge that the case company is facing is quite commonly known issues with cloud and regulations thus, we can assume that the results can be utilized in other organization that operate in the EU and are moving to cloud computing environments with some limitations.

There we some limitations regarding the academic literature of the study. There was very little earlier studies that are dealing with cloud and the GDPR. The GDPR is a new legislation and there are hardly any precedents of how the legislation has been applied. Because of this the citations in chapter covering legislation and GDPR are mostly from the official GDPR document. This is one limitation that might have some effect in the generalizability of the study if some sections of the GDPR will be applied differently than what was interpreted in this study.

The researcher is relatively inexperienced with interpreting legal documentations which can be seen as a limitation for the parts where the GDPR is covered. To mitigate this limitation, researcher consulted GDPR professionals from the case company with the writing style of the chapter and interpreting articles that are related with cloud. This is the first empirical study the researcher has conducted. That is why the competence of the researcher and the personal views of the topic can be seen as a limitation of the study. This limitation was mitigated by actively consulting more experienced researchers, mainly the supervising professor.

7.3 Suggestions for Future Research

This chapter presents interesting topics for the future research relating to the topic of the research. Proposed future research topics relates to the results and observations that were not profoundly investigated in this study or topics that are not yet mature to be studied.

Organizations are increasingly investing in cloud related projects although there are still many challenges and unanswered questions related to the privacy and security of the cloud and the legislation regulating it. There has been several studies covering cloud computing as technology and privacy and security of cloud computing. The topic that was least covered in academic literature was how cloud environments are regulated and what organizations need to do in order to be compliant with the regulations. Cloud is relatively new technology and it has been categorized as disruptive technology. The regulations regulating the cloud and how data should be handled in the cloud are even more recent.

An interesting topic for the future research that could not be covered in this study is to investigate what kind of incidents lead to GDPR sanctions and what is the root cause for those incidents. The GDPR is still relatively new legislation and there are no precedents that would show how GDPR is interpreted. This would help organizations to understand the legislation more fluently. The investigation to the root causes of the incidents would also help organization to ensure their security safeguards and privacy controls to avoid said incidents from happening that could lead to a substantial monetary sanction.

Another interesting topic for the future research regarding the regulations is how national regulations work together. One issue that was raised in the research was that there is significant problems when combining different national or multinational regulations. One of these issues is how can organizations still be GDPR compliant if they are using cloud services that are operating from the USA. Although the data is geologically located inside the EU there might still be some root access from USA. This is a problem when the regulations from the EU and the USA are reviewed. In the USA there are legislations that require citizens to hand over information they have access on without informing the data subject to ensure the national security. The problem emerges when that data happens to be covered by the GDPR which should prevent the data from EU citizen to be transferred outside the EU area. This is an especially interesting topic for future research.

Relating to the size of the sampling it would be interesting to expand the research to cover more organization. In the future studies it would be interesting to research does the challenges organizations are facing differ when moving to a different field or different country in EU. Other thing the wider sampling would allow to study is the maturity of the companies. The maturity of the companies relating to the cloud adoption would be an interesting topic. It could be studied by widening the sampling to cover multiple companies from multiple fields and multiple countries.

8 CONCLUSION

This is the concluding chapter of the Masters' Thesis. The research objective of the thesis was to investigate how cloud computing environments differ from traditional on-premise information systems and what actions organizations need to ensure the privacy, security and compliance with regulations when operating in the cloud. The topic is interesting and current due to the enactment of the GDPR and the popularity of cloud adoption among organizations.

This Masters' Thesis included literature review and empirical case study that was conducted using semi-structured interviews as a method. The literature review created a theoretical foundation for the empirical case study. Literature review is presented in chapters 2-5. Chapter 2 defines the cloud computing as a term and technology. Chapter 3 reviews the security and privacy in cloud computing environments. Chapter 4 clarifies the goals of the GDPR. Chapter 5 concludes the literature review and presents the research model for the empirical research. After the literature review, the empirical research is presented in chapters 6-8. Chapter 6 presents the research methodology. Chapter 7 presents the results of the study. Chapter 8 presents the discussion, which addresses the theoretical contributions of the study, limitations of the study and suggestions for interesting topics for the future researches. The final chapter, chapter 9, is the conclusion.

The conclusion of the study indicate that cloud differs from traditional on-premise information systems in many ways, but the existing practical security mechanisms can be utilized to ensure security and privacy in the cloud. This requires comprehensive understanding of the cloud among the organizations. The amount of control over the system decreases when moving to a cloud but this can be mitigated by contracts and agreements and proper security mechanisms. The official guidelines organizations get need to be updated to cover the tangible actions organizations need to take to ensure that following the regulations does not become too complex. As a precaution, organizations need to invest in improving the general awareness of cloud computing among the employees that will simplify the designing of the security mechanisms that are utilized with the cloud. Clouds are open to the internet and it requires a new kind of thinking when it comes to security. The awareness among organization can mitigate the security and privacy risk of sensitive data being stored and processed in cloud service or systems with insufficient security level.

REFERENCES

Al Morsy, M., Grundy, J., & Müller, I. (2016). An Analysis of the Cloud Computing Security Problem. arXiv preprint arXiv:1609.01107.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R. H., Konwinski, A., Lee, G., Patterson, D. A., Rabkin, A., Stoica, I. & Zaharia, M. (2010) A view of cloud computing. Communications of the ACM, 53(4).

Arora, A., Khanna, A., Rastogi, A., & Agarwal, A. (2017). Cloud security ecosystem for data security and privacy. In 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence (pp. 288-292). IEEE.

Botta, A., De Donato, W., Persico, V., & Pescapé, A. (2016). Integration of cloud computing and internet of things: a survey. Future generation computer systems, 56, 684-700.

Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J., & Brandic, I. (2009). Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation Computer Systems, 25(6), 599–616.

Carr, N. (2009). The big switch: Rewiring the world, from Edison to Google. WW Norton & Company.

Chen, D., & Zhao, H. (2012). Data Security and Privacy Protection Issues in Cloud Computing. In 2012 International Conference on Computer Science and Electronics Engineering (Vol. 1, pp. 647-651). IEEE.

Chow, R., Golle, P., Jakobsson, M., Shi, E., Staddon, J., Masuoka, R., & Molina, J. (2009). Controlling data in the cloud: outsourcing computation without outsourcing control. In Proceedings of the 2009 ACM workshop on Cloud computing security (pp. 85-90). ACM.

Cloud Security Alliance [CSA]. (2011). Security guidance for critical areas of focus in cloud computing V3.0 2011. San Francisco, California.

Coppolino, L., D'Antonio, S., Mazzeo, G., & Romano, L. (2017). Cloud security: Emerging threats and current solutions. Computers & Electrical Engineering, 59, 126-140.

Danezis, G., Domingo-Ferrer, J., Hansen, M., Hoepman, J.-H., Metayer, D. Le, Tirtea, R., & Schiffner, S. (2015). Privacy and Data Protection by Design - from policy to engineering. Cryptography and Security (Vol. abs/1501.0).

Data Protection Working Party. (2012). ARTICLE 29 – EU Data Protection Working Party. October, Article 29, 1–11.

Dawoud, W., Takouna, I., & Meinel, C. (2010). Infrastructure as a service security: Challenges and solutions. In 2010 the 7th International Conference on Informatics and Systems (INFOS) (pp. 1-8). IEEE.

Dikaiakos, M. D., Katsaros, D., Mehra, P., Pallis, G., & Vakali, A. (2009). Cloud computing: Distributed internet computing for IT and scientific research. IEEE Internet computing, 13(5), 10-13.

Duncan, B. (2018). Can EU General Data Protection Regulation Compliance be Achieved When Using Cloud Computing? Barcelona, Spain CLOUD COMPUTING 2018 Editors AutoManSec 4 CloudIoT-Autonomic Management and Security for Cloud and IoT View project Secure Data Engineering Lab View project.

Esage, A. (2018). Data breach in Salesforce. Retrieved 15.4.2019 from https://www.securitynewspaper.com/2018/08/04/data-breach-in-salesforce/.

European Comission. (2019). Data protection in the EU. Retrieved 5.4.2019 from https://ec.europa.eu/info/law/law-topic/data-protection/data-protection-eu_en

European Data Protection Supervisor (EDPS). (2018). Guidelines on the use of cloud computing services by the European institutions and bodies, (16 March 2018).

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008). Cloud Computing and Grid Computing 360-degree compared. Grid Computing Environments Workshop, GCE 2008, 1–10.

Gartner. (2009). Gartner highlights five attributes of cloud computing. Gartner Press; 2009. Releases June 23.

General Data Protection Regulation (GDPR). (2016). REGULATION (EU) 2016/679 OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).

Heiser, J., & Nicolett, M. (2008). Assessing the Security Risks of Cloud Computing. Gartner Research, (June), 1–6.

Hirsjärvi, S., & Hurme, H. (2014). Tutkimushaastattelu: Teemahaastattelun teoria ja käytäntö. Helsinki: Yliopistopaino.

Hirsjärvi, S., & Hurme, H. (2000). Tutkimushaastattelu: Teemahaastattelun teoria ja käytäntö. Helsinki: Yliopistopaino.

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. Qualitative health research, 15(9), 1277-1288.

Information Society Code. (2014). Information Society Code (917/2014). Translation from Finnish. Retrieved 21.9.2019 from https://www.finlex.fi/fi/laki/kaannokset/2014/en20140917.pdf.

Järvinen, P. (2012). On research methods. Opinpajan kirja.

Kandukuri, B. R., Ramakrishna, P. V., & Rakshit, A. (2009). Cloud security issues. SCC 2009 - 2009 IEEE International Conference on Services Computing, 517–520.

Lincoln, Y. S. & Guba, E. G. 1985. Naturalistic Inquiry. Beverly Hills, CA: Sage Publications.

Linthicum, D. S. (2016). Emerging hybrid cloud patterns. IEEE Cloud Computing, 3(1), 88-91.

Mather, T. (2009). Praise for Cloud Security and Privacy: an enterprise perspective on risks and compliance. " O'Reilly Media, Inc.".

Marty, M. (2011) Cloud Application Logging for Forensics. Proceedings of the 2011 ACM Symposium on Applied Computing - SAC '11.

Mell, P., & Grance, T. (2011). The NIST Definition of Cloud e National Institute of Standards and Technology.

Mogull, R., Arlen, J., Gilbert, F., Lane, A., Mortman, D., Peterson, G., & Rothman, M. (2017). Security Guidance for Critical Areas of Focus in Cloud Computing v4.0.

Okoli, C., & Schabram, K. (2010). Working Papers on Information Systems A Guide to Conducting a Systematic Literature Review of Information Systems Research. Working Papers on Information Systems, 10(26), 1–51.

Pearson, S., & Benameur, A. (2010). Privacy, security and trust issues arising from cloud computing. Proceedings - 2nd IEEE International Conference on Cloud Computing Technology and Science, CloudCom 2010, 693–702.

Ramachandra, G., Iftikhar, M., & Khan, F. A. (2017). A comprehensive survey on security in cloud computing. Procedia Computer Science, 110, 465-472.

Ruan, K., Carthy, J., Kechadi, T., & Baggili, I. (2013). Cloud forensics definitions and critical criteria for cloud forensic capability: An overview of survey results. Digital Investigation, 10(1), 34–43.

Safonov, V. O. (2016). Principles and Concepts of Cloud Computing. First Edition. Wiley-IEEE Computer Society Pr.

Salesforce.com (2018) Marketing Cloud April 23, 2018 Security Issue. Retrieved 15.4.2019 from https://help.salesforce.com/articleView?id=000313931&language=en_US &type=1&mode=1.

Samarati, P., di Vimercati, S. D. C., Murugesan, S., & Bojanova, I. (2016). Cloud security: Issues and concerns. Encyclopedia on cloud computing, 1-14.

Schwartz, M. J. & Ross, R. (2018). Salesforce Security Alert: API Error Exposed Marketing Data. Retrieved 15.4.2019, from https://www.bankinfosecurity.com/salesforce-security-alert-api-error-exposed-marketing-data-a-11278.

Singh, A., & Chatterjee, K. (2017). Cloud security issues and challenges: A survey. Journal of Network and Computer Applications, *79*, 88-115.

SLA Management Team. (2004). SLA Management Handbook – Volume 4: Enterprise Perspective. TMF document reference GB917, Version 2.0, Volume 4. The Open Group, October 2004.

Soares, J., Goncalves, C., Parreira, B., Tavares, P., Carapinha, J., Barraca, J. P., … Sargento, S. (2015). Toward a telco cloud environment for service functions. IEEE Communications Magazine, 53(2), 98–106.

Subashini, S., & Kavitha, V. (2010). A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications, 34, 1–11.

Sultan, N. (2013). Knowledge management in the age of cloud computing and Web 2.0: Experiencing the power of disruptive innovations. International journal of information management, 33(1), 160-165.

Takabi, H., Joshi, J. B. D., & Ahn, G.-J. (2011). Security and Privacy Challenges in Cloud Computing Environments. IEEE Security & Privacy, 8(6), 24-31.

The Office of the Data Protection Ombudsman. (2019). Rights of the data subject. Retrieved 14.5.2019 from https://tietosuoja.fi/en/rights-of-the-data-subject.

Tolsma, A. (2019). GDPR and the impact on cloud computing. Retrieved 5.4.2019 from https://www2.deloitte.com/nl/nl/pages/risk/articles/cyber-security-privacy-gdpr-update-the-impact-on-cloud-computing.html.

Vaquero, L. M., Rodero-Merino, L., Caceres, J., & Lindner, M. (n.d.). A Break in the Clouds: Towards a Cloud Definition. ACM SIGCOMM Computer Communication Review, 39(1), 50-55

Varghese, B., & Buyya, R. (2018). Next generation cloud computing: New trends and research directions. Future Generation Computer Systems, 79, 849–861.

Yin, R. K. 2003. Case Study Research: Desingn and Methods. Third Edition. California: Sage Publications.

Yu, X., & Wen, Q. (2010). A view about cloud data security from data life cycle. 2010 International Conference on Computational Intelligence and Software Engineering, CiSE 2010, (4072020), 1–4.

Zawoad, S., Dutta, A. K., & Hasan, R. (2013). SecLaaS: Secure Logging-as-a-Service for Cloud Forensics

APPENDIX 1 INTERVIEW FRAME

BACKGROUND INFORMATION

Who are you and what is your position in Case Company?

How long have you been working in your current position?

How long have you been working for Case Company?

How long have you been in the same field?

THEME 1 – TRANSITIONING TO CLOUD

1. How is the cloud at issue in your job?
2. What does moving to cloud change?
3. What should be done before moving to cloud?
4. How does the cloud change the responsibilities?
5. What is good or what kind of opportunities can be found in cloud?
6. What are the bad sides or what kind of challenges can be found in cloud?
7. Have you faced difficulties or challenges with the cloud in your work?

THEME 2 – THE CHANGE IN CONTROL

8. How should we relate to losing control in the cloud?
9. How can we prepare for losing control in the cloud?
10. What is the real change that losing control in the cloud causes?
11. Does losing control also mean losing responsibility in some way?

THEME 3 – THE NEEDED TANGIBLE ACTIONS

12. What is important when choosing a cloud operator?
13. How can we ensure that the cloud operator is trustworthiness / sufficiency?
14. What should be taken into account in contracts with cloud operator?
15. How can control be preserved in cloud computing environments?
16. How should logging be done in cloud computing environments?