Cloud computing data storage: Usage, security and privacy issues for individual users

Dean O’Gorman

A dissertation submitted to the University of Dublin in partial fulfilment of the requirements
for the degree of MSc in Management of Information Systems

1st September, 2015

Declaration

I declare that the work described in this dissertation is, except where otherwise stated, entirely my own work, and has not been submitted as an exercise for a degree at this or any other university. I further declare that this research has been carried out in full compliance with the ethical research requirements of the School of Computer Science and Statistics.

Signed: ________________________
Dean O’Gorman
31st August 2015

Permission to lend and/or copy

I agree that the School of Computer Science and Statistics, Trinity College may lend or copy this dissertation upon request.

Signed: ________________________
Dean O'Gorman
31st August 2015

Acknowledgements

I would like to thank my wife Emer and our two young boys Ruairí and Fionn. The most difficult part about writing this dissertation was the time that was spent away from my family.

I would also like to thank my supervisor Susan Leavy for her advice and constant support throughout this work, especially when I had developed “The Fear”.

Finally, I would like to thank the lecturers and staff at Trinity College Dublin and also my classmates. This has been a highly engaging and fun two year experience.

Abstract

Internet usage continues to grow globally and as part of this, the take-up of personal cloud data storage solutions has also been increasing. This has been assisted by the continued rise in internet access speeds as well as an increase in the numbers of vendors and the resulting completion between them. These services can be very attractive to the end user; it becomes possible to access your data from multiple locations, on various devices and easily share data and collaborate with others.

This research aims to explore the usage of these cloud data storage solutions along with the perceived security and privacy risks that are associated with using them. Are users of these solutions aware of these risks and are they taking necessary steps to protect themselves against them? The location of where the data is stored is an important factor in this situation as there may be contrasting privacy laws between different countries.

Semi-structured interviews and observation sessions were used to gather data from a group of 18 participants. Both qualitative and quantitative methods were used to analyse the data from these sessions.

The findings outlined a number of key points. There was a significant uptake of cloud data storage solutions amongst the group of participants with many having multiple accounts with different providers, although only one participant paid for a cloud data service. A minority were unaware of the potential risks with the technology. However, it was discovered that in general, the majority did not care about these risks but still used the technologies nonetheless, although most chose not to store sensitive or confidential data with their cloud data storage provider.

Table of contents

List of Tables and Diagrams

Abbreviations

DoS	Denial-of-Service
EFSS	Enterprise File Synchronisation and Sharing
EMEA	Europe, the Middle East and Africa
EU	European Union
EULA	End-user License Agreement
FAQ	Frequently asked question
HIPPA	Health and Human Services Health Insurance Portability and Accountability Act
IaaS	Infrastructure as a Service
IT	Information technology
NIST	National Institute of Standards and Technology
PaaS	Platform as a Service
PIN	Personal Identification Number
SaaS	Software as a Service
SLA	Service Level Agreement
SME	Small and Medium Enterprises
SOX	Sarbanes-Oxley
SQL	Structured Query Language
ToS	Terms of Service
USA	United States of America

1. Introduction

1.1 Background and Context

The internet has become intertwined in our everyday lives and the number of people using its services continues to grow at a rapid rate. According to the latest Global Internet Report from the Internet Society (2014), there are now over three billion internet users worldwide, up from one billion users in 2005. People use the internet for different purposes such as interacting via social media, communicating using email and online meeting tools, shopping on the various E-commerce websites and storing data online. Cloud computing has become an important enabling factor for these services. It is arguably one of the most important technological advances over the last decade and it has the potential to revolutionise the delivery of IT services further (Brynjolfsson and Jordan, 2010; Marston et al., 2010).

Cloud data storage enables users to store their data online instead of using traditional options such as local hard drives and flash drives. There are many benefits to this practise including the ability to access your data from virtually any location on many devices such as computers, tablets and smartphones. It is also a relatively simple task to share online data with other individuals or groups of people by sending a link in order to allow them to gain access to a shared folder or to download a single file.

Uptake of these cloud data services has been significant and this has been helped in no small part by the number of providers that are competing to convince people to sign up to use their particular service. Such competition between vendors has resulted in the amount of free storage that individuals can gain access to continually increase, while there has been a constant reduction in the cost of much higher capacity paid-for data storage.

But, there are risks associated with these solutions and because data is stored remotely with the service provider, it can be argued that the cloud is intrinsically insecure (Ren, Wang and Wang, 2012). Users of these cloud data storage solutions should be aware of any such risks and should be in a position to protect themselves against any potential loss of data that might occur. The issue of privacy also needs to be considered. For example, can the service provider or government where the data server is hosted access your online data?

1.2 Rationale of the Study

This study aims to investigate if individual users of cloud data storage solutions are aware of the potential risks of using such services. Users may be unaware of such risks or may simply be unconcerned. This may be down to the level of technical competence of the actual user or perhaps the data that is being stored online is not confidential in nature.

1.3 The Research Question

The objective of this research is to examine levels of awareness of security and privacy risks among users of computers and mobile devices. This objective forms the basis for the following research questions:

1. How do users of computers and mobile devices currently utilise cloud data storage solutions and do they pay for such solutions?

2. What is the perception among users of potential security and privacy risks associated with using cloud data storage solutions?

1.4 Importance of this Study

Because this research is investigating the potential security and privacy risks that are associated with using cloud data solutions, it is hoped that it would be of benefit to users of such services. These users could use the research to gain a deeper understanding of the underlying issues relating to cloud data storage security and privacy risks in order to help mitigate against them.

Additionally, because the underlying technologies do not differ widely for businesses users, small to medium enterprises (SMEs) that are considering utilising a cloud data storage solution in the workplace for their staff may find this research useful.

1.5 Scope and Boundaries

The focus of this study is on personal cloud data storage solutions and not Enterprise File Synchronisation and Sharing (EFSS) solutions that are only typically used in large organisations.

To answer the research question, data was gathered by means of face-to-face interviews along with short observation sessions with each participant. A total of 18 individuals participated in the study and these were of varying age and technical background.

Participants were asked a number of questions regarding their technical background, how they currently utilise cloud data storage solutions and finally, issues pertaining to the security and privacy of using these solutions. In order to gain knowledge from a cross section of age groups, a number of individuals from each age range were approached to participate and this resulted in a smaller pool of participants that took part in the study than if no selection criteria were used. This relatively small sample size could possibly result in a narrow set of findings, depending on the participants.

1.6 Chapter Roadmap

The structure of this dissertation is divided into the following chapters:

Chapter 1: This introduction chapter gives an overview of the research questions as well as the background of, and the rationale for, the study. This chapter also defines the scope and boundary of this research.

Chapter 2: The literature review chapter will examine the literature relating to cloud computing technologies as well as the benefits and possible drawbacks to users of the services. This section will also investigate existing literature focusing on cloud data storage solutions and why they are becoming very popular with individuals. Finally, the security and privacy issues relating to the usage of such solutions is discussed

Chapter 3: The methodology and fieldwork chapter will describe the various methodological approaches available to a researcher to complete a study. It will outline why particular methodologies were chosen and how the research itself was conducted. The techniques used for data collection in this study and the interview process will also be explained.

Chapter 4: The findings and analysis chapter will present the findings of the research and analysis of these findings. These findings will determine if users of cloud data storage solutions are aware of potential security and privacy risks and if they are concerned by them.

Chapter 5: The conclusions and future work chapter will show if the research has answered the research query or if any new or interesting results were found as part of the research. This chapter also contains important observations and possible future research that may arise because of this work.

2. Literature Review

2.1 Introduction

The objective of this literature review is to examine and critically analyse the body of published work which exists in the field of cloud computing, with particular focus on usage, security and privacy pertaining to cloud data storage. Firstly, it is necessary to briefly define what cloud computing is because of the amount of confusion that has arisen due to contrasting definitions on what exactly this term refers to. Interpretations differ in various research papers (Grossman, 2009; Wang et al., 2008). Indeed, terms such as "cloud computing", "on-demand computing", "software as a service" and "the internet as a platform" have become somewhat interchangeable even though they have very different meanings (Hayes, 2008).

Although a deep explanation of cloud computing is out of the scope of this paper, by properly defining the technology, this will allow a reader without strong background knowledge in the area to grasp the general principals of cloud computing along with its potential benefits and drawbacks. This will allow for a more in-depth discussion on security and privacy issues that individuals may face when using cloud computing solutions, particularly cloud data storage solutions.

Cloud computing can have major benefits to individuals and can be a very attractive solution to users of the service to store their data online as it can offer many advantages. The next section in this chapter will focus on these advantages along with associated disadvantages and will also investigate different cloud data storage offerings and how they vary from each other. The majority of academic publications and industry articles found during this research focus on enterprise cloud data storage solutions and not personal use options, so the differences between individual and enterprise solutions will be discussed as well as important aspects such as the location of where the service provider stores the data as this is directly relevant to privacy issues.

Finally, cloud computing security and privacy will be discussed along with examples of how measures can be implemented to ensure that data is properly secured as well as auditing options that allows for the reporting and the upkeep of security. Because of recent high profile issues that highlight cloud data security breaches such as the leaking of celebrity photographs from their Apple iCloud accounts, features such as encryption, location of data and possible reasons loss of data will be examined.

2.2 Overview of Cloud Computing

References to cloud computing can be found as far back as the 1960's where visionaries such as J.C.R. Licklider and John McCarthy saw a future where individuals could access data and programs at any site, from anywhere (Kaufman, 2009; Mohamed, 2009). But the term was not used officially until George Favaloro published a business plan in Compaq in 1996 (Regalado, 2011). Virtualisation is regarded as being the main enabling technology for cloud computing and shares many characteristics such as the ability to create a scalable system of independent machines, high agility and lower costs (Hamdaqa and Tahvildari, 2012). Salesforce.com, a well-known cloud computing only company, introduced one of the first functional cloud computing applications in 1999 and established the notion of delivering enterprise services through a website (Kaufman, 2009). Cloud computing is a disruptive technology which now has implications for the entire IT sector and beyond (Subashini and Kavitha, 2011).

The cloud computing sector continues to grow and more than half of U.S businesses are now using cloud computing in one form or another (Cohen, 2013). Forrester has predicted that the size of the cloud computing market will be more than $240 billion by 2020 (Forrester, 2011).

At an individual level, personal cloud services such as Gmail and Dropbox have become highly accessible. The only requirement is a computer, smartphone or tablet along with an internet connection to get access to a myriad of cloud computing services including email, data storage and hosted machines, many of which are available at no charge to the user.

The utilisation of cloud computing allows for global access. The locations that the user is accessing the service from is unimportant because the service is hosted remotely by the service provider. Once online, the user can access the service from anywhere, at any time and in many cases, from virtually any device. According to Lin and Chen (2012), cloud services such as webmail, Flickr and YouTube have been widely used by individuals for some time and these can be accessed by many devices, not just a desktop computer or laptop. Gartner identified both "Computing Everywhere" as well as "Cloud/Client Computing" as two of the top strategic technology trends for 2015 (Gartner, 2014a).

Ease of use is also a very important factor to the adoption of a cloud computing service. As can be seen in figure 2.1 below from Pew Research, 51% of the people surveyed cited ease of use and convenience as the main reason for using cloud applications and data storage. In all areas of cloud computing there is a lot of competition between vendors and therefore a significant choice of products for a user to select from. If a product is not easy to use then the indivudal will simply chose another available option.

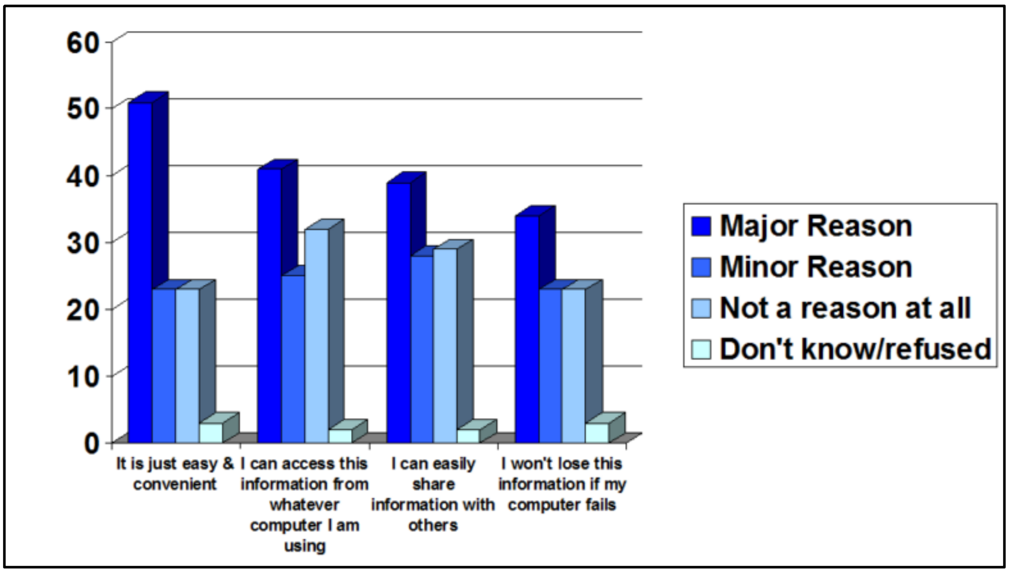


Figure 2.1: Why people use Cloud applications (source: Pew Research Centre, 2008)

Additionally, by leveraging cloud computing technologies, IT professionals can help further their career because it gives them access to additional technologies that they would not otherwise have access to, thereby allowing them to acquire additional skills (Armbrust et al., 2010).

2.3 Service models

Cloud computing solutions that are available to the individual vary widely from email to hosted virtual servers and in order to discuss these offerings effectively, it is necessary to first describe the technologies that cloud computing is built upon. The service models that cloud systems are based upon are particularly important as they define the scope of that solution.

Cloud Computing is often described as a stack and in effect it consists of a number of services built on top of one another and this is covered by the term “Cloud”. The generally accepted definition of Cloud Computing comes from the US-based National Institute of Standards and Technology (NIST). The NIST definition essentially states that cloud computing is a model for enabling 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 (Rackspace, 2013; NIST 2012).

Figure 2.2 below is a graphical representation of the three cloud computing service models and their relevant target audience. All of the listed examples are available to the individual user.

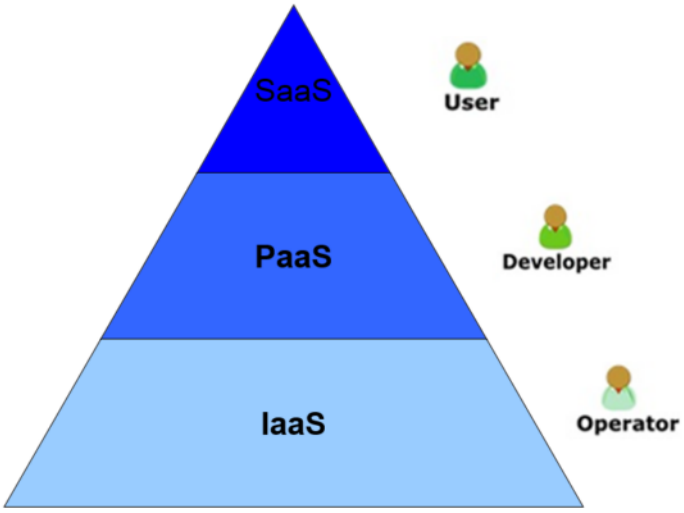


Figure 2.2: Cloud computing service models

2.3.1 Infrastructure as a Service (IaaS)

Infrastructure as a Service (IaaS) is a method of delivering cloud computing infrastructure such as storage, servers, network and operating systems as an on-demand service over the internet. Instead of going through the process of purchasing servers, datacentre space, network equipment or software outright, customers instead use these resources as a fully outsourced service on demand (Kepes, 2011; Shi et al., 2008). Additionally, the IaaS provider is responsible for the upkeep and maintenance of the underlying infrastructure while the customer is responsible for the operating systems running on the IaaS environment (Lin and Chen, 2012). The top IaaS providers on the market today are Amazon Web Services, Microsoft Azure and Rackspace.

2.3.2 Platform as a Service (PaaS)

Platform as a Service (PaaS) can be defined as a computing platform including operating system, programming environment, database connector and web server that is typically aimed for use by software developers and allows for the development, testing and deployment of web applications quickly and easily, without the cost and complexity of managing the underlying software and infrastructure (Rackspace, 2013). The customer is only responsible for the upkeep of the actual applications that are running in the PaaS environment. Leading examples of PaaS market leaders are Google App Engine, Heroku and Engine Yard (Sullivan, 2014).

2.3.3 Software as a Service (SaaS)

Software as a Service (SaaS) allows for the delivery of an application as a service across the internet, usually (but not always) to be run in a web browser. The benefit for the customer is that the provider is responsible for the entire delivery of the application as a service and this includes all hardware, operating system, database, storage, networking and security. In addition, the service provider is also responsible for the ongoing updating and patching of the application, allowing for reduced administration and support at the local level (NIST, 2012). Top SaaS applications based on revenue are Gmail, LinkedIn and Facebook.

2.4 Deployment Models

There are four major cloud computing deployment models (Zhang et al., 2010) and are described in the following subsections. These are related to the usage of cloud computing and each deployment model has its own advantages and drawbacks.

2.4.1 Public cloud

A public cloud is wholly hosed by a third party service provider and is open for public use. Individuals and professional users share the cloud services on a multi-tenancy basis and all resources are connected to over the internet. This is the cheapest and most flexible model available for use, however there may be security issues because the cloud services are made available over a public network. When a cloud service is made available in a pay-as-you-go manner it is referred to as a public cloud (Armbrust et. al, 2010).

2.4.2 Private Cloud

A private cloud infrastructure is operated solely for a single agency or organisation. The services can still be hosted by a third party service provider or internally by the organisation itself. This gives the organisation a much higher level of control over the security and management of the cloud, however there is a significant increase in costs associated with this deployment model.

2.4.3 Community Cloud

A community cloud refers to the sharing of cloud infrastructure or services between several groups that share a common set of concerns such as security, compliance or jurisdiction considerations. As with the private cloud deployment model, these services can be hosted externally or internally, however the cost is shared between the organisations concerned.

2.4.4 Hybrid Cloud

A hybrid cloud comprises of two or more clouds (public, private or community) from different service providers (Bittman, 2012). There are various use cases for hybrid clouds, one of the most popular being the ability to extend the capacity of a private cloud service by aggregation with another cloud service, possibly public in this scenario. This type of leverage allows for increased capacity but is flexible so has the ability to only be used during busy periods.

It is important to highlight the above deployment models and how they affect the individual user. Only larger organisations will have the resources to implement a cloud model other than a public cloud, and this is applicable to the personal user also.

2.5 Characteristics of Cloud computing

There are five essential characteristics (NIST, 2012) of cloud computing and these are outlined below. It is possible to disregard the confusion surrounding the definition of cloud computing so long as we understand the actual characteristics of the technology itself (Gong et al., 2010). These characteristics can help individuals to better understand the cloud computing solutions that they are using compared to traditional applications and solutions.

2.5.1 On-demand self service

A user can provision computing resources such as server time and networking as required directly, without the need for human interaction with the service provider (NIST, 2012; Sakr et. al., 2011).

2.5.2 Broad network access

Resources are available over the internet and are accessible with standard equipment such as desktop computers, laptops, tablets and smartphones (NIST, 2012; Sakr et. al., 2011).

2.5.3 Resource pooling

The cloud service providers' resources are pooled in order to deliver services to multiple consumers based on a multi-tenant setup. These resources are then dynamically provided to the customer based on their changing requirements (NIST, 2012; Sakr et. al., 2011).

2.5.4 Rapid elasticity

Resources can be elastically delivered and released depending on the demand from the customer. From the customers' point of view, these services must be available for appropriation in any quantity at any time (NIST, 2012; Sakr et. al., 2011).

2.5.5 Measured service

Cloud systems automatically control and optimise the resources used by the customer by leveraging a metering function at some level of abstraction appropriate to the type of service (e.g., storage, processing, bandwidth, etc.). Resource usage can be monitored, controlled, and reported, providing transparency for both the provider and consumer of the utilised service and allowing for a pay-per-use service (NIST, 2012; Sakr et. al., 2011). There is an increased perception that because of this pay-per-use factor, that the cloud model will elevate computing to the 5th major utility after water, electricity, gas and telephony (Buyya et al., 2009).

2.6 Benefits of Cloud computing

Why should people adopt cloud computing technologies? There are a number of benefits associated with its usage and the following is a summary of such benefits (NIST 2012; Armbrust et al., 2010; Marston et al., 2011) –

2.6.1 Cost efficiency

Cloud computing services are typically pay as you go so there is no need for capital expenditure at the outset and this allows for the conversion of capital expenses to operating expenses. Additionally, licensing costs and on-going administrative tasks such as upgrades are handled by the provider.

2.6.2 Improved accessibility

As long as the end user has an internet connection, they can access applications and data anytime on demand from anywhere, allowing for easier access to information and a more flexible way of working (Buyya et al., 2009). This is not only good for individual users but for businesses as well as it allows their employees to work more flexibly. Additionally, applications to access cloud resources can be run from smartphones and tablets in addition to laptops and desktops, resulting in further flexibility.

2.6.3 Faster time to value

Cloud computing gives the advantage of quick deployment compared with traditional computing systems. Because the resource can be provisioned directly online, there is no need to procure and configure the hardware and install the software for a comparable in-house solution.

2.6.4 Backup and recovery

Since all of the data is stored in the cloud by the service provider, backing it up and restoring the same is relatively much easier than storing the same on a physical device. It must be clearly stated as part of the Service Level Agreement (SLA) between the vendor and the client that this is the case as not all vendors offer this service by default and if this is not the case, the task of backing up the data remains the responsibility of the customer. Furthermore, there is no longer a need for complex offsite disaster recovery plans, again since the data is stored with the service provider and would not be under threat as it would with a local data server. This is easier for a business to negotiate with the service provider but the individual user must be aware of the service providers' functionality and terms and conditions. For example, if some stored data is accidentally deleted, is it recoverable and if so, for how long? Is version history an option as standard?

2.6.5 Improved flexibility

By utilising cloud computing services, this gives the freedom to access seemingly unlimited computing power and resources when, and only when, it is required. 65% of respondents to an Information Week survey said the ability to quickly meet business demands was an important reason to move to cloud computing (Biddick, 2008).

2.7 Challenges of Cloud computing

Along with the benefits that are associated with cloud computing, individuals must be aware of potential drawbacks. The following is a brief discussion of some of the challenges facing the adoption of cloud computing. Some of these issues will be discussed in more detail in section 2.9.

2.7.1 Availability

One of the major disadvantages of cloud computing is the dependency on the provider. An outage by the service provider could have a huge impact on all of its users. For example, over the 48 hour period of Christmas Eve and Christmas day in 2012, Dropbox had a total downtime of 106 minutes (Macpherson, 2013) which caused disruption to its hundreds of millions of users, preventing all of them from accessing their online data.

A traditional, robust IT infrastructure calls for a setup without a single point of failure, yet the adoption of a cloud computing service provided by a single provider is in itself a single point of failure. Does the provider have a single data centre? If so, it is therefore venerable to a number of possible disasters, natural or otherwise, that could render it inoperable and take its services offline. Even if the service provider has multiple data centres in different countries, it may share the same underlying software which may be vulnerable to failure or attack. Finally, the service provider itself may go out of business and in that situation, what would happen to an individual’s data? Armbrust et al. (2010) argue that the only way to circumvent this possible single point of failure is to employ multiple cloud computing providers, but this would add significant cost to the cloud solution and would be overly complex to the individual user.

2.7.2 Security

Security remains one of the most often cited reasons for not adopting cloud computing services (Subashini and Kavitha, 2011). Individual users may simply be fearful when deciding not to use a cloud option and instead opting for a traditional solution that is local on their computer. Without appropriate security and privacy solutions specifically designed for cloud usage, there is a risk that this potentially revolutionising computing technology could become a failure (Takabi, Joshi and Ahn, 2010). There are physical issues that must be addressed by the cloud service provider. They must be responsible for the physical access to their data centres and control who has access and at what time. Additionally, it is vital that any hardware that is to be disposed of is done so correctly such as a hard disk is wiped clean to ensure that no data can be retrieved from it (Quick and

Choo, 2013a). The provider must also guard against denial-of-service (DoS) attacks which could render all of its services inaccessible to its customers.

Encryption is necessary to encapsulate users' data from each other but also from the service provider – under no circumstances should the provider have access to a customer's data. All data in transit should also be encrypted between the end user and the cloud service to ensure that the conversation remains secure. Encryption technology is normally built into the tools for accessing cloud computing services, so this functionality should be transparent to the end user but it is important that the end user ensures that all traffic is encrypted. This information would be found in the terms of service from the provider. Encryption in itself does not prevent the data from being intercepted, but it denies the actual content of the message to the person that intercepted it.

2.7.3 Dependency and vendor lock-in

One possible disadvantage of cloud computing when purchasing a service is vendor lock-in whereby it can be difficult and sometimes almost impossible to move to a new provider. If an individual wishes to switch to a new service provider, it could be very inconvenient to transfer large amounts of data from the old provider and in some situations, it may simply not possible to do so (Lin and Chen, 2012). For example, if an individual was using the Dropbox paid service, this allows for up to 1TB of data to be stored with the service provider. If for some reason the user then became unhappy with the service being provided and wished to switch provider, the amount of data that is stored with the original provider may become an inhibiting factor preventing such a switch.

2.7.4 Support

Compared with most packaged or custom-built software, cloud-based services do not always have the highest level of customer care support available. It may be difficult to get in contact with the vendor on the phone or by e-mail, and they often do not provide illustrated user manuals. Instead, they ask their customers to rely on FAQ pages and online community support, which may not be always easily searchable which and can be time consuming to find a resolution to an issue.

For example, for a free user of Dropbox data services do not have access to dedicated support, but Dropbox Pro paid-for service users do have access to priority support which will result in much faster response times.

2.8 Cloud Data storage

Cloud data storage, also known as a file hosting service, allows users to store their data remotely with the service provider instead of locally on their hard drive. The result of this is that the data can then be accessed over the internet from any location from devices such as desktop computers, laptops, tablets and smartphones which offers increased flexibility and accessibility (Quick and Choo, 2013a). The user is required to provide a username and password to access their data and it is then possible to share files publicly or keep them password-protected (Google Drive Support website, 2015).

Another important aspect of cloud storage is the ability for an end user to synchronise their data to a local folder on their computer, which can hold a full copy of the data from the remote server (Drago et al., 2012). Any changes to documents are automatically synchronised between the local folder and the cloud service provider. The user is not limited to a single device for synchronisation; this can be configured on multiple apparatuses and the result is that an up-to-date version of the data appears to be in the same folder regardless of which device is used to view it. (Geel, 2013).

Mobile devices such as smartphones also come with automatic backup-to-cloud options and the result is that when a configured device is connected to the internet, either directly or through a computer, it is updated to the cloud service provider. For example, data such as photographs and contacts can be automatically backed up from an iPhone to Apple's iCloud service (Apple, 2015). Users of Samsung or HTC Android smartphones can be configured to automatically backup to their Dropbox account. It should be noted though that a user may not be actually aware that they are using a cloud data storage account to handle the backing up of their device. Signing up for this service may be a single step when configuring their smartphone or tablet for the first time.

Dropbox, a leading company for individual cloud storage, now claims to have over 300 million users worldwide (Hong, 2014), an increase of over 100 million users since November 2013. Cloud data storage continues to gain popularity. Figure 2.3 depicts the result of research carried out by Business Insider in 2014 showing that 29% of computer users currently utilise cloud storage for their data. At the launch of Windows 8 in 2012, Windows Division President Steven Sinofsky stated that SkyDrive (now known as OneDrive) had over 200 million customers storing data with their free SkyDrive product which consists of 11 billion photos, 550 million documents, storing over 14 PB of data and every month two petabytes of data are added (Microsoft, 2012).

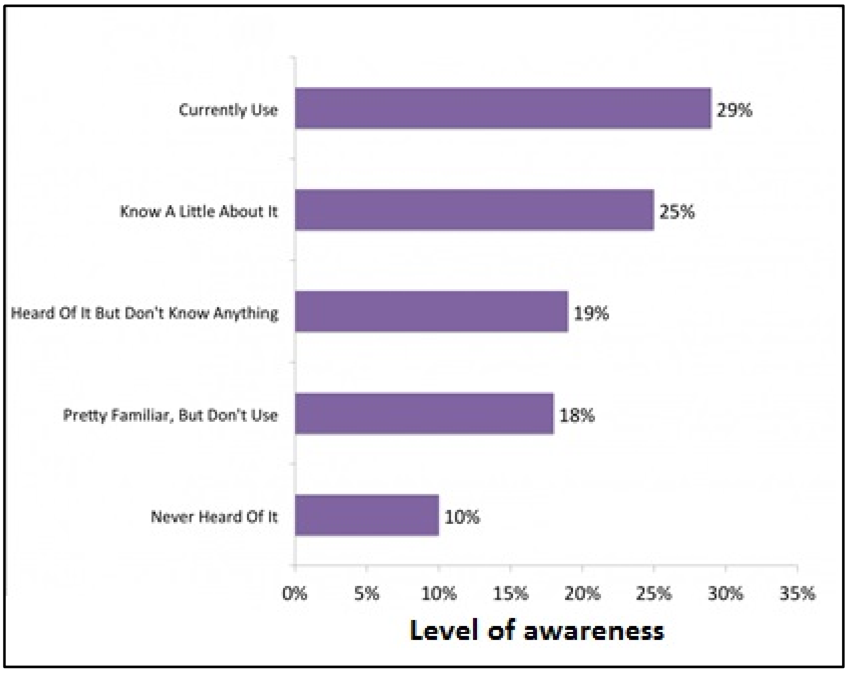


Figure 2.3: Cloud storage awareness (source: Business Insider, 2014)

The number of hosting companies is also continues to grow and competition is rife in the sector resulting in an increasing number of vendors for the indivudal to select from. The Gartner magic quadrant for Enterprise File Synchronisation and Sharing published in 2014 lists a total of 19 data hosting companies alone (Gartner 2014b). This is focused on the enterprise only and does not even include well known and used personal cloud data services such as Apple iCloud and Amazon.

There is no reason to believe that this growth will subside, particularly as the amount of data that is stored in the cloud continues to grow. In a separate document published in 2012, Gartner states that consumers will store over a third of their digital content in the cloud by 2016 (Gartner, 2012). Figure 2.4 shows the major cloud data storage companies along with the number of subscribers for each as of August 2014. As of 2011, Dropbox alone was handling over 500 million file uploads from its users on a daily basis (Bergen et al., 2011).

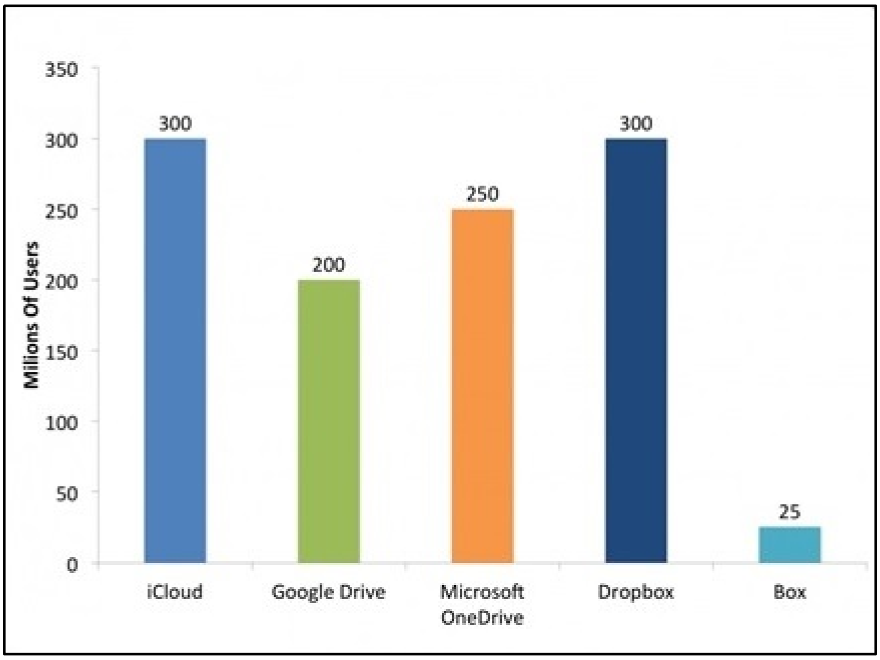


Figure 2.4: Number of users per Cloud Storage Service (source: Business Insider, 2014)

Cost is an important factor to a user when deciding on a vendor or service. All service providers charge for their cloud data storage, however they all also offer a limited amount of storage for free and this ranges from 5GB to 15GB depending on provider (Quick and Choo, 2013b). To gain additional storage space, a user simply needs to sign up for a subscription with the provider and this will incur a monthly fee. Because of intense competition in the area of cloud data storage, the cost to the consumer continues to fall. For example, in September 2014 Dropbox slashed its cost for a monthly subscription for 1TB of data by 90% (Chowdhry, 2014) to keep up with competitors such as Google who had previously reduced their monthly subscription for 1TB of data by 80%.

Table 2.1 below shows the current cost of this additional space for personal users with these service providers. Because not all vendors offer Euro pricing, US dollars was used for this comparison. Pricing information was obtained from each vendor's website and is valid as of February 2015. As can be seen, Microsoft offer the cheapest solution with their OneDrive product which costs under $7 per month for 1TB of data.

Provider	Free Storage	Cost per month	Storage amount
Apple iCloud	5GB	$19.99	1TB
Google Drive	15GB	$9.99	1TB
Microsoft OneDrive	15GB	$6.99	1TB
Dropbox	2GB	$9.99	1TB
Box	10GB	$4	100GB

Table 2.1: Cost of subscription cloud data storage as per vendor advertised figures

2.9 Cloud Security and Privacy

The majority of previous research carried in this area focuses on security and privacy issues for organisations rather than individuals. However, some of this information can be used when discussing personal users of the solutions because the underlying technology is directly comparable. Additionally, there has been much research in similar technologies such as e-commerce focusing on personal users and their security and privacy perceptions of the technology and this too is a good source of information.

As cloud computing continues to gain momentum and the number of users of its services continues to rise, initial enthusiasm has somewhat given way to a critical evaluation of the benefits that can be drawn from these services (Martens and Teuteberg, 2012). Figure 2.5 below from Carroll et al. (2011) illustrates how security remains the highest risk to cloud computing and this is a concern to users who are considering cloud computing options.

Even from an enterprise perspective, companies are looking at Cloud Computing as an ideal way of cutting costs, but do they know how private and secure the service really is (Mather, Kumaraswamy and Latif, 2009)? Security and privacy remain the two main obstacles to wider adoption of cloud computing services. One of the risks regarding cloud computing security is that that service providers have to potentially manage millions of customer user accounts and this presents a challenge in itself (Ohlman, Eriksson and Rembarz, 2009).

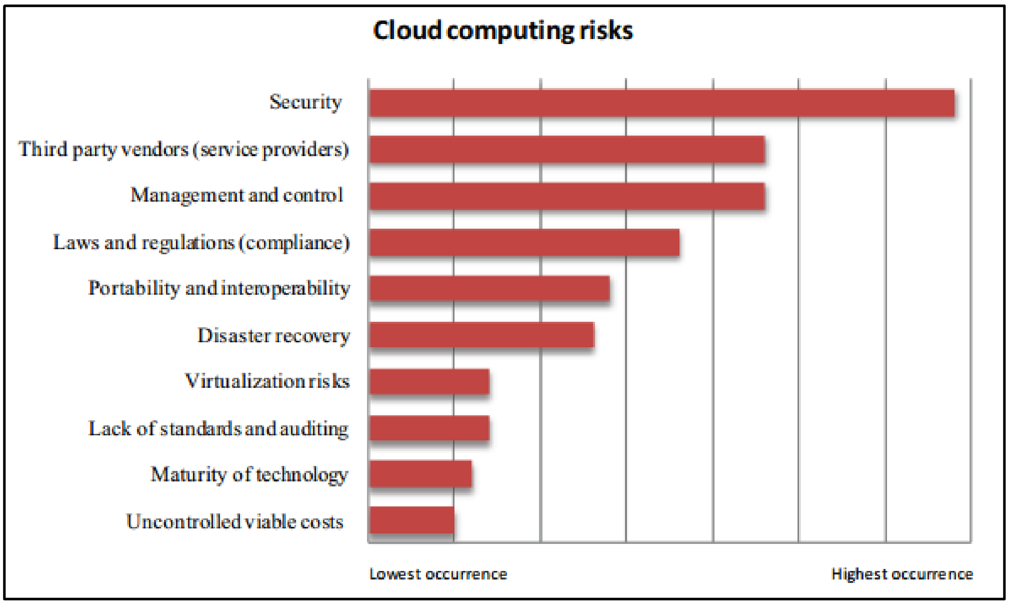


Figure 2.5: Cloud computing risks (source: Carroll et al., 2011)

Users need to be vigilant in understanding the risks associated with data breaches in the cloud service environment (Subashini and Kavitha, 2011) and how these may possibly affect them. Once users no longer physically possess their data, its confidentially and integrity is effectively at risk. Additionally, public cloud solutions may appear to be cost effective, but they also present additional security and privacy issues that must be mitigated against. It can therefore be argued that the cloud is intrinsically insecure (Ren, Wang and Wang, 2012).

All cloud service models are open to possible attack. For example, Ristenpart et al. (2009) demonstrate how a Virtual Machine (VM) residing with a cloud service provider can be used to mount cross-VM attacks on other VMs residing on the same host server in order to extract information. Further examples of possible vulnerabilities include accessibility vulnerabilities, virtualization vulnerabilities, web application vulnerabilities such as Structured Query Language (SQL) injection and cross-site scripting, physical access issues, privacy and control issues arising from third parties having physical control of data, identity and credential management issues, data verification issues, tampering, integrity, confidentiality, data loss and theft, authentication issues and IP spoofing (Subashini and Kavitha, 2011).

All service maintenance is carried out by the service provider which may leave the customer unaware of potential downtime. Additionally, the data as well as certain processing is handled on remote servers by the service provider so the customer has to trust the provider on the availability of data and the security. The SLA is the only legal agreement between the vendor and customer and must protect them against these issues (Kandukuri, Paturi and Rakshit, 2009). As an illustration, an SLA document should include the following:

- Definition of Services
- Performance Management
- Problem Management
- Customer Duties and Responsibilities
- Warranties and Remedies
- Security
- Disaster Recovery and Business Continuity
- Termination

2.9.1 Privacy issues

What is privacy and how does it relate to cloud data storage? The definition of privacy varies across publications, countries and cultures. For example, the Oxford English Dictionary defines it as “a state in which one is not observed or disturbed by other people” and “a state of being free from public attention.” Focusing on cloud computing legislation, according to Mather et al. (2009), privacy is a basic right in the European Union whereas in the United States, privacy is more focused on avoiding harm.

It is argued that the majority of security and privacy issues related to cloud computing are due to the lack of control that the customer has over the physical infrastructure (Subashini and Kavitha, 2011). Furthermore, privacy becomes even more important for a company when they manage confidential data such as customer information. For all users, there must be trust that the cloud service provider will protect their data from unauthorised users. Recent mishaps such as the 5 million Gmail usernames and passwords that were leaked in 2014 (Gordon, 2014) does nothing to instil confidence in the general public.

Privacy is important to both organisations and the individual user because personal or sensitive information could be stored with the service provider but it is not yet completely understood whether the cloud computing infrastructure will be able to support the storing of sensitive information without making organisations liable for breaking privacy regulations (Carlin and Curran, 2011). This is especially relevant in multi-tenancy cloud environments as it is possible to extract a third party’s private information through traffic patterns and other side-channel information allowing for a possible “privacy leak” (Ristenpart et al., 2009).

2.9.2 Control of data

Enabling a third party service to take custody of personal data raises important issues regarding control and ownership (Hayes, 2008). Questions must be asked such as, if you delete your account, can you be certain that the data has been expunged from the service providers' servers? If you would like to move to a different service provider, can you migrate your data while still under contract? Finally, if you are paying for a cloud service such as data storage, what happens if you do not pay a bill? Does this mean that you will lose access you your data?

Each service provider will have their own terms of service (TOS) and privacy policy which must be confirmed when signing up to the service by clicking the ubiquitous box. If the box is not checked then the user will be denied access to the service, but if the box is checked without carefully reading the TOS and privacy policy then this can have a negative impact on the user's legal rights (Kesan et al., 2013). It is vital that the user of the service is aware of what exactly they are agreeing to when accepting the TOS and possible impact there may be on their privacy rights.

2.9.3 Auditing

To comply with Sarbanes-Oxley (SOX) and compliancy such as Health and Human Services Health Insurance Portability and Accountability Act (HIPPA), the cloud service provider must allow the ability to audit it services as well as other measures (Armbrust et al., 2010, Tate 2014). In a standard non-cloud environment, the organisation alone would be responsible for this auditing, but in a cloud environment, the provider must be relied upon to meet some if not all of the auditing requirements. Additionally, the provider needs to show their customers that they are providing appropriate security measures that will protect their data and build up confidence for their service. One way they can achieve this is through the use of third party auditors (Mikkilineni & Sarathy, 2009).

Auditing also helps to ensure that cloud service providers are kept honest. For example, a provider may discard data which has not been accessed or rarely accessed to save the storage space or keep a lower number of replicas than agreed in the SLA. Moreover, they may decide to hide a possible data loss and claim that the data still remains stored in the cloud (Wang et at., 2009). As a result, data owners need to be persuaded that their data remains correctly stored (Yang and Jia, 2011).

2.9.4 Cloud data storage security issues

The type of data stored with a cloud provider ranges from publically sourced data which has low security concerns to highly sensitive information such as personal information and medical records. The storage of personal data in the cloud creates its own set of regulatory concerns that directly or indirectly impacts security. Some examples of these concerns include:

- Who has jurisdiction over data as it travels across international borders?
- Can governments access the data as it changes jurisdiction?
- Is there more risk in storing personal information with a single data centre compared to data stored in multiple data centres?

Legislation has not yet been written to address these issues in detail. Resolving these security and regulatory related concerns will take years, and will significantly influence the overall evolution of cloud computing (Kaufman, 2009). This could well have a negative impact on the speed of adoption of cloud computing data storage services.

Weak security measures resulted in the recent leakage of compromising celebrity photographs from their Apple iCloud accounts (Lewis, 2014). Normally, a set number of failed logon attempts would result in an account becoming locked for a period of time but in this situation, the hacker could continually guess the username and password of the individual until they gained access. It transpired that many of these passwords were weak, thereby allowing the hacker to easily accomplish their goal and gather personal materials such as photographs and videos.

As mentioned previously, one of the benefits of cloud data storage is that it allows the user to synchronise their data to multiple devices in order to allow for easy access to their data. But, what happens if the device is then lost or stolen? If the device itself is not password protected, a third party can then gain full access to the users' data. All cloud data storage providers allow for a 'remote wipe' functionality to help in this situation, but studies have shown that data remnants can still remain on certain devices (Quick and Choo, 2013a; Quick and Choo, 2013b).

The use of personal cloud data storage solutions is also causing problems for businesses as individuals are accessing their storage accounts in the workplace. According to an Enterprise Strategy Group Report, 70% of organizations know or suspect their employees are using personal online file sharing accounts without formal IT approval. (Kao, McClure and Oltisk, 2012). This has led to what is commonly known as 'The Dropbox Problem' in the IT industry and because of this, Dropbox is now the number four banned application in the enterprise (Smith, 2012).

2.9.5 Location of Data

If an individual relies on a cloud computing service for data storage, this could well result in the data being stored on unseen computers, whereabouts unknown, possibly scattered across the continents (Hayes, 2008). For example, at the end of 2014 Dropbox claim to have over 300 million users, 70% of which are outside the United States (Waters, 2014). However, Dropbox do not have a data centre outside of the United States. Microsoft and Amazon have a different data storage infrastructure which spans multiple continents (Slatman, 2013). In this setup, for users of Microsoft OneDrive services their data is normally stored in their closest data centre but as part of Microsoft's own business continuity planning, this data may be mirrored to different locations so it is not possible for Microsoft to guarantee the actual location where your data is stored.

The European Union (EU) formalised a system of privacy legislation known as the Data Protection Directive (Directive 95/46/EC) which was implemented in 1998. This legislation effectively prevents companies based in the EU sending personal data to countries outside the EU unless there are guaranteed levels of protection. The 'Safe Harbor' privacy framework allows companies in the US to register their certification if they meet EU privacy requirements. This allows for the transfer of personal data to the US in a more streamlined manner and offers a higher level of privacy protection to European citizens (Export.gov, 2012).

But the Safe Harbor programme is not without its critics. It was formally adopted by the EU in 2000 but reviews carried out by the EU in 2002 and 2004 resulted in negative findings and the growing number of false claims made by organisations represent a new and significant privacy risk to consumers (Connolly, 2008). Users from the EU need to be aware of this situation and it is important to understand that the levels of privacy protection to the individual user is higher in the EU than other parts of the world.

In many situations, the indivudal simply does not know where the data is being stored and this may be an issue because of the different levels of protection offered. For example, in many South American and EU countries, certain types of data are not permitted to leave the country because it contains potentially sensitive information (Subashini and Kavitha, 2011). In this situation, if an individual's data resided on a data centre in the United

States a US judge can issue a subpoena to order access to that data. It is irreverent what nationality the user is or in what country they reside.

To add a level complexity to this concern, in 2014 a US judge ordered access to data which was held in Microsoft's data centre in Ireland (Kennedy, 2014). However, under Irish Law, the data cannot be handed over without approval from an Irish Court. The judgement is currently being appealed by Microsoft but the outcome of case will have ramifications for other US cloud computing companies with operations outside the US as well as its customers.

2.10 Conclusion

Cloud computing is maturing as a technology and it is clear to see that it is here to stay; in fact it is set to become even more intertwined in our digital lives as additional services continue to become available over time. There are clear benefits to the user; it is possible to gain access to a wide range of applications and cheap data storage supplied by their service provider on a free or pay-as-you-go basis. This has been largely facilitated by increases in internet connection speeds and stability.

Increasing competition has resulted in more vendor choices for the consumer and in the case of cloud data storage, a drastic reduction in costs over the past year. Even if a user does not wish to pay for such a service, there are many options available with vendors offering large amounts of space at no cost to the individual user. The service provider can provide these services at low cost because of the ability to construct and operate large-scale data centres at low cost locations. They can then offer their services on a multi-tenancy basis, charging users as they use their services.

The ability for an individual to access their data from any location on many different types of device offers great degree of flexibility. It is a simple task to save and synchronise data to the cloud service provider, making it available on as many devices that have been configured for use with the service.

Individuals need to be aware of the potential risks that come with using cloud computing services. Because data is stored remotely with the service provider in a public, multi-tenancy environment, this should be considered as intrinsically insecure and the individual user should ensure that the service provider has appropriate security measures in place to mitigate against potential risks. All data traffic should be encrypted to protect any intercepted data from being useable.

Cloud data storage for example offers flexibilities such as the ability to access data from any location at any time on virtually any device but there are potential drawbacks that individual users must be aware of. Is the average user of these cloud service aware of these risks? For example, if all of their data stored with a cloud provider and it goes out of business, what happens to their data and can they retrieve it? Are they aware of any potential privacy issues regarding who can gain access to their data and under what circumstances? Finally, is the location of where the data is to be stored of importance to the individual or organisation? If so, this needs to be confirmed before agreeing to a service providers' terms and conditions.

Despite possible disadvantages and security issues, cloud computing remains an excellent option for the individual and the business and it has great potential for the future. Its user base is growing constantly and as more vendors are added to it, better, more fine-tuned services will become available. Stiff competition, in the area of cloud data storage for example, will ensure that usage costs for the individual will remain low. Service providers however must ensure that the end user trusts their service to encourage adoption.

3. Methodology and Fieldwork

3.1 Introduction

There are some common research methodologies and philosophes available to a researcher and these were investigated to find the most appropriate to address the research question. A research methodology is a structured framework used to describe, explain and justify the various methods for conducting research (Saunders, Lewis, & Thornhill, 2012).

The objective of this research is to examine levels of awareness of security and privacy risks among users of personal cloud data storage solutions. The following sections outline the methodological approaches that were considered and why an interview along with an associated observation session were the methods chosen for the research. This is followed by a description of how the research strategy was implemented using data collection and analysis as well as the ethical considerations that were considered during the process. Finally, the lessons that were learned during the research process are discussed.

3.2 Purpose of the Research

The purpose of this research study is to gain an insight into whether personal users of cloud data storage solutions are aware of the associated potential risks. These risks are focused on security and privacy issues. It is intended to ascertain if the participant knows, or is even concerned, about such issues. For example, data located within an individual’s cloud data storage account may be stored in a location outside of Ireland which could lead to risks such as who potentially might be able to gain access to it. It is intended to ascertain if users of these solutions store personal or confidential information with their cloud data provider and if they are concerned with these risks.

Additionally, it is intended to investigate if individuals are using cloud data storage unbeknownst to them. For example, it is possible for a user to create a cloud data storage account while going through the process of setting up a new device such as a smartphone or tablet. A step is also included when setting up a Mac computer to create an iCloud account and this is equally true for newer Windows operating systems (for the Microsoft OneDrive service). This may be important to individuals, especially if they are unaware that an account may be in use or where their data is being stored.

3.3 Research Methodologies

According to Bryman & Bell (2011), a researcher should select a research method which best fits the purpose of the research being undertaken from the available options. The selection of a particular methodology will have its associated advantages and disadvantages so it was important to consider all available methodologies before making a decision.

Saunders, Lewis & Thornhill (2012) divide the research process into six distinct stages – Philosophy, Approach, Strategies, Choices, Time Horizon and Techniques and Procedures. This is commonly referred to as a research "onion" because of the many layers involved and can be seen in figure 3.1 below. This framework was used to guide the methodological decisions during the development of this research.

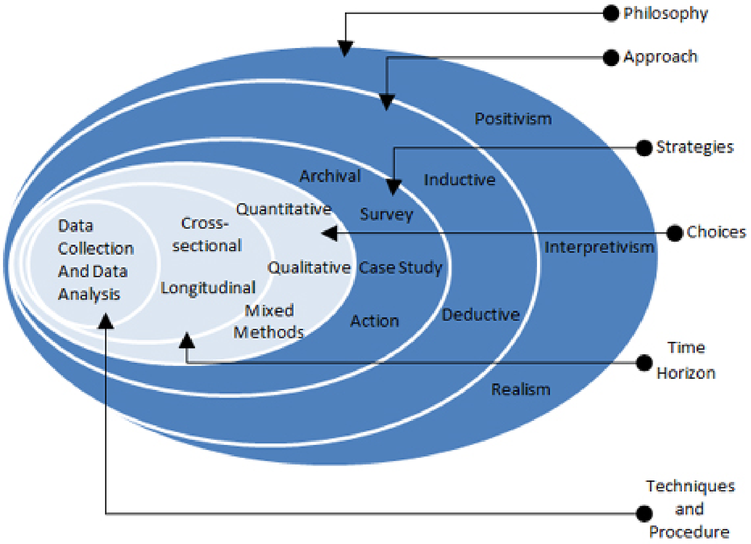


Figure 3.1: The Research 'Onion' (Saunders, Lewis, & Thornhill, 2012).

3.4 Research Philosophy

The outer layer of the research "onion" is concerned with the research philosophy. A research philosophy describes how the data within a research is collected, analysed and interpreted (Yin, 2003). For Information Systems research, four main research philosophies are commonly employed: Positivism, Interpretivism, Realism and Pragmatism.

3.4.1 Positivism

Positivism generates research that is considered to be factual and objective. It is a highly structured methodology which allows for replication and it involves working with an observable social reality with an end product of 'law-like' generalisation as in physical/natural sciences (Remenyi & Williams, 1998). The research is considered to be value free, but by adopting such a philosophy it can be argued that the researcher has already taken a value driven position. Data is gathered by methodologies such as observation for example and is generally quantifiable in nature.

3.4.2 Interpretivism

Interpretivism is a very different approach to positivism in that it takes into account social aspects during the research process and uses qualitative methods. The belief is that research is carried out on people, not objects and human behaviour cannot be quantified in the same way as physical sciences. An interpretivist philosophy is suited to research in business related fields as they are often complex and unique and the researcher should aim to understand the motives, intentions and actions of the subjects involved with the study.

3.4.3 Realism

Realism is similar in many respects to positivism however it differs in that it recognises that humans are affected by social forces and processes so cannot be studied in the same manner as natural sciences. Realism is an appropriate philosophy to adopt when studying Information Systems where a quantitative analysis is required. Realists take a scientific approach to the collection, analysis and development of data but view their findings as evidence-based probabilities. (Guest, Namey and Mitchell 2013).

3.4.4 Pragmatism

The pragmatic philosophy is based on the assumption that no other philosophy meets the requirements of the research and may in fact have a limiting effect on the research process itself. The belief of the pragmatist is that the most important factor in the selection of the research philosophy is the research question and this drives the research. The focus is on a practical approach using different methods to collect and analyse data and this allows for multiple approaches to be used to answer the research question. Because mixed methods can be useful in gathering data, it can then be analysed using quantitative and\or qualitative methods where appropriate.

3.4.5 Selected Research Philosophy

A number of philosophies were investigated for their appropriateness to this study. Initially, Positivism was selected as the research philosophy because there was an expectation that the research would be completed by means of a survey and that the results would be fully quantifiable. However, as the research planning developed, it was discovered that this approach would not be appropriate as more interaction with participants would be necessary and furthermore, the results would not be quantifiable only in nature. Another reason that the data gathering process would have to be a more interactive process than a survey was because it is assumed that some participants may not be aware of the fact that they are using cloud data storage.

Additionally, an interpretivist philosophy was not deemed as appropriate because at some level, there would be baseline results that would be quantifiable. Although there would be significant data gathered that would be qualitative by nature, the expectation was that there would also be data available where quantitative analysis would be the most appropriate method to use.

For these reasons, pragmatism was selected as the research philosophy to be used for this study as it provides a practical approach and embraces mixed methods. This combination of methods results in an analysis revealing greater insights (Kaplan & Duchon, 1988) and allows for both quantitative and qualitative methods. This philosophy also aligns with the mixed research methods that were selected to gather the primary data.

3.5 Research Approach

The next stage when following the research onion model is the selection of the research approach. There are two possible approaches: deduction and induction. The deductive approach involves developing a hypothesis and testing it using an explicitly designed research strategy while the inductive approach builds a theory based on the data collected (Matthews and Ross, 2010). According to Saunders, Lewis & Thornhill (2012), neither approach is tightly aligned with a particular research philosophy, but the deductive approach has stronger attributes of positivism while induction has attributes more associated with interpretivism.

Driven by the research question and the adopted pragmatist research philosophy, it was possible to use elements of both research approaches for this study. During the literature review process, it was possible to produce theoretical framework with regards to individuals and their usage of cloud computing services. By using the deductive approach, this theory was then tested against participants of the study to see if it held true. Furthermore, it was possible to use an inductive approach to gain further information from participants and draw additional conclusions from this data.

3.6 Research Strategy

There are several research strategies available to a researcher in order to complete a study. Examples of these include experiments, surveys, case studies, action research, grounded theory, ethnography, archival reviews, interviews and observation.

Each of these options were reviewed in detail and the majority were ruled out for reasons of applicability to the study in question and it was determined that by using two strategies, this would be the most appropriate and beneficial approach. According to Yin (2013), a mixed-method approach like this can increase the validity of the research by examining the same incident in different ways, thereby minimizing any weaknesses that may result from a single approach. These strategies are outlined below -

3.6.1 Semi-structured Interviews

The primary method that was considered to be the most appropriate for this study were semi-structured interviews of participants on a one-to-one basis over a set period of time. Interviews are an appropriate method to use when the researcher’s objective is to understand the participants’ experiences and views on a specific topic (Easterby-Smith, Thorpe & Jackson, 2008). Semi-structured interviews were preferred to a fully structured interview method because it gave the ability to use a set of structured questions associated with fully structured interviews but allowed enough flexibility to enable the interviewee to elaborate on any subject raised during the interview process.

The interview session focused on direct questions such as cloud data storage usage and also if the interviewee was aware of associated security and privacy issues. Furthermore, it was important to ask questions to ascertain if they were concerned about such risks. However, as with a survey method, there was still the possibility that the interviewee would be unaware that they were actually using cloud data storage so it was decided that an associated observation session would be required. Additionally, the interviewer could deviate from the pre-planned question list if there was a subject that warranted further investigation.

Prior to proceeding, a list of interview questions was developed to ask each participant. After some refinement, the final list of 35 interview questions (listed in Appendix 3) was drawn up under the following headings –

- Basic background information
- Technical background
- Cloud data usage
- Security & Privacy

Extensive notes were taken during each interview to allow for analysis of the information that was collected.

3.6.2 Observation

It was expected that not every participant would not know for certain that they are actually using a cloud data storage account, so it was deemed appropriate to have an observation session with the participant following the semi-structured interview.

The observation session was designed to be a secondary session and would be short. The purpose was to quickly ascertain if the participant was using cloud data storage without their knowledge. For example, they could have signed up for a cloud account while setting up a smartphone or computer and be unaware that data is actually being synchronised to their online account. Additionally, the observation session allowed for the inspection of their level of general security awareness when performing tasks such as signing on to secure websites.

The participant was asked to check their smartphone. If it was an iPhone then the investigation would take the course of an iCloud account. If it was an Android phone then we would look at the possibility of the user having a Dropbox or Google Drive account. Finally, if it was a Windows Phone, the participant would be guided to check to see if a Microsoft OneDrive account was present. As a further step, the participant would be asked to connect their smartphone to their computer to observe if any synchronisation of data (photographs, contacts, etc.) takes place.

As a final step in the observation session, the participant was asked to log onto their computer to see if there were any desktop synching tools installed from cloud data storage providers. Such tools allow for the automatic synchronisation of data from the local computer to the cloud provider.

3.7 Population

A research population is the total number of individuals or objects that are the main focus of the study. The population of this study is the number of individuals in Ireland that use personal cloud data solutions either on their computers, smartphones or both. Collecting information from the entire population is not feasible for this study so a sample or subset of the population will be considered instead. Findings will be extrapolated based on the data received from this sample population.

3.8 Sample Selection and Data Collection

It was projected that each session would take up to one hour to complete – a maximum of 50 minutes for the semi-structured interview and 10 minutes for the observation session. It was planned to go through a total of 18 sessions with various participants with ages ranging from 18 to 65.

There were a number of ways that the sample population could have been selected. For example occupation and technical competence were considered but it was decided to group participants according to age group. It was expected that this would give the most random level of results and there would be enough mixed data, independent of participants' technical background or occupation. Table 3.1 below shows the participants of the study divided into the sample ranges.

Age range	Number of Participants	Technical Background
18 – 30	4	Various
31 – 45	5	Various
46 – 60	5	Various
61+	4	Various

Table 3.1: Sampling table

The participants of this study were all from Ireland, the majority of which came from the greater Dublin area.

3.9 Research Ethics

Prior to arranging the interview and observation sessions, approval was gained from the Ethics Committee of the School of Computer Science and Statistics from the University of Dublin, Trinity College. The application included a Project Proposal form that fully outlined the details of the research and ensured any issues such as conflicts of interest were identified and that all relevant documentation was made available to each participant. This included an Informed Consent Form and Information Sheet. These gave a detailed description of the study, its purpose and important information such as the participant's right not to respond to any specific questions without prejudice to the research and that they could withdraw from either the interview or observation process at any time.

It was also noted that upon the unlikely event that any illicit activity or materials were discovered during the interview or observation sessions, there would be an obligation on the researcher to report this to the relevant authorities. The sessions were not to be recorded and retained, however detailed notes would be taken and this was explained to every participant. A guarantee that anonymity would be preserved was included in the Informed Consent Form and it was highlighted that any data retained will be in accordance with the Data Protection Act. After supervisor review and submission for approval, permission to proceed with the gathering of data was received in May 2015 from the Ethics Committee.

3.10 Lessons Learned

Selecting the overall research topic proved to be a very difficult process. A research area was initially selected, but then it was found that a student in a previous year of the course has covered the same topic. This wasted a significant portion of time early during the dissertation process and it would have been an important benefit if the previously submitted dissertations had been checked at an earlier stage.

It proved to be a time consuming process to select an appropriate research strategy for this particular study. Initially, a survey was thought to be the most beneficial way of gaining a large amount of data quickly, but upon developing a test survey and sampling it, it was discovered that participants would not necessarily be in a position to answer all of the questions correctly. All strategies needed be fully considered before selecting the appropriate strategy for the research and a semi-structured interview was then suggested to be the best form of data gathering. There still remained a limitation with this strategy in that the interviewee may not be aware of their cloud data usage so an associated observation session had to be included with each participant.

Because an interview is a more time consuming and interactive process, getting people to participate proved to be difficult, especially because of the additional requirement to have access to their personal computer for the observation process. This meant travelling to each participants home if a desktop computer was their everyday computer.

Finally, asking people questions such as ‘rate your technical knowledge’ is a very subjective question and the response varied according to the interviewee’s opinion. Perhaps asking more direct questions to challenge their answer would give a more objective standard of answer.

3.11 Limitations

There are a number of limitations to this study. Firstly, the small sample size would result in a narrow set of findings depending on the participants. This would be further skewed somewhat because a number of the participants work in the IT industry and may have a deeper understanding of cloud computing and cloud data storage than the average user. The expectation was that because the participants invited to take part in the study came from a wide range of ages, this should mitigate against this potential issue.

The demographic was not very varied and it would be a big benefit to the study if it was widened. For example, since cloud data storage is becoming more prevalent, and in a lot of situations it is the de facto storage option, it would be interesting to focus this study on a younger age group because as was shown as part of this research, this demographic have embraced the technology more freely.

The chosen methods themselves have underlying limitations. Semi-structured interviews regularly suffer from data quality issues related to reliability, bias and validity and generalisability (Saunders, Lewis & Thornhill, 2012). In order to address these limitations, the interview questions were worded as concisely as possible in order to not lead the participants in a particular direction with their answers.

In terms of this research, two potential areas of bias have been identified. Firstly, individuals that are working in the IT industry or are very technically literate may be more open to the adoption of cloud data storage solutions. They may also have a better understanding of the potential risks and how to best mitigate against them. Conversely, individuals that are not used to the technology or have a low level of computing experience may be wearier of using cloud data storage solutions. This is especially true if a solution is deemed to be over-complicated and not user-friendly because the resultant take up by less technical individuals would be lower.

4. Findings and analysis

4.1 Introduction

This chapter presents the findings and analysis from the data collected during the research. As explained in the previous chapter, the primary source of data was through the use of semi-structured interviews along with associated short observation sessions with each participant. A total of 18 participants took part in these sessions as part of the study and they were of various ages and technical backgrounds.

A framework of 35 questions formed the core of the semi-structured interviews, however these were used as a guideline only and the conversations would generally go beyond the pre-defined questions where necessary to obtain further information if it was deemed to be relevant or interesting. The observation session then followed with each participant. The primary purpose of this was to discover if they were using cloud data storage solutions that they did not explicitly sign up for. Additionally, the observation session allowed for the inspection of their level of general security awareness when performing tasks such as signing on to secure websites, etc.

From the literature review it was found that cloud computing was a maturing technology and the number of individuals that are using its services has been increasing. Specifically, the uptake of cloud data storage services has seen significant growth and this allows individuals to easily share data with others and access it from multiple devices from many locations (Lin and Chen, 2012). The intention of this research was to discover whether the following research questions hold true:

1. How do users of computers and mobile devices currently utilise cloud data storage solutions and do they pay for such solutions?
2. What is the perception among users of potential security and privacy risks associated with using cloud data storage solutions?

4.1.2 Qualitative analysis and theme development

From an analysis of the notes gathered from the semi-structured interview and observation sessions, common trends and observations emerged and these were highlighted by the use of coding of the transcripts. Some of these common trends were identified across all participants, whilst others were only identified in a smaller grouping of participants. Refining the common trends led to the development of the themes and this was an iterative process whereby some elements clearly fitted together from different interviews while others required new themes or sub themes to be developed from within a main theme. From this process, three main themes were derived:

- Technical background and technology usage
- Cloud data usage
- Security and privacy: perception versus practice

4.1.3 Quantitative analysis

In contrast to qualitative analysis, quantitative analysis involves some form of statistical, mathematical or computational techniques to be applied to the collected data and it implies measurement. The output is always in a numerical form such as statistics. Because a number of questions or topics discussed with the participants resulted in yes\no or short descriptive answers, this allowed for the collected data to be analysed and formed into graphs or tables. This held true for the observation sessions also; a number of the tasks that the participants were asked to perform resulted in outcomes that could be easily noted and compared. For example, asking a user to log into a website and noting if they used stored usernames and passwords or manually enter them each time.

4.2 Participant background and technology usage

The first stage was to analyse the information that was gathered about the participants themselves and their current uses of technology. This allowed for a baseline to be created regarding their technical backgrounds and also discover relevant information to the research relating to their use of technology, up to and including cloud data storage usage.

The participants were first asked to declare their level of technical competency. It was explained that the intention was not necessarily to focus directly on IT or computers and was simply to get a benchmark of how they perceive themselves technically in a general sense. The purpose was to build a guideline from which the remainder of the interview could be guided where possible and also there was an expectation that less technically-minded participants would not be using cloud data storage solutions to their full potential. Making cloud computing solutions easy to use for all users of all levels of ability is an important factor for the adoption of such applications (Pew Research Centre, 2008). Enquiring about the technical level of each participant and measuring their usage of cloud data solutions investigated if this statement held true.

Figure 4.1 below shows the results of this question and as can be seen, more than half of the participants consider themselves to be above average with a score of 6 or above, with 7 being the most popular self-assigned mark.

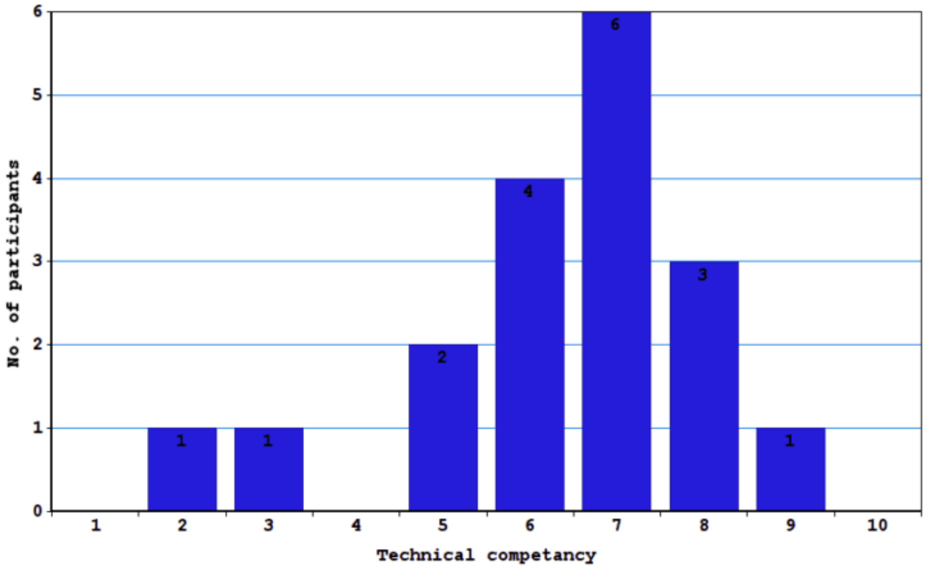


Figure 4.1: Self-rating of technical knowledge

Analysing the information further and focusing on the different age groupings, table 4.1 below shows the average self-declared technical competency for each of the different age groups.

Age range	Number of Participants	Avg. Technical competency
18 - 30	4	7.6
31 - 45	5	8.2
46 - 60	5	7.3
61+	4	4.8

Table 4.1: Self-declared technical competency by age group

It can be seen that the 31-45 age group were the most confident with technology. Whether this translates to actual competency is not necessarily true; it merely states that the group on average felt that they were more confident with declaring a higher level of technical competency. This was not what was expected based on previous research such as from OFCOM (2014) which suggested that the younger group of participants would more comfortable with technology in general. In this situation perhaps confidence or acknowledging ones limitations was a factor for choosing a lower score.

It should also be noted that all bar one of the older participants needed to have cloud computing explained to them, albeit at a very high level. They simply did not know if they were using it or not. For example, they had no recollection of going directly to Dropbox or Google Drive and signing up for storage, but they did have a Gmail account which automatically creates a Google Drive account for that user, along with other applications from the Google ecosystem such as YouTube and Google+. In situations like this, they may have had a cloud data storage account setup for them when they were signing up for a different application.

But these types of accounts may have been created automatically without the user necessarily knowing what they were doing. For example, signing into a Winnows 8\8.1 computer for the first time will create a Microsoft account for the user along with an associated OneDrive account automatically. Registering an Apple product such as a Mac laptop or iPhone for the first time will create an iCoud account for the user while a user has the option of creating a Googe Drive or Dropbox account on an Android smartphone for ‘backup’ purposes. This may lead to confusion and it shows that users of these technologies may simply not realise that they are creating an account with a cloud data storage provider.

Finally, the different employment sectors for the group of participants can be seen in figure 4.2. The relevance of this is to give some further background on the participants and perhaps give insight into their experience with computing technologies. It was also to demonstrate that the participants worked in a spread of different employment sectors.

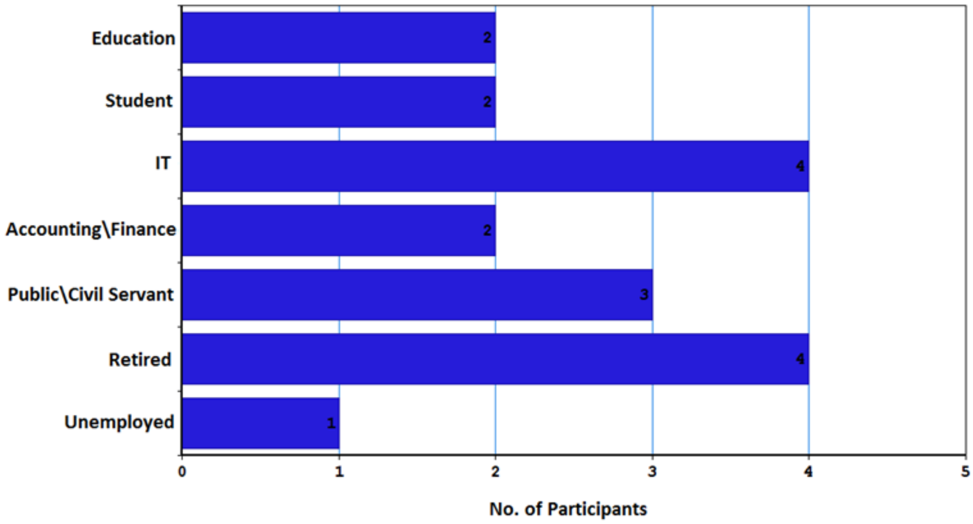


Figure 4.2: Employment sectors

4.2.1 Cloud Data Storage usage

All apart from one of the participants stated that they currently used a cloud data storage solution, which resulted in a rich resource of interview transcripts to analyse. This is not to say that all remaining participants declared that they use cloud data storage, some simply did not know and could not answer for sure when asked outright. This large uptake of cloud data storage enforces the research from Gartner (2012) which states that consumers will store over a third of their digital content in the cloud by 2016.

The only participant to declare that they did not use cloud data storage for certain described themselves as 'very conservative' by nature and relied on local file storage options only. Questioning this participant further led to them explaining that they do not use any social media applications such as Facebook and they only spent limited time on the internet.

4.2.2 Smartphone and tablet usage

All participants used smartphones with three actually having two smartphones; one for personal use and the other for business use. Not one of the participants used a standard (non-smart) mobile phone. The breakdown of the types of smartphones in use are shown in table 4.2 below.

Smartphone type	**Number**
Android	14
iPhone	5
Windows Phone	1
Blackberry	1

Table 4.2: Smartphone ownership

Even though all participants owned a smartphone, two do not use any of the 'smart' functionalities and use it simply for making calls and sending and receiving texts. Both of these participants are from the older age group.

Additionally, the tablet ownership can be seen in table 4.3 below. Four of the participants did not own a tablet, whereas some owned two or more. Whilst most people use their tablet for simple operations such as internet browsing or watching video files, some did use them for more advanced features such as accessing their cloud data storage via their Dropbox app. As one person stated:

> “This allows me to access my data while on the go and it’s great for when I’m away on business travel”. (Tony)

Tablet type	Number
Android	14
iPad	5
Windows	1

Table 4.3: Tablet ownership

The relevance of this to the research is because cloud data storage is now integrated into almost all mobile devices, allowing people to access their data on their phones or tablets from many locations. Additionally, features such as auto sync and auto backup are now becoming more standard with these mobile devices. Traditionally, this functionality occurred when they connected their mobile device to their computers, but now this has evolved to the device backing up or synchronised automatically every time it is connected to WiFi for example. Of course, this functionality is completely configurable by the owner of the device to best suit their needs.

During the observation session, each participant was asked to connect their phone to their computer. This was to discover if the user was prompted to back-up the contents of their phone to their cloud provider. Not all of the participants had explicitly signed up for cloud data storage, meaning that they went directly to the providers’ website and registered for an account. In some situations such as when setting up a smartphone for the first time the user is prompted to register for cloud storage. This may be for general data storage or simply for backup purposes but it is not always clear.

This was certainly the case for five of the participants; two iPhone owners and three Android smartphone owner were surprised to learn that they were synchronising certain contents of their phone to cloud data storage. All of this was apparently created as a step during the setup of their phone and they simply never realised. Technically, this does not impinge on their privacy rights as outlined by Mather et al. (2009) as they would have accepted an end-user license agreement (EULA) during this task. This may have been in the form of a simple check-box and this is possibly why it was not noticed at the time by the individual.

4.2.3 Computer usage

All 18 participants use Windows as the primary operating system on their laptop or desktop computers so the usage analysis was focused on this operating system only and not alternatives such as Mac OS and Linux. In order to give an indication of how many different locations the individual accesses the internet, participants were asked if they used computers at home, at work/university or both. As can be seen in table 4.4 below, the majority of participants use computers both at home and at work or university, and this is in line with observations from Lin and Chen (2012) regarding accessing cloud computing services from multiple locations.

Computer usage	Number
Home	5
Work\University	0
Both	13

Table 4.4: Computer usage

This usage shows that, in conjunction with the high level tablet and smartphone ownership of the group, each individual has a number of different options available for accessing their cloud data storage from multiple locations.

These are computers at non-mobile locations such as at home or at university or at more mobile locations, facilitated by laptops connected to WiFi hotspots. None of the participants had a laptop that was capable of connecting to the internet via cellular networks allowing for full mobility.

Finally, all participants have broadband connections at home with the bandwidth varying from 100Mbps to 240Mbps. 12 of the participants have UPC as their internet Service Provider while the remaining 6 have Eircom. This was necessary information to ascertain as noted by Bergen et al. (2011) because bandwidth is an important factor for accessing online services. As all participants had a high-speed internet connection available to them, there was no negative impact on accessing their online data because of a lack of available bandwidth.

4.3 Cloud Data Usage

4.3.1 General usage

Table 4.5 below shows the cloud data storage providers that the participants use. As mentioned in the previous section, one participant does not use any cloud data storage at in any form. All but five of the remaining participants use multiple providers for storing their data online. The most accounts that any one participant has was four, using Dropbox, Google Drive, iCloud and OneDrive.

Cloud provider	Number of users
Dropbox	14
Google Drive	11
OneDrive	8
iCloud	6
Amazon	1

Table 4.5: Cloud data provider usage

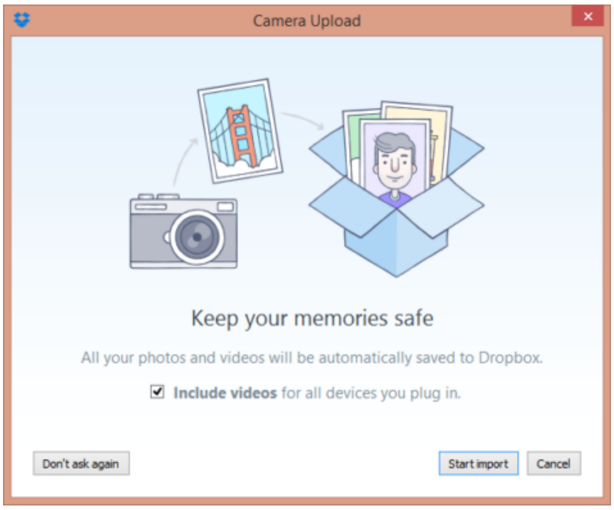
The main reason that the participants gave for signing up for multiple providers was because of the free space that is included when signing up. When asked about this, the responses were that it allows them to maximise the amount of free space and to use different providers for different functions, such as one for storing documents, one for storing music files, etc.

Although some participants mentioned different use cases such as accessing Dropbox or Google Drive on their computer while using iCloud on their iPhone or iPad only. One participant in particular took the time to go through each of their cloud data accounts (Dropbox, Google Drive and Amazon) to demonstrate what was being stored in each and describe what each one was being used for. They discussed how keeping a logical separation for each cloud data storage account based on usage made the solution more workable. For example, they would use Amazon for photos and document storage, Google Drive for their music library and a small number of movies to access when they travel and finally Dropbox was used for “everything else” such as sharing files with others. This shows the different level of understanding and usage of cloud data storage solutions amongst the group of participants.

As mentioned in the previous section, two participants from the older age group do not use the ‘smart’ functionalities of their phones at all. Therefore, they do not use their smartphones to access their data.

It did take the observation session to uncover some of this information. For example, during the interviews, the iPhone and iPad owners were asked if they had iCloud accounts. Two people said that they did not but they were asked to check their accounts on the mobile device and all confirmed that there were actually iCloud accounts setup. It was a similar situation for the Android users and in one situation, when the person was asked to connect their smartphone to their computer, they were actually prompted to confirm if they wanted to back-up to two separate providers, Dropbox and Google Drive.

The example prompt window shown in figure 4.3 below was taken during one observation session. Once the smartphone was connected to the individuals’ computer, a Dropbox prompt appeared asking them if they would like to import their photos and videos from their phone to their Dropbox account.



Keep your memories safe

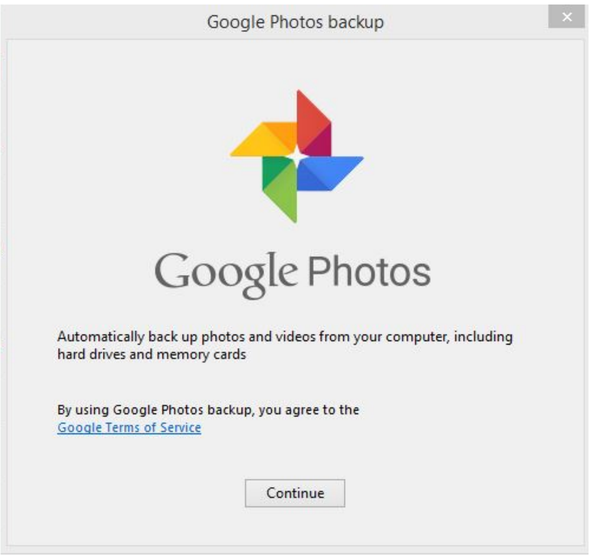
All your photos and videos will be automatically saved to Dropbox.

☑ **Include videos** for all devices you plug in.

Figure 4.3: Dropbox sync prompt

Automatically backing up data in this fashion does have certain advantages to the user and the majority of the participants have this functionality configured and are pleased with it. For example, if their phone is lost, stolen or damaged, all of their photos are not lost along with it because they are being stored in the persons cloud data storage as well as their smartphone.

This functionality is not limited to Dropbox alone. Figure 4.4 below is a screenshot taken during another observation session when the participant was being prompted by Google Drive to synchronise their photos. Similarly, iCloud offers a similar process for iPhone users.



Google Photos backup

Google Photos

Automatically back up photos and videos from your computer, including hard drives and memory cards

By using Google Photos backup, you agree to the Google Terms of Service

Figure 4.4: Google Drive sync prompt

As discussed in the previous section, participants have access to the internet via computers at home or in work\university, via their smartphones or via their tablets. The majority of participants access their online cloud data from these multiple options, confirming Quick and Choo's (2013a) observations that data can then be accessed over the internet from any location from devices such as desktop computers, laptops, tablets and smartphones which offers increased flexibility and accessibility to the user.

4.3.2 Cloud computing usage

There are many different uses for cloud data storage and the feedback from the participants was mixed. Some simply used it for basic storage of files while others used more advanced features such as creating folders to share data with others and directly storing BitTorrent downloads to a defined folder. Figure 4.5 below shows the breakdown of this feedback.

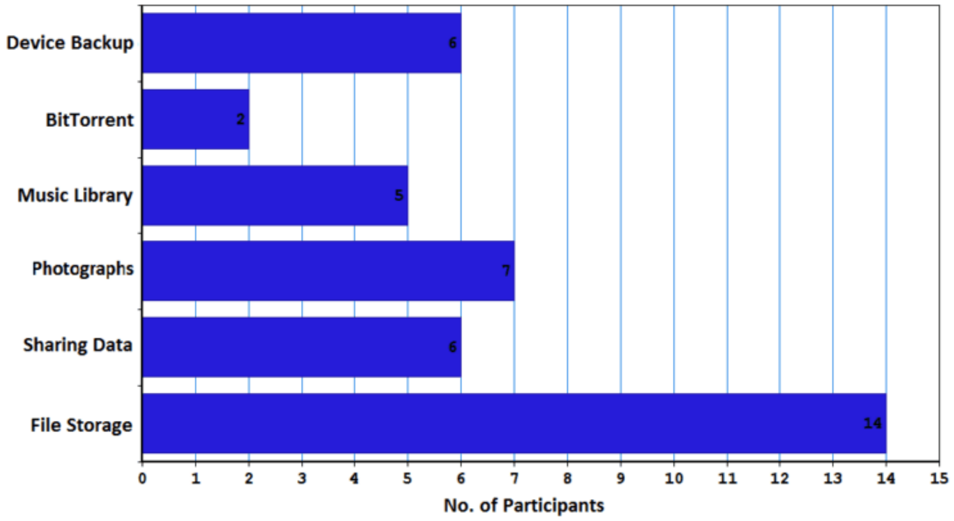


Figure 4.5: Cloud computing usage

In particular, participants that had graduated from university in recent years were more used to features such as sharing data with others because they were encouraged to do this for group collaboration and project work while they were still students. The general perception was that this was an easy task to setup shared folders and invite people to have access and was very beneficial because it allowed multiple people to work on documents or projects.

An important point that one person mentioned was that it is not possible to set the rights of the invited collaborators on a shared folder of the free versions of the products such as Dropbox. This should be considered as a security concern for some users of this functionality because people that are invited to access the shared folder have the ability to delete the files within that folder. Upon further research, it was found that in order to be able to set permissions for a shared folder to read-only, it is first necessary for the user to upgrade to the paid service.

4.3.3 Free versus paid-for solutions

Only person that took part in the research currently pays for cloud data storage and this is for the Google Drive product. Out of the remaining 17 participants, seven did mention that they would consider paying for the service in the future, particularly when we discussed how the cost has declined rapidly over the recent past. When asked if they had ever run out of space in their account(s), from the few that said that they had, the resolution was to simply open another free account; either with a different provider or with the same provider using a different email address.

Not one person knew how much 1TB of online data would cost per month and when asked for a rough estimate, every guess was higher than the actual cost of the most popular vendors. From this, it could be suggested that there would be more uptake of the paid services if more people know what the costs are, especially as Chowdhry (2014) highlighted that the cost of personal cloud data solutions has been dramatically reduced recently.

Additionally, the one person that pays for the service could not remember how much they are currently paying. They also belived that they were using the cheapest provider (Google) but this was incorrect and they were surprised to hear that Microsoft OneDrive was then the cheapest. They said that at the time when they were signing up for cloud data storage, they researched all of the vendors and Google was the most competitively priced. This person also mentioned:

> "To be honest I love using the cloud now for storage and I won't go back to using local storage as my primary option again. I store everything I can in the cloud now and the great thing is that I can access all of my data from anywhere. Even if I lose or break my laptop, I don't really care because the data is the important thing and that is in the cloud". (Paul)

4.3.4 Cloud data storage benefits

A lot of research (for example NIST 2012; Armbrust et al., 2010; Marston et al., 2011) – focuses on the benefits of using cloud computing in general. To apply this to this study, the participants we asked why they used cloud data storage and what they felt the benefits were. The following responses were identified:

- Zero or low cost: there are free options for all personal cloud data providers. The one person that pays for storage is happy that the cost is very cheap compared to traditional storage. Additionally, they mentioned that cloud data storage is cheap whilst purchasing a tablet with a large amount of memory is very expensive. The cost of cloud data storage can be easily offset in this situation and a cheaper tablet could be purchased and the data accessed from the cheaper cloud storage whilst connected to the internet.
- Capacity: the amount of free storage varies from provider to provider and one participant mentioned that between all of their accounts, they have over 100GB of storage that they do not pay for.

- Convenience: it is not difficult for participants to access their data either via the webpage or by installing the application onto their desktop, allowing them to access their synchronised data directly on their computers. This is an important aspect of cloud data usage as noted by Drago et al., (2012).
- Accessibility: the desktop applications also make it possible for users of these services to access their data while offline and during the interviews, the majority of participants remarked that this was a major benefit. Some also cited the ability for them to access their data from multiple locations and from many devices as an additional benefit, although during the observation session, it was discovered that less than half of participants had the applicable app installed on their smartphone or tablet.
- Sharing – Not many of the participants use their cloud data storage account to share files and folders with others, but those that did spoke about its convenience and ease of use.

It is important to note that not one participant mentioned security or privacy as a benefit of using cloud data storage.

4.3.5 Sync tool usage

Table 4.6 below indicates the usage of sync tools on the participants' computers. As can be seen, the majority of participants' use the sync tools to access their data locally as well as from the website. This is an important point and can be seen as a good indication of ease of use. This was the general consensus when the participants' were asked and one reply explained this well: "if it wasn't easy then I wouldn't be using it".

Sync tool usage	Number
Home	3
Work\University	2
Both	8

Table 4.6: Sync tool usage

The high number using the sync tools in work\university or both work\university and at home was contrary to expectations based on Kao, McClure and Oltisk's (2012) findings that 'The Dropbox Problem' has resulted in over 70% of organisations knowing or suspecting that their employees are using personal online file sharing accounts without formal IT approval. This has resulted in Dropbox being the number four banned application in the enterprise globally (Smith, 2012) and sync tools from other providers are equally prohibited. Only two participants reported that they could not get access to their cloud data storage in the workplace which again was contrary to the findings of Kao, McClure and Oltisk (2012), which suggests that this number should have been much higher. Furthermore, this was equally true for accessing their cloud data accounts via the providers' website; only the same two individuals mentioned that it was not possible to gain access in their workplace.

Additionally, none of the participants reported that the use of usb flash drives were disabled in their workplace. As with allowing access to cloud data applications, this should be highlighted as a potential security risk on behalf of the employers. Between these physical devices and allowing access to cloud data storage, this opens the door for the removal of potential sensitive data from the workplace without the employer's knowledge.

4.3.6 Mobile usage

Table 4.7 below is an extension of the previous table indicating smartphone ownership. From this, it can be seen that of the 21 smartphones, 10 have apps installed allowing users direct access to their cloud data.

Smartphone type	Number	App installed
Android	14	7
iPhone	5	2
Windows Phone	1	1
Blackberry	1	0

Table 4.7: Cloud apps usage on Smartphones

Similarly, table 4.8 below is an extension of table used earlier to demonstrate tablet ownership. A total of 9 of the 20 tablets had an app installed from a cloud data provider. It must be noted that some of this information was uncovered during the observation stage; not all participants were aware that they actually had an app installed on either their smartphone or tablet. The use of apps on mobile devices should be highlighted as a serious security risk, especially if the owner is unaware that they are installed. If the smartphone or tablet was misplaced and left unlocked or had a weak personal identification number (PIN) code, this could give an unscrupulous individual an opportunity to gain full access to the data on their cloud storage.

Tablet type	Number	App installed
Android	14	5
iPad	5	4
Windows	1	0

Table 4.8: Cloud apps usage on Tablets

Although just slightly under half of participants' smartphones and tablets have the relevant apps installed, the number of people that actually use their mobile devices to access their cloud data is actually quite low. Only two participants responded that they would use a mobile device on a regular basis (more than twice a week) to access their data or work on it. The majority of participants simply use the providers' website or the local copy of their data storage that is automatically available via the desktop sync tool on their computer(s). This highlights the different usage scenarios for cloud data storage usage as per the research question but this time conflicts with observations from Quick and Choo (2013a) regarding individuals accessing their data from multiple types of devices because mobile usage from this group was found to be low.

4.4 Security and privacy: perception versus practice

When asked the direct question "how security conscious are you?" the vast majority of participants responded with affirmative answers such as "very" or "extremely". Not one person answered that they do not take security seriously. From across the different age groups and levels of technical competencies, all participants had a basic understanding of what security meant in the context of general IT and cloud computing.

Perceived risk can be described as "the felt uncertainty regarding the possible negative consequences of adopting a product or service" (Cunningham, 1967). Research such as Loske et al. (2014) highlights that the perception of security is crucial for the adoption of a technology such as cloud data storage. This was directly relevant to this study because it was necessary to show the individuals perception and expectation of security versus their actual implementation.

4.4.1 Password usage

To gather more information on this subject, the participants were asked about their usage of passwords. All apart from two used passwords to access their home computers or laptops. Additionally, every person that used a computer in the workplace or at university were required to use passwords to log onto the computers. The usernames and passwords are also required to gain access to their cloud data storage account(s) and the participants were then asked about the complexity and usage of these.

In the document “Choosing and Protecting Passwords” published in 2013, the United States Computer Readiness Team describes what is suggested as a complex password and this is based on a function of length, complexity, and unpredictability of the password itself. It was not feasible to ask the participants what their actual passwords were but it was possible to have a conversation to get a description of how complex each participant’s passwords were and if they followed guidelines regarding having separate usernames and passwords for different accounts and if cached usernames and passwords were stored in their web browsers. Common guidelines for strong passwords are as follows:

- Minimum password length
- Lowercase and uppercase characters
- Numbers and symbols
- Avoid using dictionary words, names, dates and letter\number sequences
- Different passwords for different accounts

Figure 4.6 below is an indication of the strength of passwords used by the number of participants.

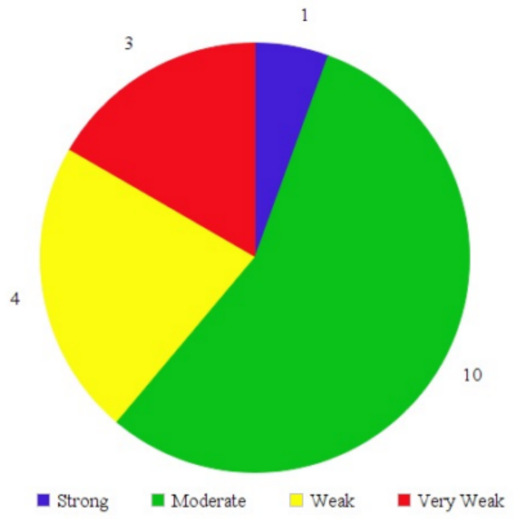


Figure.4.6: Password strength

None of the participants used a password management application such as LastPass or Keychain to securely manage passwords for separate accounts. Only one of the participants followed all of the above guidelines but they did mention that they used their internet browser to cache the usernames and passwords for their accounts. They felt that it was "too difficult to keep track of all of them". Additionally, as a backup they used a note stored in their corporate Microsoft Outlook as a backup of their passwords, although they did mention that they were "slightly changed" so would not be easily useable if found by another person.

Of the remaining participants, 10 used what would be considered as strong passwords but they replicated these across multiple accounts. These passwords may not have all of the elements listed above but would have some of the required complexity. For example, a common word or date might be in use but this would be combined with uppercase and lowercase characters and symbols. As before, caching in the internet browser was common practice.

Four participants did not use strong passwords and these consisted mainly of a dictionary word mixed with numbers, either a date or a sequence of numbers.

The remaining three participants used very weak passwords and did not follow good practice for password usage. Their passwords would be short and use a single word wherever possible. The only time that they would add complexity to a password was when the website forced them to when they were creating their account(s). They would then generally use a variation of their 'common' password for this new password with some slight added complexity such as the year of their birth.

A concerning aspect of username and password management best practice was uncovered during the observation sessions. A total of six participants had usernames and passwords written down and stored in an unlocked drawer or left openly on the desk beside their computer. They referred to these when asked to log onto their cloud data storage account for example. This finding is directly related to the 2nd research question regarding the potential security risks of cloud data storage solutions. If individuals were unconcerned with basic security practices such as username and password protection, it could be suggested that they would be equally unconcerned about using cloud data solutions. As Subashini and Kavitha (2011) point out, users need to be vigilant in understanding the risks associated with using cloud services and this should be applied to basic security principals as well.

4.4.2 Cloud data security

Security remains one of the most oft cited reasons for not adopting cloud computing services (Subashini and Kavitha, 2011), but that observation does not hold true with this research as 94% participants use cloud computing for data storage and many have multiple accounts with different service providers.

The participants were asked the direct question “Do you consider cloud computing to be inherently risky or possibly insecure?” The responses to this question can be seen in figure 4.7 below. The general response was that the majority felt safe using cloud computing for data storage with only 2 respondents commenting that they felt that it was a risky technology and should be treated with caution.

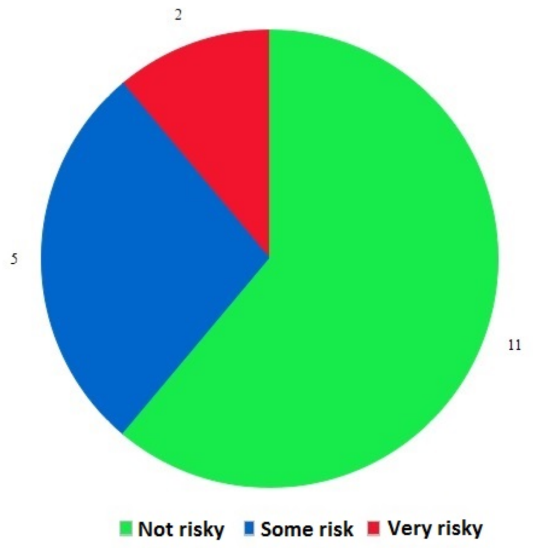


Fig.4.7: Perceived risk level of cloud computing

This general perception goes against common research such as Ren, Wang and Wang, (2012) who state that there are risks associated with cloud computing solutions and because data is stored remotely with the service provider, it can be argued that cloud data storage is intrinsically insecure. Most participants simply felt that they trusted the service provider and had little reason to feel at risk. There was an assumption from some that because the provider was well known that they would be secure. During the developing conversations, some individuals were surprised to hear about recent downtime issues (Dropbox) and the leakage of user account information from some of the well-known providers.

Just under half of the participants did not care about the potential risks associated with using cloud data storage. This contrasts with findings from previous research such as Carroll et al. (2011) suggesting that users of this technology might be more conservative with its usage. It should be noted though that very few participants said that they stored confidential documents in their cloud storage, especially the more technically minded of the group who stated that they only store non confidential data there. This highlighted the overall low expectation of privacy from the group.

One piece of relevant information that suggested why the majority of the group showed an inherent trust with the storage provider was that of the 18 participants, not one had ever had an online account of theirs hacked, nor had they lost a device such as a smartphone that may have held sensitive information. From discussing this topic, it could be suggested that they might be more cautious with their usage of cloud data storage if they experienced a negative incident such as one of these mentioned.

In order to get a comparison to the security measures that the participants currently take with traditional, non-cloud data storage such as on their computers or portable drives, some questions were asked regarding the alternative methods that each person currently utilize. Although people seemed to be more concerned with the data that they store online than locally, it was found that not one person explicitly encrypted any of their usb or flash drives and very few kept them in a locked drawer. Most people kept usb flash drives in highly accessible locations such as in the pockets of their bags or on key chains and when asked they said that they did on occasion keep personal or confidential information on these drives. Many accepted that this was potentially more unsafe than personal cloud data storage solutions.

As discussed previously, some participants did not know that their mobile devices were synchronising to their cloud account automatically and many were surprised to find that certain data was being backed-up without prompting each time. Additionally, as part of the observation processes, each participant was asked to log into their cloud data storage account(s) to confirm how many, if any, devices were registered for either desktop synchronisation or mobile access.

For example, with Dropbox it is possible to check in each users account settings how many devices are registered to access their online data. Some were surprised to see computers listed that they no longer used. For example, computers from previous employers where showing up as registered devices and if the hard drive of these computers have not been formatted since the person left that company, a full local copy of their data remains on that computer, albeit from the last time that they logged in to it. When informed of this situation, the relevant participants expressed concern and wished that they had deleted the synchronization tool and any local data at the time of leaving the company.

4.4.3 Protection of data

Protection of data relates to the end user of the cloud data application protecting themselves against the potential loss of data either by accidental means or to an unwanted party gaining access to it. To mitigate against the previously discussed topic regarding multiple mobile devices and computers being able to access and store a copy of a user’s cloud data, some vendors offer a poison pill or self-district functionality whereby all data on the remote device will be deleted, leaving the data stored in the cloud untouched.

Only three of the participants knew of this functionality. This would not be considered a surprise as it could be acknowledged as advanced technical functionality and only four of the participants work in Information Technology. Although, when asked if any of their providers had this service available, none of these participants could respond for sure but the collective assumption was that it would be available.

The participants were then asked about encryption and if they knew at a high level what it referred two. 72% of the participants could give at least a reasonable response to this question, although when asked if their cloud data storage provider employed encryption by default, only 4 of the group could say that they did for certain. The remainder of the group assumed that the service provider did encrypt all data by default, while none assumed that there was no encryption used.

Again, to compare to their current practice with their own computers, the participants were asked if they used encryption software such as BitLocker on their personal computers. Only one of the group did employ this additional security step but 7 of the group did believe that their work computer did have encryption of some form installed but none felt the need to have this on their personal computers. From this, it could be suggested that if individuals do not take steps to secure their locally held data then they would not be overly concerned with how secure their cloud data storage is.

Enabling a third party service to take custody of personal data raises important issues regarding privacy, control and ownership (Hayes, 2008). The participants were asked questions such as who could access their online data and figure 4.8 below illustrates the responses that were provided.

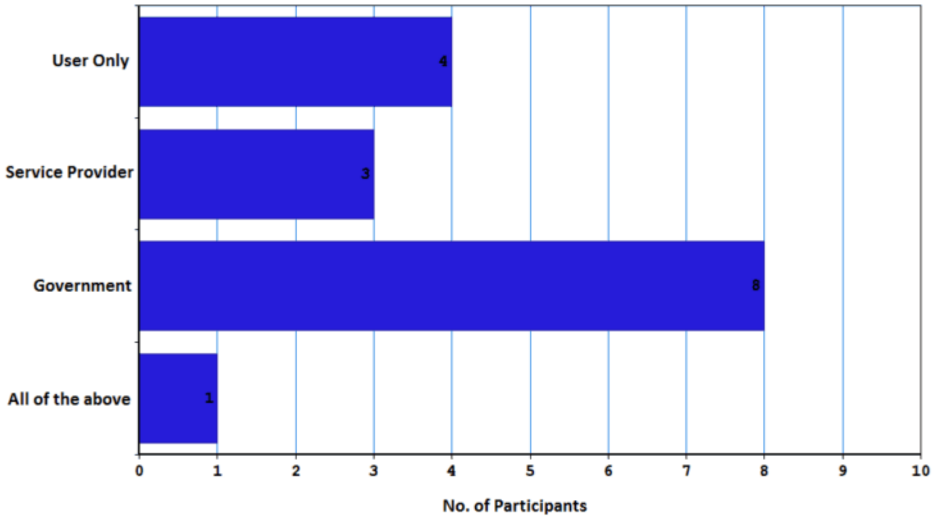


Figure 4.8: Do participants feel that cloud computing is inherently risky

Only four participants believed that they had sole access to their online data. The remainder of the group believed that either the service provider or the government where the data resides could gain access to their data.

These findings suggest that the majority of the group do not assume privacy with regards to their online data.

4.4.4 Location of data

The location of where the individuals' data is stored is relevant because each country has different security and privacy laws that are applied to individuals' data (Hayes, 2008). Regarding the location of the storage of data, no participant could say for certain that they knew in what country they data was stored. The majority guessed that it was in the USA and for providers such as DropBox and iCloud this is correct, but for Microsoft and Google the data for European users is actually stored in their EMEA data centers, both of which are in Citywest, Dublin.

The participants were then asked, given the option, where they would prefer to have their data stored. Figure 4.9 below shows a representation of the responses and as can be seen, a majority of 10 said that they would prefer to have their data stored in Ireland while a further five said any location within the European Union would be acceptable. Two participants responded that they had a preference for the USA and the reasons given ware that it the belief was that they would not be able to use the same provider if they did not want their data located in the USA. Only one participant responded that they did not care where their data was located, implying that they had a possible low expectation of privacy.

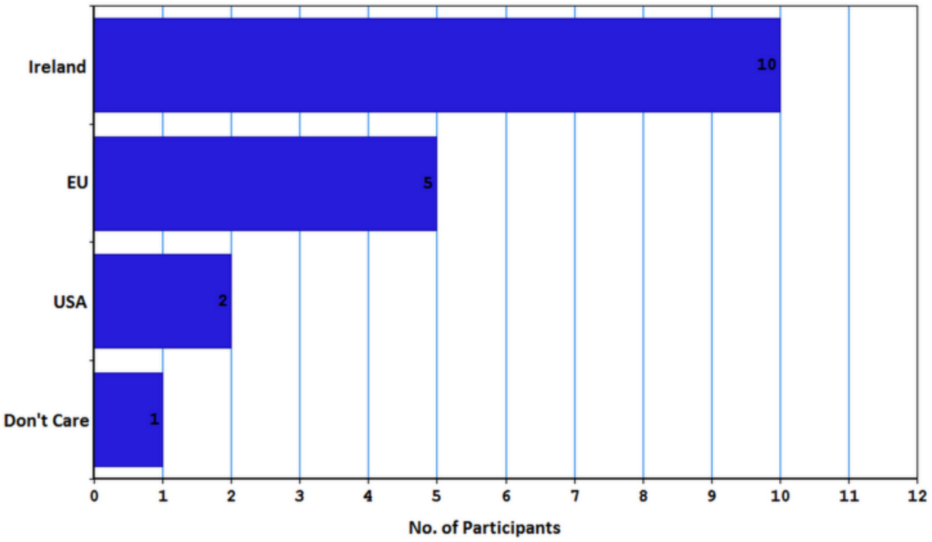


Figure 4.9: Preferred location of cloud data

The location where the users' data is stored is important as privacy laws vary from country to country. As discussed in the previous chapter, the laws in the EU are more protective of an individual's privacy rights when compared to the USA and the 'Safe Harbor' programmed was introduced to offer more protection to EU citizens, but it is not without its critics such as a publication from Connolly (2008) who cited negative findings in reports from 2002 and 2004. There were a total of 20 Dropbox and iCloud accounts used by the group of participants so this situation directly affects the owners of those accounts.

There was a general feeling of acceptance from these participants regarding having their data stored in the USA and the lower privacy regulations that at in force there. Some participants even quoted the Patriot Act and the fact that this gives the US government more freedom to access individual's data, but the acknowledgement was that this is simply a part of using a cloud data storage service that is hosted in the USA.

4.5 Summary

This chapter presented the findings and analysis from the data collected as part of the research. The primary source of data was through the use of semi-structured interviews along with associated short observation sessions with a total of 18 participants. A pragmatic philosophy approach was selected to interpret the data as this allows for mixed quantitative and qualitative methods and helps to reveal greater insights.

The participants of the study were from the ages of 18 to 65, of varying technical backgrounds and employed in different sectors. All apart from one of the participants used computing data storage. Background information such as smartphone and tablet ownership was obtained and how these applied to cloud data usage. The majority that used cloud computing had explicitly setup their own accounts but through the observation sessions, it was found that some had been setup inadvertently when the individual had setup a smartphone or tablet.

The participants were quizzed on topics such as which cloud computing data storage providers they used and if they accessed this data from multiple locations such as home and work. They were also asked about what tools they used to access their online data and also if they would consider paying for a cloud data storage service.

General security behavior was discussed to get an overview of each individual’s perception and practice. The issue of privacy was also discussed and how it related to cloud data storage, in particular the physical location of where each individuals cloud data provider stored their data.

5. Conclusions and Future Work

5.1 Introduction

This chapter presents the conclusions form the research that was carried out as part of the study and also demonstrates that the research questions have been answered. Additionally, there will be a discussion on what this research is claiming along with any new or interesting findings. The generalisibility of the research results are also highlighted along with the associated limitations. Finally, any possible future directions for research in this area are outlined.

5.2 Research Claims

The following is a summary of the findings of the research:

- 94% of the group used cloud data storage
- Only one of the group paid for cloud data storage
- The majority of the group explicitly created their cloud data storage accounts
- 83% of the participants had multiple cloud data storage accounts
- All participants had a high-speed internet connection
- All of the participants owned a smartphone while 78% owned a tablet
- 72% of participants accessed their cloud data from home and work\university
- Only one participant followed guidelines for strong password usage
- 39% followed weak or very weak guidelines. 56% followed moderate guidelines
- Only 11% of the group considered cloud computing to be risky or insecure
- 61% of the group believed that there was low risk in using cloud data storage
- None of the participants could state where their data was being stored
- 56% said that they would prefer Ireland while 28% had a preference for the EU

This research determined that amongst the group of participants that took part in the interview and observation sessions, cloud data storage was a prevalent technology with a high rate of adoption and various uses amongst the group members.

There was a very low rate of paid usage but it was found that a large majority of participants simply did not know the costs associated with cloud data storage and some commented that they would be more open to paying for a service once they had a better understanding of the costs.

It was found that security in general was not implemented to a high degree amongst the group, although a moderate level of security practice was observed as being the most common. This situation translated to cloud computing data storage. A very low number of participants of this research actually found using cloud computing for data storage a risky mechanism. This contradicts most academic research on the topic of cloud computing in general such as Subashini and Kavitha (2011), which lists security as the main reason for individuals not adopting these services.

Finally, although some participants knew about the privacy concerns of using cloud computing, particularly when using a data center in the USA instead of the EU, the general consensus amongst the group was that the participants were unconcerned about this and were more interested in the quality of the provider(s) or the amount of free storage space that they could obtain as part of the service.

5.3 Demonstration that the Research Question has been answered

The objective of this research was to examine levels of awareness of security and privacy risks among users of computers and mobile devices. This objective formed the basis for the following research questions:

1. *How do users of computers and mobile devices currently utilise cloud data storage solutions and do they pay for such solutions?*

The output of this research showed that a broad cross-section of users had many different uses for cloud data storage ranging from basic document storage, centrally locating music libraries, sharing files with others and also for the backup of devices. Additionally, the vast majority had apps installed on their smartphones and\or tablets which allowed them to access their data while mobile. The perception was that cloud data storage is not difficult to use and the main vendors such as Dropbox, Apple, Microsoft and Google have all relatively mature products that are easy to operate for the end user. This was reflected in the large usage of these services amongst the group.

Only one of the group currently paid for cloud data storage with Google as their selected provider. After discussing the actual costs with others, more said that they would consider paying for the larger capacity that would be made available and also for the extra functionality such as being able to set folder permissions on shared folders with Dropbox. This is an important point; could the providers be doing more to inform the current users of their free accounts exactly how much the paid version would cost and what the benefits would be? Would this result in a higher uptake of paid versions of the services by end users? Broadly speaking, the group interviewed as part of this study simply viewed cloud data storage as a free service, akin to email services such as Gmail and Hotmail.

Another important finding relating to this research question was that participants that had graduated from college or university were more comfortable with using the technology and were the most carefree amongst the group. There was almost a complacent approach as cloud data storage was treated no differently than local storage such as a hard drive on a computer or a usb flash drive. It was considered the norm to use cloud data storage and this had developed from a strong encouragement to use it as a sharing platform whilst they were students.

Most people signed up for their cloud data provider explicitly. Only two did not realise that they had an account and the assumption was that they registered an account during the steps of setting up a smartphone.

2. *What is the perception among users of potential security and privacy risks associated with using cloud data storage solutions?*

As part of this research, participants were interviewed about their perceived security and privacy concerns that are associated with the use of cloud data storage solutions and how this directly affected them. Additionally, observation sessions were employed in order to capture any additional information that the participants may have inadvertently omitted or otherwise provided incorrect information. The data gathered as a result of this showed that across the group, users were generally not aware of the potential security and privacy risks of utilizing cloud data storage solutions.

There were exceptions to this and a small number of the group were fully aware of the security and privacy risks. Although most could not say for certain in what country their online data was stored, most assumed that it was in the USA because their service provider was an American company and because of this, the US government and\or the service provider could gain access to their data. In reality, the location of this data was only correct for some providers but it transpired that this was immaterial because the underlying assumption was that the service was unsafe and people would take steps to mitigate against this by not storing sensitive or confidential data online.

Although security remains one of the most oft cited reasons for not adopting cloud computing services (Subashini and Kavitha, 2011), the opposite was discovered as part of this research. Only a single participant out of a total of 18 did not use cloud computing for data storage and from those that did, there was general acceptance of if it's usage in their everyday computing lives. Broadly speaking, the participants of this study had a very casual attitude towards using cloud data storage.

Because of research such as Loske et al. (2014) highlighting that the perception of security is crucial for the adoption of a technology such as cloud data storage, it was expected that people would be concerned with the potential security and privacy risks that are associated with using cloud data storage solutions. In the majority, it was found that this perception did not hold true with this research. Some participants simply ignored the security and privacy risks and were unconcerned while others accepted these risks and were still happy to use the technology, but in some situations the risks were mitigated against by the avoidance of storing sensitive or confidential data with their cloud data provider.

Research such as Loske et al. (2014) highlights that the perception of security is crucial for the adoption of a technology such as cloud data storage. This was directly relevant to this study because it was necessary to show the individual's perception and expectation of security versus their actual implementation. 61% of the group stated that they did not consider using cloud computing for data storage as a risk. Only 11% said that they felt that it was inherently risky, one of this small group does not use cloud data storage solutions in any form because of this.

5.3.1 Research objective conclusion

The objective of this research was to examine levels of awareness of security and privacy risks among users of cloud data storage solutions on their computers and mobile devices. In general, it was found that participants were not aware of these risks but still used cloud data storage solutions nevertheless. The majority were also unconcerned with these security and privacy risks but did not store confidential or sensitive data with their cloud data storage provider.

5.4 Limitation of the research

This study was very broad. Even with a group of 18 it was difficult to have tightly focused results that may have been more meaningful if it was a smaller group from a similar background such as age or technical background.

Additionally, cloud computing is a dynamic topic and the technology is evolving rapidly. This should be taken into account when reading this research paper and some of the findings may not still hold true. For example, it is stated within this document that Dropbox had announced that they had 300 million users in the latter half of 2014. It is rumoured that Dropbox are now (July 2015) close to announcing that they have reached 400 million users.

5.5 Areas for future research

It was found during the literature research process that there were very few studies in the area of personal cloud data storage usage. The majority found were either in general cloud computing usage or on focused topics such as security or Enterprise File Synchronisation and Sharing (EFSS) systems. There is certainly an opportunity for additional research on a similar topic to the usage and risks associated with personal cloud data storage.

As suggested in the limitation section previously, it would be worth considering performing a similar area of research but this time focusing on a single group of participants. For example, this focus could be based on age, technical background or sector of employment. This would allow the researcher to focus on a narrower set of results and perhaps result in more concise findings.

As this research was focused on the greater Dublin area, a study based on a different geographical location or wider region would be interesting to see how usage and perception to risk of using cloud data storage solutions differ. This would be of particular interest where high speed broadband might not be necessarily readily available for example.

Finally, although the underlying technologies are very similar, there are many vendors to choose from and each have slightly different options available to the end user. It could be considered a benefit if a study could be focused on one particular vendor as this would remove some of the discrepancies that arose when comparing products that were not necessarily exactly alike.

The research approach that was employed during this study worked well; having an interview alone would not have gathered all of the required data necessary for this report. The observation sessions collected additional information from participants that they might not necessarily have known such as how security conscious were they with the use of passwords and if any cloud data storage accounts had been created that they were not aware of. Therefore, it would be recommended that the same process could be used in any similar future study.

References

Apple. 2015. Backup and restore your iPhone, iPad, or iPod touch using iCloud or iTunes - Apple Support. [Online] Available at: https://support.apple.com/en-ie/HT203977 [Accessed 7 March 2015].

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stocia, I. and Zaharia, M. 2010. A View of Cloud Computing. Communications of the Acm, 53**,** 50-58.

Biddick, M. 2008. Time To Think About Cloud Computing. InformationWeek. [Online] Available at: http://www.informationweek.com/cloud/software-as-a-service/time-to-think-about-cloud-computing/d/d-id/1073198? [Accessed 11 February 2015].

Bittman, T. J. 2012. Mind the Gap: Here Comes Hybrid Cloud. Gartner Blog Network. [Online] Available at: http://blogs.gartner.com/thomas_bittman/2012/09/24/mind-the-gap-here-comes-hybrid-cloud/ [Accessed 10 February 2015].

Bergen, A., Coady, Y. and McGreer, R. 2011. Client Bandwidth: The Forgotten Metric of Online Storage Providers. 2011 Ieee Pacific Rim Conference on Communications, Computers and Signal Processing (Pacrim)**,** 543-548.

Bryman, A. and Bell, E. 2011. Business Research Methods, 3rd edition. Oxford University Press.

Brynjolfsson, E. and Jordan, J. 2010. Cloud computing and electricity: beyond the utility model. Communications of the ACM 53(5), pp. 32-34.

Business Insider. 2014 Number of users per Cloud Storage Service. [Online] Available at: http://www.businessinsider.com.au/cloud-services-explain-win-users-2014-8 [Accessed 14 March 2015].

Buyya, R., Yeo, C. S., Venugopal, S., Broberg, J. and Brandic, I. 2009. Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility. Future Generation Computer Systems-the International Journal of Grid Computing and Escience, 25**,** 599-616.

Carlin, S., and Curran, K. 2011. Cloud Computing Security 2 Cloud Architecture 3 Cloud Deployment Models. International Journal of Ambient Computing and Intelligence, 3(1), 38–46.

Carroll, M., Van Der Merwe, A., and Kotzé, P. 2011. Secure cloud computing: Benefits, risks and controls. 2011 Information Security for South Africa - Proceedings of the ISSA 2011 Conference.

Chowdhry, A. 2014. Dropbox Drops Price Of 1TB Storage To $9.99 Per Month. Forbes. [Online] Available at: http://www.forbes.com/sites/amitchowdhry/2014/08/27/dropbox-drops-price-of-1tb-storage-to-9-99-per-month/ [Accessed 7 March 2015].

Cohen, R. 2013. The Cloud Hits the Mainstream: More than Half of U.S. Businesses Now Use Cloud Computing - Forbes. [Online] Available at: http://www.forbes.com/sites/reuvencohen/2013/04/16/the-cloud-hits-the-mainstream-more-than-half-of-u-s-businesses-now-use-cloud-computing/ [Accessed 5 February 2015].

Connolly, C. 2008. The US Safe Harbor - Fact or Fiction? Privacy Laws and Business International, 96, 26–32.

Cunningham, Scott M. 1967. The Major Dimensions of Perceived Risk in Risk Taking and Information Handling in Consumer Behavior. Harvard University Press, 82 – 108.

Drago, I., Mellia, M., M Munafo, M., Sperotto, A., Sadre, R., and Pras, A. 2012. Inside Dropbox. The 2012 ACM Conference, 481.

Easterby-Smith, M., Thorpe, R. and Jackson, P. R. 2008. Management research, 3rd edition London, SAGE Publications.

European Commission. 2012. Unleashing the Potential of Cloud Computing in Europe. [Online] Available at: http://euapm.eu/wp-content/uploads/2013/04/STAFF-WORKING-DOCUMENT-Unleashing-the-Potential-of-Cloud-Computing-in-Europe.pdf [Accessed 2 February 2015].

Export.gov. 2012. U.S.-EU Safe Harbor Homepage. [Online] Available at: http://export.gov/safeharbor/eu/eg_main_018365.asp [Accessed 9 March 2015].

Forrester. 2011. Sizing the Cloud. [Online] Available at: https://www.forrester.com/Sizing+The+Cloud/fulltext/-/E-RES58161 [Accessed 7 March 2015].

Forrester. 2012. Understand the true cost of cloud services. [Online] Available at: http://www.forrester.com/Understand+The+True+Cost+Of+Cloud+Services/fulltext/-/E-RES61608. [Accessed 28 January 2015].

Gartner. 2012. Gartner Says That Consumers Will Store More Than a Third of Their Digital Content in the Cloud by 2016. [Online] Available at: http://www.gartner.com/newsroom/id/2060215 [Accessed 8 March 2015].

Gartner. 2014a. Gartner Identifies the Top 10 Strategic Technology Trends for 2015. [Online] Available at: http://www.gartner.com/newsroom/id/2867917 [Accessed 2 March 2015].

Gartner. 2014b. Magic Quadrant for Enterprise File Synchronization and Sharing. [Online] Available at: https://www.gartner.com/doc/2788017/magic-quadrant-enterprise-file-synchronization [Accessed 7 March 2015].

Geel, M. 2013. Cloud Storage: File Hosting and Synchronisation 2.0. [Online] Available at: https://www.vis.ethz.ch/de/visionen/pdfs/2012/visionen_2012_3.pdf?end=15&start=11 [Accessed 7 March 2015].

Google Drive Support. 2015. How to share - Drive Help. [Online] Available at: https://support.google.com/drive/answer/2494822?hl=en [Accessed 7 March 2015].

Gong, C., Liu, J., Zhang, Q., Chen, H., and Gong, Z. 2010. The characteristics of cloud computing. In Proceedings of the International Conference on Parallel Processing Workshops, 275–279.

Gordon, W. 2014. 5 Million Online Passwords Leaked, Check Yours Now. [Online] Available at: http://lifehacker.com/5-million-gmail-passwords-leaked-check-yours-now-1632983265 [Accessed 3 March 2015].

Grossman, R. L. 2009. The case for cloud computing. IT Professional, 11, 23–27.

Guest, G., Namey, E. and Mitchell, M. L. 2013. Collecting Qualitative Data: A Field Manual for Applied Research. Sage Publications

Hamdaqa, M. and Tahvildari, L. 2012. Cloud Computing Uncovered: A Research Landscape. In: MEMON, A. (ed.) Advances in Computers, Vol 86. San Diego: Elsevier Academic Press Inc.

Hayes, B. 2008. Cloud Computing. Communications of the ACM, 51, 9-11.

Hong, K. 2014. Dropbox Reaches 300M Users. [Online] Available at: http://thenextweb.com/insider/2014/05/29/dropbox-reaches-300m-users-adding-100m-users-just-six-months/ [Accessed 2 March 2015].

Internet Society. 2014 Global Internet Report 2014. [Online] Available at: https://www.internetsociety.org/sites/default/files/Global_Internet_Report_2014_0.pdf [Accessed 6 March 2015].

Kandukuri, B. R., Paturi, V. R. and Rakshit, A. 2009. Cloud Security Issues. 2009 Ieee International Conference on Services Computing, 517-520.

Kao, T. McClure, T. Oltsik, J. 2012. Online File Sharing and Online Collaboration: Information Security Challenges and Requirements - ESG Research - Enterprise Strategy Group. [Online] Available at: http://www.esg-global.com/research-briefs/online-file-sharing-and-collaboration-security-challenges-and-requirements/ [Accessed 9 March 2015].

Kaplan, B. and Duchon, D. 1988. Combining qualitative and quantitative methods in information systems research: a case study. MIS Quarterly, 571-586.G.

Kaufman, L. M. 2009. Data Security in the World of Cloud Computing. Ieee Security & Privacy, 7, 61-64.

Kepes, B. 2011. Moving your Infrastructure to the Cloud. Diversity. [Online] Available at: http://diversity.net.nz/wp-content/uploads/2011/01/Moving-to-the-Clouds.pdf [Accessed 7 April 2015].

Kesan, J. P., Hayes, C. M. & Bashir, M. N. 2013. Information Privacy and Data Control in Cloud Computing: Consumers, Privacy Preferences, and Market Efficiency. Washington & Lee Law Review, 70, 341-472.

Kennedy, J. 2014. NYC judge orders Microsoft to hand over data stored on Irish servers. Silicon Republic. . [Online] Available at: http://www.siliconrepublic.com/enterprise/item/37847-nyc-judge-orders-microsoft/ [Accessed 17 March 2015].

Lewis, D. 2014. iCloud Data Breach: Hacking And Celebrity Photos. . [Online] Available at: http://www.forbes.com/sites/davelewis/2014/09/02/icloud-data-breach-hacking-and-nude-celebrity-photos/ [Accessed 9 March 2015].

Lin, A. and Chen, N. C. 2012. Cloud computing as an innovation: Perception, attitude, and adoption. International Journal of Information Management, 32, 533-540.

Loske, A., Widjaja, T., Benlian, A. and Buxmann, P. 2014. Perceived IT security risks in cloud adoption: the role of perceptual incongruence between users and providers. Twenty Second European Conference on Information Systems, Tel Aviv 2014.

Macpherson, S. 2013. How Reliable is Dropbox? Uptime Finally Revealed: Digital First. [Online] Available at: http://www.digitalfirst.com/2013/03/28/dropbox-availability-finally-revealed/ [Accessed 1 March 2015].

Marston, S., LI, Z., Bandyopadhyay, S., Zhang, J. and Ghalsasi, A. 2011. Cloud computing - The business perspective. Decision Support Systems, 51, 176-189.

Martens, B. and Teuteberg, F. 2012. Decision-making in cloud computing environments: A cost and risk based approach. Information Systems Frontiers, 14, 871-893.

Mather, T. Kumaraswamy, S. and Latif, S. 2009. Cloud security and privacy: an enterprise perspective on risks and compliance. O'Reilly Publishing

Matthews, B. and Ross, L. 2010. Research Methods: A practical guide for the social sciences, Pearson Longman publishing.

Microsoft. 2012. Steven Sinofsky, Steve Ballmer, Julie Larson-Green, and Michael Angiulo: Windows 8 Launch. Microsoft News Centre. [Online] Available at: http://news.microsoft.com/2012/10/25/steven-sinofsky-steve-ballmer-julie-larson-green-and-michael-angiulo-windows-8-launch [Accessed 21 February 2015].

Mikkikineni, R. and Sarathy, V. 2009. Cloud Computing and the Lessons from the Past. 10th International Conference on Intelligent Tutoring Systems, Jun 14-18 Jun 29-Jul 01 2010 Pittsburgh, PA

Mohamed, A. A history of cloud computing. 2009. [Online] Available at: http://www.computerweekly.com/feature/A-history-of-cloud-computing [Accessed 05 January 2015].

NIST. 2012. Cloud Computing: A Review of Features, Benefits, and Risks, and Recommendations for Secure, Efficient Implementations. Itl. [Online] Available at: http://csrc.nist.gov/publications/nistbul/june-2012_itl-bulletin.pdf [Accessed 2 February 2015].

OFCOM. 2014. Techie teens shaping communications. [Online] Available at: http://consumers.ofcom.org.uk/news/cmr-2014 [Accessed 23 July 2015].

Ohlman, B., Eriksson, A. and Rembarz, R. 2009. What Networking of Information Can Do for Cloud Computing. 10th International Conference on Intelligent Tutoring Systems, Jun 14-18 Jun 29-Jul 01 2010 Pittsburgh, PA

Oxford English Dictionary Online. 2010. Oxford University Press. [Online] Available at: http://dictionary.oed.com [Accessed 23 April 2015].

Pew Research Centre. 2008. Use of Cloud Computing Applications and Services. [Online] Available at: http://www.pewinternet.org/2008/09/12/use-of-cloud-computing-applications-and-services/ [Accessed 2 April 2015].

Quick, D. and Choo, K. K. R. 2013a. Digital droplets: Microsoft SkyDrive forensic data remnants. Future Generation Computer Systems-the International Journal of Grid Computing and Escience, 29, 1378-1394.

Quick, D. and Choo, K. K. R. 2013b. Dropbox analysis: Data remnants on user machines. Digital Investigation, 10, 3-18.

Rackspace.com. 2013. Understanding the Cloud Computing Stack: SaaS, PaaS, IaaS. Rackspace Knowledge Centre. [Online] Available at: http://www.rackspace.com/knowledge_center/whitepaper/understanding-the-cloud-computing-stack-saas-paas-iaas [Accessed 13 February 2015].

Regalado, A. 2011. Who Coined “Cloud Computing”? MIT Technology Review. [Online] Available at: http://www.technologyreview.com/news/425970/who-coined-cloud-computing/ [Accessed 4 February 2015].

Remenyi, D. and Williams, B. 1998. Doing research in business and management: an introduction to process and method, Sage Publications Ltd.

Ren, K., Wang, C. and Wang, Q. 2012. Security Challenges for the Public Cloud. Ieee Internet Computing, 16, 69-73.

Ristenpart T., Tromer, E., Shacham, H. and Savage, S. 2009. Hey, You, Get Off of My Cloud: Exploring Information Leakage in Third-Party Compute Clouds. Ccs'09: Proceedings of the 16th Acm Conference on Computer and Communications Security, 199-212.

Sakr, S., Liu, A. N., Batista, D. M. and Alomari, M. 2011. A Survey of Large Scale Data Management Approaches in Cloud Environments. Ieee Communications Surveys and Tutorials, 13, 311-336.

Saunders, M., Lewis, P. and Thornhill, A. 2012. Research Methods for Business Students, 5th edition. Prentice Hall publishing.

Shi, X., Jiang, H., He, L., Jin, H., Wang, C., Yu, B. and Chen, X. 2013. Developing an optimized application hosting framework in Clouds. Journal of Computer and System Sciences, 79, 1214-1229.

Slatman, H. 2013. Opening Up the Sky: A Comparison of Performance-Enhancing Features in SkyDrive and Dropbox. 18th Twente Student Conference on IT.

Smith, G. 2012. The top 10 apps being blacklisted in the enterprise - TechRepublic. [Online] Available at: http://www.techrepublic.com/blog/10-things/the-top-10-apps-being-blacklisted-in-the-enterprise [Accessed 10 March 2015].

Subashini, S. and Kavitha, V. 2011. A survey on security issues in service delivery models of cloud computing. Journal of Network and Computer Applications, 34, 1-11.

Sullivan, D. 2014. PaaS Providers List: Comparison And Guide. [Online] Available at: http://www.tomsitpro.com/articles/paas-providers,1-1517.html [Accessed 5 February 2015].

Takabi, H., Joshi, J. B. D. and Ahn, G.-J. 2010. Security and Privacy Challenges in Cloud Computing Environments. Ieee Security & Privacy, 8**,** 24-31.

Tate, A. R. 2014. HIPAA and cloud computing: What you need to know – IBM Thoughts on Cloud. [Online] Available at: http://thoughtsoncloud.com/2014/03/hipaa-cloud-computing-healthcare-compliance [Accessed 3 March 2015].

United States Computer Readiness Team. 2013. Choosing and Protecting Passwords, Security Tip (ST04-002). [Online] Available at: https://www.us-cert.gov/ncas/tips/ST04-002 [Accessed 26 March 2015].

Wang, C., Wang, Q., Ren, K., Cao, N. & Lou, W. J. 2012. Toward Secure and Dependable Storage Services in Cloud Computing. Ieee Transactions on Services Computing, 5**,** 220-232.

Wang, L., Tao, J., Kunze, M., Castellanos, A. C., Kramer, D., and Karl, W. 2008. Scientific cloud computing: Early definition and experience. In Proceedings - 10th IEEE International Conference on High Performance Computing and Communications, HPCC 2008, 825–830.

Waters, R. 2014. Dropbox steps up international expansion - FT.com. Financial Times. [Online] Available at: http://www.ft.com/cms/s/0/29000be6-1021-11e4-80b1-00144feabdc0.html#axzz3T242oxri [Accessed 23 March 2015].

Wohlsen, M. 2014. Dropbox Slashes Its Price as the Cost of a Gigabyte Nears Zero. Wired. [Online] Available at: http://www.wired.com/2014/08/dropboxs-plan-to-stay-relevant [Accessed 21 March 2015].

Yang, K. and Jia, X. H. 2012. Data storage auditing service in cloud computing: challenges, methods and opportunities. World Wide Web-Internet and Web Information Systems, 15**,** 409-428.

Yin, R, K. 2013, Case Study Research: Design and Methods, 5th edition. Sage Publications.

Zhang, Q., Cheng, L., & Boutaba, R. 2010. Cloud computing: State-of-the-art and research challenges. Journal of Internet Services and Applications, 1, 7–18.

Appendices

Appendix 1: Information Sheet for Prospective Participants

TRINITY COLLEGE DUBLIN

INFORMATIONSHEETFORPROSPECTIVEPARTICIPANTS

LEAD RESEARCHER: Dean O'Gorman

BACKGROUND OF RESEARCH: The purpose of the research is to gain an insight into whether users of computers and devices such as smartphones and tablets use cloud data storage and if they are aware of the associated potential risks associated with this.

DETAILS OF RESEARCH PROCESS: The researcher requests an interview of participants followed by a period of observation. The interview will focus on areas such as perceived level of technical understanding, if the participant uses cloud data storage and if the participant is aware of associated security and privacy issues. An observation session will then take place to investigate if the participant is using cloud data storage without being aware of it and if this data is being replicated to additional devices such as smartphones or tablets. The interview and observation session is expected to last for a total of approximately 1 hour.

Several participants in this research study are friends and work colleagues of the principal investigator and the principal investigator is taking advantage of his existing personal and professional relationships in order to make progress in this research study. Participation in this research is completely voluntary and the interviewee has the right to withdraw and to omit individual responses without penalty. There are no identified risks to the participant and there may be a learning process regarding cloud data storage that may be of benefit to the participant.

At the request of the participant, a debriefing can take place after the sessions are complete but this will not be by default.

It is not intended to use audio or video recordings during the interview or observation sessions, but extensive notes will be taken. Any information or data which is obtained during this research with me will be treated confidentially. Information will be presented in summary format and individual participants will not be identified in any publications or presentations. In the extremely unlikely event that illicit activity is reported I will be obliged to report it to appropriate authorities.

Appendix 2: Informed Consent Form

TRINITY COLLEGE DUBLIN

INFORMED CONSENT FORM

LEAD RESEARCHER: Dean O'Gorman

BACKGROUND OF RESEARCH: The purpose of the research is to gain an insight into whether users of computers and devices such as smartphones and tablets use cloud data storage and if they are aware of the associated potential risks associated with this.

PROCEDURES OF THIS STUDY: This study will be in the form of an interview so the researcher can gain an insight into what (if any) forms of cloud data storage the interviewee is using. The interview will also include high level security questions such as the risks associated with storing personal data in the cloud and the location of data.

Additionally, observation will be used with the interviewee to see how they are using their computer but more specifically, their mobile device. The expectation is that some may be using cloud data storage and not be aware of it. Each session (interview and observation) will last approximately 1 hour.

PUBLICATION: The outcomes of this research will be published as part of a dissertation for and MSc Management of Information Systems, Trinity College.
Individual results may be aggregated anonymously and research reported on aggregate results.

DECLARATION:

- I am 18 years or older and am competent to provide consent.
- I have read, or had read to me, a document providing information about this research and this consent form. I have had the opportunity to ask questions and all my questions have been answered to my satisfaction and understand the description of the research that is being provided to me.
- I agree that my data is used for scientific purposes and I have no objection that my data is published in scientific publications in a way that does not reveal my identity.
- I understand that if I make illicit activities known, these will be reported to appropriate authorities.
- I freely and voluntarily agree to be part of this research study, though without prejudice to my legal and ethical rights.
- I understand that I may refuse to answer any question and that I may withdraw at any time without penalty.
- I understand that my participation is fully anonymous and that no personal details about me will be recorded.
- I understand that if I or anyone in my family has a history of epilepsy then I am proceeding at my own risk.
- I have received a copy of this agreement.

PARTICIPANTS NAME:

PARTICIPANTS SIGNATURE:

DATE:

Statement of investigator's responsibility: I have explained the nature and purpose of this research study, the procedures to be undertaken and any risks that may be involved. I have offered to answer any questions and fully answered such questions. I believe that the participant understands my explanation and has freely given informed consent.

RESEARCHERS CONTACT DETAILS: Dean O'Gorman
ogormade@tcd.ie
087 8221289

INVESTIGATOR'S SIGNATURE:

Date:

Appendix 3: Interview questions

TRINITY COLLEGE DUBLIN

INTERVIEW QUESTIONS

The primary focus of the interview process is to gain an understanding of the level of technical ability of the participant and if they use cloud data storage. It is then intended to investigate further based on this to ascertain if they are aware of associated risks such as the location that this data is stored and if it is easily accessible on additional devices. The interview process will be semi -structured and the data collected will undergo qualitative analysis.

In order to help with forming conclusions, personal information such as age range and level of education will be asked, but it will be repeated again before the interview process begins that all questions are voluntary.

To follow up the interview process, a short observation session will then take place whereby the researcher will ask the participant to demonstrate how they access their cloud data storage account (if they have one). Additionally, if the participant has a smartphone, they will be asked if they knowingly synchronise data from their handset to a cloud data provider. The expectation is that not all users are aware of this functionality and this is why the observation session is necessary.

SECTION 1: Basic background information

1. What is your age range?
2. What is your level of education?
3. Do you now or have you ever worked in IT or telecommunications?
4. Do you use the internet?
5. How many hours a day would you normally spend on the internet?
6. What do you mainly use the internet for?

SECTION 2: Technical background

1. What do you believe your level of technical of competency to be?
2. Do you use a computer at work or at home?
3. Do you own a Smartphone? What is the make\model?
4. Do you own a Tablet? What is the make\model?
5. What is your internet connection speed (if known) at home?

6. Who is your Internet Service Provider?

7. Do you stream TV\Movies using services such as Netflix or Amazon Prime?

8. Do you own a Smart TV? Do you use services such as the RTE Player or BBC iPlayer to watch programs?

SECTION 3: Cloud data usage

1. Do you have a cloud data storage account? If yes, what is the service provider? E.g. Dropbox, iCloud, OneDrive, Box, etc. Can you go into detail about what you use each service for?

2. Do you know what amount of data storage that you have with each provider?

3. What data do you store online? E.g. photos, documents, etc. Do you store any sensitive data online?

4. Do you currently pay for cloud data storage? Would you consider paying for this in the future? Are you aware of the cost of 1TB of online data storage?

5. Do you synchronise data from your smartphone to your computer? Is this then synchronised (or directly synchronised) to a cloud data storage provider?

6. Have you ever lost your smartphone that had data synchronization enabled? Did you encounter issues with somebody accessing your data?

7. Do you use a desktop sync tool to automatically synchronise data from your cloud provider to your desktop computer(s) or laptop(s)?

8. Has someone ever gained access to your synchronised data on a desktop or PC? What would happen if your computer was stolen? Are you aware of the potential risks and how to mitigate against them?

9. How many devices do you access your cloud data from? Do you synchronise a copy off all of your data to each device?

10. Do you share data which is stored online with anyone? Do you ever email large amounts of data to anyone?

11. Have you ever run out of online space?

SECTION 4: Security and Privacy

1. How security conscious are you? Do you know what a ‘strong’ password refers to and how strong are your passwords? Do you use different passwords for all of your online accounts or a single username and password for all?

2. Does data privacy matter to you? Do you know if the cloud data storage provider can gain access you your data without your consent? Do you know if governments can gain access to your data without your consent?

3. Do you store all of your data locally on a computer (i.e. do not knowingly use cloud data storage)? Do you use external hard drives to store data? Are they stored securely?

4. Do you know what two-step verification means? Is it available with your provider? Do you have it setup?

5. Do you know what remote wipe means? Is it available with your provider? Have you ever used it?

6. Do you access your personal cloud data storage on your work\university network? Is it enabled by default or is it blocked by administrators?

7. Do you know what encryption means? Does your cloud data storage provider employ this?

8. Do you know if your account has ever been hacked? Has someone gained access your account by a PC being left unlocked?

9. Are you aware where this data is stored? Is the location important to you?

10. Would you prefer to have your data stored in Ireland, the EU or the USA for example?