

**FINAL YEAR REPORT**

**Auto-Backup System with AES 256-bit Encryption**

**By**

**VERIKO QBERT LIMBERTA**

**TP057263**

**APU3F2305CS(CYB)**

A report submitted in partial fulfilment of the requirements for the degree of

B.Sc. (Hons) Computer Science (Cyber Security)

at Asia Pacific University of Technology and Innovation

**Supervised by Ms Noris Binti Ismail**

**2nd Marker: Ms Nur Azyyati Binti Ahmad**

**09-August-2023**

**Acknowledgement**

Firstly, this investigation report contains the initial process of planning and understanding of the project better. I would like to express my greatest gratitude to Ms Noris Binti Ismail that has guided me in the making of this report. Ms Noris has helped me from the very beginning where we started off just discussing my Final Year Project (FYP), to the point of finishing this investigation report. Ms Noris has also shown huge amount of support and provided me with advice that would ultimately make this investigation report better. Ms Noris also helped me chapter by chapter making sure that I am on the path that was clear and right. Ms Noris also checked chapter by chapter to make sure that the information and requirements needed in each chapter was achieve. With this, I would once again thank Ms Noris Binti Ismail for the hard work and her efforts in making this report a better report.

I would also like to thank Ms Nur Azyyati Binti Ahmad on commenting on my work and going through my work again. Ts. Umapathy Eaganathan is another person I am thankful off. He has been giving information and have guided me through the whole FYP process and has also provided me with the resource and guidance on how to do well in this project. Lastly, I would like to thank all the individuals that took part in my survey and providing me with the necessary data to be able to conduct this research.

I would finally thank all the coffeeshops that has accommodated me throughout the journey as staying in my room had me tempted with many different stuff. Thus, going out and doing this project is more effective and efficient.

**Table of Contents**

[Chapter 1: Introduction of Study 8](#_Toc142424198)

[1.1 Background to the project 8](#_Toc142424199)

[1.2 Problem Context 9](#_Toc142424200)

[1.3 Rationale 10](#_Toc142424201)

[1.4 Potential Benefits 11](#_Toc142424202)

[1.4.1 Tangible Benefits 11](#_Toc142424203)

[1.4.2 Intangible Benefits 11](#_Toc142424204)

[1.5 Target Users 11](#_Toc142424205)

[1.6 Aim and Objectives 12](#_Toc142424206)

[1.6.1 Aim 12](#_Toc142424207)

[1.6.2 Objectives 12](#_Toc142424208)

[1.6.3 Deliverables 12](#_Toc142424209)

[1.6.4 Nature of Challenge 12](#_Toc142424210)

[1.7 Overview of this Investigation Report 13](#_Toc142424211)

[1.8 Project Plan 13](#_Toc142424212)

[Chapter 2: Literature Review 15](#_Toc142424213)

[2.1 Introduction 15](#_Toc142424214)

[2.2 Domain Research 15](#_Toc142424215)

[2.2.1 Backup System 15](#_Toc142424216)

[2.2.2 Cyber Security 16](#_Toc142424217)

[2.2.3 Ransomware 17](#_Toc142424218)

[2.2.4 Blockchain 19](#_Toc142424219)

[2.3 Similar System 20](#_Toc142424220)

[2.3.1 IDrive 20](#_Toc142424221)

[2.3.2 Storj 21](#_Toc142424222)

[2.3.3 Sia 22](#_Toc142424223)

[2.3.3 Comparison 23](#_Toc142424224)

[2.4 Summary 24](#_Toc142424225)

[Chapter 3: Technical Research 25](#_Toc142424226)

[3.1 Programming Language 25](#_Toc142424227)

[3.1.1 Python 25](#_Toc142424228)

[3.1.2 Java 26](#_Toc142424229)

[3.1.3 Programming Language Chosen 26](#_Toc142424230)

[3.2 Interactive Development Environment (IDE) Chosen 27](#_Toc142424231)

[3.3 Operating System Chosen 28](#_Toc142424232)

[3.4 Summary 28](#_Toc142424233)

[Chapter 4: System Development Methodology 29](#_Toc142424234)

[4.1 Overview 29](#_Toc142424235)

[4.2 Waterfall Methodology 29](#_Toc142424236)

[4.2.1 Background of Waterfall Methodology 29](#_Toc142424237)

[4.2.2 Overview of Waterfall Methodology 30](#_Toc142424238)

[4.3 Agile Methodology 31](#_Toc142424239)

[4.3.1 Background of Agile Methodology 31](#_Toc142424240)

[4.3.2 Overview of Agile Methodology 32](#_Toc142424241)

[4.4 Comparison of System Development Methodology 33](#_Toc142424242)

[4.4.1 Comparison Table 33](#_Toc142424243)

[4.4.2 Selected System Development Methodology and Justification 34](#_Toc142424244)

[4.5 Application of Agile Methodology on this Project 34](#_Toc142424245)

[Chapter 5: Research Methods 36](#_Toc142424246)

[5.1 Introduction 36](#_Toc142424247)

[5.1.1Research 36](#_Toc142424248)

[5.1.2 Sampling Methods 36](#_Toc142424249)

[5.1.3 Data Collection Method 37](#_Toc142424250)

[5.2 Design 38](#_Toc142424251)

[5.3 Summary 45](#_Toc142424252)

[Chapter 6: Requirements Validation 46](#_Toc142424253)

[6.1 Analysis of data 46](#_Toc142424254)

[6.1.1 Demographic Data 46](#_Toc142424255)

[6.1.3 Analysis of Data 47](#_Toc142424256)

[6.2 Summary 54](#_Toc142424257)

[Chapter 7: Conclusion 55](#_Toc142424258)

[References 56](#_Toc142424259)

[Appendices 64](#_Toc142424260)

[Log Sheets 64](#_Toc142424261)

[Log Sheet 1 64](#_Toc142424262)

[Log Sheet 2 65](#_Toc142424263)

[Log Sheet 3 66](#_Toc142424264)

[Log Sheet 4 67](#_Toc142424265)

[PPF 68](#_Toc142424266)

[Ethics Form 75](#_Toc142424267)

[Gantt Chart Semester 1 79](#_Toc142424268)

**Table of Figures**

[Figure 1 : Country Against Ransomware 18](#_Toc142025117)

[Figure 2: Python Logo 25](#_Toc142025118)

[Figure 3: Java Logo 26](#_Toc142025119)

[Figure 4: PyCharm Logo 27](#_Toc142025120)

[Figure 5: Waterfall Methodology Diagram 29](#_Toc142025121)

[Figure 6: Agile Methodology Diagram 31](#_Toc142025122)

[Figure 7: Agile Diagram for the Project 34](#_Toc142025123)

[Figure 8: Background about the Research 39](#_Toc142025124)

[Figure 9: Demographic Question 39](#_Toc142025125)

[Figure 10: Section B Background 40](#_Toc142025126)

[Figure 11: Question 1 40](#_Toc142025127)

[Figure 12: Question 2 41](#_Toc142025128)

[Figure 13: Question 3 41](#_Toc142025129)

[Figure 14: Question 4 41](#_Toc142025130)

[Figure 15: Question 5 42](#_Toc142025131)

[Figure 16: Question 6 42](#_Toc142025132)

[Figure 17: Question 7 42](#_Toc142025133)

[Figure 18: Question 8 43](#_Toc142025134)

[Figure 19: Question 9 43](#_Toc142025135)

[Figure 20: Question 10 43](#_Toc142025136)

[Figure 21: Question 11 44](#_Toc142025137)

[Figure 22: Question 12 44](#_Toc142025138)

[Figure 23: Question 13 44](#_Toc142025139)

[Figure 24: Question 14 45](#_Toc142025140)

[Figure 25: Responses 46](#_Toc142025141)

[Figure 26: Demographic Question 1 46](#_Toc142025142)

[Figure 27: Demographic Question 2 46](#_Toc142025143)

[Figure 28: Data of Question 1 47](#_Toc142025144)

[Figure 29: Data of Question 2 47](#_Toc142025145)

[Figure 30: Data of Question 3 48](#_Toc142025146)

[Figure 31: Data of Question 4 48](#_Toc142025147)

[Figure 32: Data of Question 5 49](#_Toc142025148)

[Figure 33: Data of Question 6 49](#_Toc142025149)

[Figure 34: Data of Question 7 50](#_Toc142025150)

[Figure 35: Data of Question 8 50](#_Toc142025151)

[Figure 36: Data of Question 9 51](#_Toc142025152)

[Figure 37: Data of Question 10 51](#_Toc142025153)

[Figure 38: Data of Question 11 52](#_Toc142025154)

[Figure 39: Data of Question 12 52](#_Toc142025155)

[Figure 40: Data of Question 13 53](#_Toc142025156)

[Figure 41: Data of Question 14 53](#_Toc142025157)

**Tables of Tables**

[Table 1: Project Plan 13](#_Toc142025205)

[Table 2: Similar System 23](#_Toc142025206)

[Table 3: Methodology Comparison Table 33](#_Toc142025207)

# Chapter 1: Introduction of Study

## 1.1 Background to the project

In these present times, it is expected that there are two billion of active computers in the world (Cadenas, 2023). Since these are active computers, they would most likely contain data, whether the data is important or not it is almost a must to backup the data that are available in the computer in the sense that if there is anything that happened towards the computer that would results in the data that was store to be lost, the user would have a backup to the data. Backups are the process when the individual take a data that is already in the computer or system and creating a copy of the data and use or recover the copy of the data when the original data is lost or corrupted. The users just have to access the backup data and they would be able to recover the data (*What Is Backup? (Data Backup) Comprehensive Guide - Acronis*, n.d.). Backups are important due to a various possibility of data lost such as human error, failure of hardware components in the computer, possible virus that can attack the computer and mess with the data, lost of power and natural disasters (Mathieu, 2023).

Backup system has evolved through-out the decade and has come to a point where it is more accessible. Traditional backup system works in a way that requires multiple components and moving parts to be able to back up the user’s data. By using this method, they can achieve the results of backing up the data, but as they are using moving parts and different components, it results in a complicated solution and a lot has to be done to be able to manage and protect the data itself. Traditional method of backing up the data just focuses on the ability to copy the files and have it stored. By having it this way it results in the possibility of the components to fail or corrupted which will lead to the data being useless. Not only that, old and traditional ways require a high maintenance to keep the data that was copied safe, with this the maintenance cost is high making it pricy to have this backup system. With the long run, the backup system will accumulate more data which will cost even more to back up the data (Evangelist, 2022).

The way this system will back up the data is by using blockchain technology to be able to store the data in a safe environment. Blockchain is a technology is a development used in digital transaction and data or record keeping. Blockchain is a technology that has been around since the year 2008. Starting development in 1979, Blockchain technology has evolved a lot in the process (Gondek, n.d.). In the current day, one of the most famous system or company that uses the blockchain technology is Bitcoin. What makes Blockchain the choice of storage for this backup system is, blockchain technology has proved itself to be one of the reliable ways to receive data that is accurate. Not only that, but blockchain also provides security by creating a permanent record and no one including the administrator could make changes to the data and record. Lastly the efficiency of blockchain technology is also another reason why blockchain technology was chosen to be used in the backup system itself (*What Is Blockchain Technology - IBM Blockchain | IBM*, n.d.).

## 1.2 Problem Context

**1. People Tend to Forget to Back-up Their Data**

There are many people out there that are using the internet that either do not back-up their data or do not even know what backing-up data is. Based on Backblaze website, there is user who has proper cloud back-up solution but 48% of the people that has this service does not even back-up their data and there is another 10% of people that has proper cloud back-up services that they are not even sure what they are doing or what is back-up used for. 61% of people are not confident on using the back-up system. This is a huge problem as they are more prone to cyber-attacks and once, they got attacked, they cannot do anything other than either paying the hacker and hope that they will return their data or basically just give up and know that their data that was stored was all gone and they cannot get it back (Yev, 2022).

**2. It is Annoying**

Due to the amount of data that must be back-up every day, most people might find it annoying, most people might find it annoying to have the need to always back-up and have the need to resave everything. This is understandable due to time is important. Even so, they should not just ignore all the possible threats and just carry on as there is nothing happening. These is why researcher would say that an auto-backup system is needed to allow the user to have more time on the task that they are assign with and not worry about backing-up the data that is store on their device. Not only that people find it annoying to backup due to time, but it is also annoying to find out that the backup system that they are using is not really a backup system itself. With that in mind, people get frustrated and annoy by using this third-party software (SpidersC, n.d.).

**3. More Cybercrimes are Happening**

With the current era, there is more cybercrimes happening than ever before. In the year 2023, it is expected that every 1 in 5 internet users will be affected by cybercrimes. Data breaches is one of the things that hacker will go after, with this it cost business $4.35 million in damage. Ransomware is one of the attacks that happened a lot. As cyber attackers can easily hold the user’s data and demand a large sum of cash to able to release back the data. Even if the individual has paid the Hackers, it is not a guarantee that the users will still be able to get the data back. With the more cybercrimes happening, the more data lost is happening. Users that do not backup their data are the most likely more vulnerable to the hackers than those who backup their data. This is due to the people who backup their data can just access their back-upped data and can ignore the demands from the hackers (Imber, 2023).

## 1.3 Rationale

Based on the problem that is stated above, the purpose of this project is to create an auto backup system. This is important due to the number of cybercrimes that is happening. The number of ransomware attack will keep on increasing causing data lost or the user. This is why an auto backup system is needed. The primary reason of this work is that to be able to provide user to be able to have a copy or a backup of their data to allow them to have the sense that even if an attacker has chosen them as their target, they would be in a better position to deal with it. The user does not have to blindly follow the attacker and can just access their copied or backup data and recover the data.

The secondary reason for the suggestion of this system is to implement the usage of blockchain technology to store the data itself. This is mostly due to the cause of blockchain technology has a lot of security in place for the protection of the system. Not only that, but blockchain technology also allow the fast response of data when needed, it is known that blockchain technology is currently on a uptrend causing it to be used by many types of companies in different types of industries (Zesium\_Admin, 2022).

The idea of the auto backup system first came upon when research has shown that the huge amount of people does not backup their data as they either do not have it, do have a system that automatically backup the data, or even has it but does not use it as they do not trust the auto-backup system since it is not clear. This is why the idea of the auto-backup system come into the picture (Yev, 2022).

The idea of using blockchain technology is mainly to be able to have the data being kept safely as blockchain technology has a lot and good security measures that can be implemented so that the data is safer. Blockchain technology also allows faster access to the data, this allows the user to be able to access the data faster with little to no delay as it is a unique part of the blockchain technology. They can store their data by in the form of blocks that are linked together with cryptography (Sahu, n.d.).

## 1.4 Potential Benefits

This section will go over the project's advantages from both the tangible and intangible perspectives. The term "tangible benefits" in this debate refers to quantitative, quantifiable benefits, whereas the term "intangible benefits" refers to unquantifiable benefits that can nonetheless have a significant impact on cybersecurity.

### 1.4.1 Tangible Benefits

* Minimize the disturbance from data being either breach or loss.
* Prevent the need to pay the hacker fee to be possibly get the individual’s data back.
* Maximize the security of the data that the user has.
* Minimize the chances of human error.
* Minimize the time needed for the user to have their data backup.

### 1.4.2 Intangible Benefits

* Increase trust of users towards auto-backup system
* Increase the usage of auto-backup system for users who doesn’t have an auto-backup system.
* More efficient auto-backup systems

## 1.5 Target Users

This auto-backup system will be used for Asia Pacific University. This is made available as this backup system can just be implemented to allow the university to be able to keep a copy of their student data into the system. It should not be limited to one user as organization can use the data and view the student’s information and check whether the student is from the university and their details such as grades. With the help of the backup system, the university can upload the data into the blockchain. The organization can then interact and retrieve the blocks containing the data that is required. This is then shown to the organization or anyone who is accessing it. As it is in blockchain, the data could not be modify thus, making the data reliable.(Simply Explained, 2017)

## 1.6 Aim and Objectives

### 1.6.1 Aim

The aim of this project is to design and develop an auto-backup system that enables the system to be able to automatically backup the data. The auto-backup system will be integrated with different type of security to be able to as best as possible protect the user’s data. The auto-backup system will be easy to be used by both the university and also the individuals that access it.

### 1.6.2 Objectives

Below are the objectives of this project:

* To study on the different method and current technologies that are possible to create an auto-backup system.
* To develop an auto-backup system that can be used by universities and individual that will access it.
* To implement backup system using blockchain technology to avoid tampering of data and secure.
* To test and evaluate the functionality

### 1.6.3 Deliverables

This project delivers an auto-backup system that can be used by both individual and organization. The key deliverables in this project include:

1. An auto-backup system that automatically back-up the user data from their devices.
2. Data is backup into blockchain technology to store the data.
3. Security measures such as encryption is in places for the security of the data.
4. Easy to access the data that was stored into the system.
5. A system that will be able to be used by the university to store the student details.
6. A system that is easy to be used by any individual to view the student details.

### 1.6.4 Nature of Challenge

There will be few challenges both big and small that will be faced with. One of the challenges being blockchain technology itself. This is a challenge as the researcher have not learned anything about blockchain technology and must learn it through websites and other resources to be able to implement this into the project. Another thing is the lack of freely available blockchain infrastructure, due to this, the usage of blockchain for this project will be harder. There are many other auto-backup system out there causing the user to have multiple choice and hence that is why this auto-backup system should have something special and unique to attract user to use this auto-backup system over the others.

## 1.7 Overview of this Investigation Report

In this report, the researcher will explain and research articles that will be able to help the researcher to understand the project better. Research will also be done addressing and understanding the population of their understanding and views on backup system. This report will fully go through the early process of research and understanding the project better before starting to code the project and bringing the project into a reality.

## 1.8 Project Plan

Table 1: Project Plan

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **NO.** | **TASK** | **START** | **END** | **DAYS** | **% DONE** | **STATUS** |
|  | PPF | Tue 16/05/2023 | Fri 02/06/2023 | 18 | 100% | Completed |
|  | Ethics form Fast Track | Tue 25/07/2023 | Tue 25/07/2023 | 1 | 100% | Completed |
| **1** | **Chapter 1: Introduction of Study** |  |  |  |  |  |
| 1.1 | Background to the project | Thu 22/06/2023 | Fri 23/06/2023 | 2 | 100% | Completed |
| 1.2 | Problem Context | Thu 22/06/2023 | Fri 23/06/2023 | 2 | 100% | Completed |
| 1.3 | Rationale | Thu 22/06/2023 | Sat 24/06/2023 | 3 | 100% | Completed |
| 1.4 | Potential Benefits | Thu 22/06/2023 | Sat 24/06/2023 | 3 | 100% | Completed |
| 1.5 | Target Users | Thu 22/06/2023 | Sun 25/06/2023 | 4 | 100% | Completed |
| 1.6 | Aims and Objective | Thu 22/06/2023 | Mon 26/06/2023 | 5 | 100% | Completed |
| 1.7 | Overview of Final Year Project | Thu 22/06/2023 | Tue 27/06/2023 | 6 | 100% | Completed |
| **2** | **Chapter 2: Literature Review** |  |  |  |  |  |
| 2.1 | Introduction | Tue 27/06/2023 | Tue 27/06/2023 | 1 | 100% | Completed |
| 2.2 | Domain Research | Wed 28/06/2023 | Wed 12/07/2023 | 15 | 100% | Completed |
| 2.3 | Similar system | Thu 13/07/2023 | Sat 22/07/2023 | 10 | 100% | Completed |
| 2.4 | Summary | Sat 22/07/2023 | Mon 24/07/2023 | 3 | 100% | Completed |
| **3** | **Chapter 3: Technical research** |  |  |  |  |  |
| 3.1 | Programming Language | Tue 25/07/2023 | Tue 25/07/2023 | 1 | 100% | Completed |
| 3.2 | Interactive Development Environment (IDE) Chosen | Tue 25/07/2023 | Tue 25/07/2023 | 1 | 100% | Completed |
| 3.3 | Operating System Chosen | Tue 25/07/2023 | Tue 25/07/2023 | 1 | 100% | Completed |
| 3.4 | Summary | Tue 25/07/2023 | Tue 25/07/2023 | 1 | 100% | Completed |
| **4** | **Chapter 4: System Development Methodology** |  |  |  |  |  |
| 4.1 | Overview | Wed 26/07/2023 | Wed 26/07/2023 | 1 | 100% | Completed |
| 4.2 | Waterfal Methodology | Wed 26/07/2023 | Wed 26/07/2023 | 1 | 100% | Completed |
| 4.3 | Agile Methodology | Wed 26/07/2023 | Wed 26/07/2023 | 1 | 100% | Completed |
| 4.4 | Comparison of System Development Process | Wed 26/07/2023 | Wed 26/07/2023 | 1 | 100% | Completed |
| 4.5 | Application of Agile Methodology on this Project | Wed 26/07/2023 | Wed 26/07/2023 | 1 | 100% | Completed |
| **5** | **Chapter 5: Research Methods** |  |  |  |  |  |
| 5.1 | Introduction | Thu 27/07/2023 | Thu 27/07/2023 | 1 | 100% | Completed |
| 5.2 | Design | Thu 27/07/2023 | Fri 28/07/2023 | 2 | 100% | Completed |
| 5.3 | Summary | Thu 27/07/2023 | Sat 29/07/2023 | 3 | 100% | Completed |
| 5.4 | Survey | Sat 22/07/2023 | Sat 29/07/2023 | 8 | 100% | Completed |
| **6** | **Chapter 6: Requirements Validation** |  |  |  |  |  |
| 6.1 | Analysis of Data | Sun 30/07/2023 | Tue 01/08/2023 | 3 | 100% | Completed |
| 6.2 | Summary | Sun 30/07/2023 | Tue 01/08/2023 | 3 | 100% | Completed |
| **7** | **Chapter 7: Conclusion** |  |  |  |  |  |
| 7.1 | Conclusion | Wed 02/08/2023 | Wed 02/08/2023 | 1 | 100% | Completed |

# Chapter 2: Literature Review

## 2.1 Introduction

The goal of this project is to create an auto-backup system that will be able to use by both individual and organization. To be able to do this, the comparison of multiple system that are similar or in some way helpful should be done. Not only similar system comparison but articles that help to build or develop this system should also be reviewed and looked at. This is important due to the research will allow the system to be develop better and more accurate.

## 2.2 Domain Research

### 2.2.1 Backup System

Having to back up in this era where computer stores almost all of the individual data is a very important. Having a backup is important as the user may never know when their data is lost or corrupted. By having a backup system, if the user wanted to get back the lost data or corrupted data, they could just access the backup system database where their data is stored and recovered it. The system can also backup the data as many times as the system wants and needs as the more backup the system has, the safer the data is stored. Data lost can either be by natural cost or by the user itself (Nonprofit Technology Collaboration & Baylor, 2013).

The old traditional way of backing up the data is using a disk or any other hardware devices. As this itself is not a big problem on how the disk or the hardware stored the data, the bigger problem is the data that the average user needs to store is growing more that what a disk or other hardware can hold. It is known that within the years past, the amount of storage that the storage that the disk and hardware also increase, but that also means that the data that the user will need to store is also increase. Thus, the usage of these hardware devices is good, but it has limitation (Chervenak et al., 1998).

There are many types of backup system out there. Each of them has their own unique part to them, which means there is pros and cons for each type of backup system. Currently one of the more well-known backup systems is uses the cloud storage. The other type of storage the backup system might use are, internal hard disk drives, removable storage media, blockchain. Each with their own benefit and consequences. This is why users should know what types of data that he or she is storing to be able to choose the current type of backup system (Ruggiero & A. Heckathorn, 2012).

As stated above, there are many ways to backup the data. In this research, they are focusing on tape and cloud system type of backup. Through this, the research has found out that by using cloud storage system to backup and store their data is more effective. This research is done by trying it on enterprise. They found out that the CPU and memory utilization is less overloaded meaning that the system can perform task more effectively. Though they did not able to get the accurate data, they are still able to see the different by utilizing the cloud storage. As they are not able to get the accurate data, the test that this study could be further investigated (Yarrapothu, n.d.).

### 2.2.2 Cyber Security

As we are moving into a more advanced technology area, the number of cyber attacks that happen also increases. Cyber-attacks not only happen to big cooperations, but it also happened to small or medium cooperation and also individual that uses the internet. According to Cyber security Magazine.com, around 43% of data breach happen has been linked to medium size company and 30% of small companies fallen into phishing attacks that happen. Thus, with these statistics, it is clearly shown on the number of cyber-attacks that happen not only to the big cooperation but also to the small and medium size business. Cyber-attacks can happen due to multiple reason, and according to the research paper by “Research Parks”, the `top 8 reasons also known as weakness that could cause cyber-attacks are unsecured network, communication channel that are not secure, outdated technology, bugs that could not be identify, insufficient amount of cyber-security strategy, lack of monitoring of the system, employee does not have enough training, and multiple connection points and the internet of things. (Abdumalikov, 2022)

According to “Kuwait Chapter of Arabian Journal of Business and Management Review”, the number of users of internet user of Malaysia have increase a lot. At the year 2000, the number of internet user in Malaysia is only 3,700,000 and by the year 2010 the number of internet user has increase to 16,902,600. By the year past and now 33.03 million of internet user in Malaysia at the year of 2023. This is also the reason why awareness of cyber security should also increase as if the number of internet user grow, the number of cyber attacks will also grow. Therefore, it is also the responsible of the Malaysian government to impost security measure and to spread awareness to all of the Malaysia citizen. Not only the government has to look at individual, but they also must look at organization. This is also due to the organization is also switching to internet based where they will in some way stored their data on the internet. That is why the government should have a programme that would reach the target audience (Supayah & Ibrahim, 2016).

Application and software are built properly but when they are building the application or the software, the software developer tend to forget or do not pay much attention to the security side of things, hence the application and software is having a big chance of getting infiltrated and the application or software will get compromised. Not only that, but the other reason for cyber-attacks could also be from the users itself. The way to quickly responds to the incidence or the cyber security parameters are, identify threats, identify vulnerabilities, access risk explores, establish contingency plan, respond to cyber security accident, and respond to cyber-security. As more application and software are available for individual and organization to use, they must also be careful when they are using the network, as their data may be modified or even be lost by the attackers (Ghundare et al., 2020).

As already been discuss above, the number of cyber-crimes has been increasing lately. Another way that we can help safeguard the users of the internet is by the government introducing new and stricter laws that will be able to punish the attackers fairly. As this means that the lawmakers can keep an eye out for any misused of the technology or the internet. This also allow the lawmaker to see how much work is needed to be done in the law side to be able to achieve the healthy and legal used of the internet and not allowing the attackers to use the internet illegally (Sharma & Sharma, 2023).

Cyber-attacks can also cause cyber terrorism. Cyber terrorism has almost a guarantee that the victim will cause harm to themselves in some way. One of the harms that would happen is mental injury. Cyber terrorism is a big issue as the victims not only fallen for the trap which the attacker has put in place but also the victims has lost something big to the attackers which prompt them to cause self-harm to themselves. As cyber-attacks are increasing, the number of cyber terrorisms will also increase. This is why cyber security to protect the application, software or network is very important and they must keep up on the security feature to make sure it is in the latest form so the attackers will not have a big chance to attack the users (Kalakuntla et al., 2019).

### 2.2.3 Ransomware

Ransomware is one type of attack that the attacker might do or use to the victim. Ransomwares happen when the attacker enters the victim space or computer or devices, they take away their data and ask for the victim to pay up a certain amount so that the attacker will give back their data. As the attackers are so mischief, even once after the victim pay up the amount, the attackers may not return their data and ask for more amount of money for the data to be released back. This may go on and go on until to the point where the victim will be tired and give up. About 2 to 3 percent of the population that is using the internet or has any data on the internet falls victim to ransomware in a year. With this, we can conclude that millions of users have fallen into ransomware per year. This type of attack is just one of many other types of attack that the attackers might do (Simoiu et al., 2019).

From the market report on 2023 ransomware insight, their finding shows that 69% of ransomware attack started from an email. 31% of the victim of ransomware paid ransom to the attackers hoping that they would return their data, 34% of those has then been hit for the second time meaning that the attackers attacked them again and within those, 42% of them then got hot for the third or more times. This shows that the victims or not everyone that is using the internet or has data on their deice know what ransomware is and still be able to fall into the attacker demand. Almost all the country also is not fully prepared to deal with ransomware. These countries include, the United States, United Kingdom, France, and many more. This shows also that ransomware is a big threat to not only individual, organization but also to the government (Barracuda, 2020).

A graph of blue rectangular bars with white text

Description automatically generated

Figure 1 : Country Against Ransomware

There are 2 types of ransomwares which are encrypting ransomware and locker ransomware. Encrypting ransomware involves when the attackers encrypt the data from the user devices and request the user to pay up the ransom and then the attackers would encrypt the access to files and the users hopefully will be able to see and use the data again. Examples of this type of ransomware are “WannaCrypt” and “Locky CyptoWall”. The other type of ransomware is also known as the locker ransomware. This type of ransomware attack is like its name which is locker, the attacker would basically lock the user out from the devices and demand the ransom so that they can get back to their device. It works like a locker that locks the stuff away. Example of this type of ransomware is “Winlocker” (Shah & Farik, 2017).

Ransomware has taken over 6 trillion dollars as ransom and it will continue to grow as it is expected that every 11 seconds, a device will be attack by the attackers and asking for ransom. Ransomware has also evolved from a virus that was release from floppy disk attack to now where the attackers can just enter the victims’ devices and hold it ransom. With this, we are not certain how the future ransomware will happen. As it can evolve within time to be more threatening (Humayun et al., 2021).

As more ransomware is happening, the more security features and prevention also increase. Some of these features or prevention in helping against ransomware are, defence in depth, security awareness training, simulated attacks, software-based protection, and backups. These are some of the ways to prevent an individual from falling into the ransomware and needing to pay ransom. It is important to note that even if the individual have successfully implemented of the features and prevention, there is still a chance that the attackers are still able to attack the individual device and hold the device or files for ransom (Alessandrini, n.d.).

### 2.2.4 Blockchain

Blockchain technology is a fairly new technology founded in the year 1978, but first properly in used at the year 2008. Blockchain is nothing much other than a data structure. Blockchain allows the frictionless exchange of value. Since blockchain is a fairly new technology, there is still some limitation on implementing it. As blockchain technology highly dependent on broad adoption. As blockchain technology is going to improve within time, the access of using blockchain will also increase. As some features of blockchain has already been adapted and used, some of the other features will still be in development for many years to come before they could properly be adapted and used (Fuchs, n.d.).

The first concept of blockchain technology came from the first known cryptocurrency or well known as Bitcoin. Blockchain is a decentralized and distributed ledger technology that gives information so that it could record it and stored it and maintain the data or record and share it with the community. With this, no entity can change or modify or delete the transaction history. It is also nearly impossible to attack or hack or corrupt the entire system of the blockchain technology as this technology is distributed. Thus, making blockchain a good way to be able to store the data and information (Mgcub, n.d.).

Blockchain stores their data in blocks linked with other blocks by a chain or blockchain is also known as a chain of chronological blocks. The blocks contain data that collected and processed through a process called as mining. Each of the blocks then get their own cryptographic hash and timestamp to be able to uniquely define them. The block also contains the hash of the previous block to be able to make the block in chronological order from the first block created. The first ever block create in the blockchain is also known as the “Genesis Block” (Saha, 2019).

Blockchain can be used both in the financial industry and also non-financial industry. Since blockchain does not allow the modification, delete of the data or the block, it is good to store important data that might also be confidential. Some of the possible data that could store in blockchain is legal documents, health records, fingerprints and many more. Blockchain is now know as a new technology but a reliable technology as this technology has good security records and have been successful in keeping the data or blocks safe from attackers (Crosby et al., 2015).

Blockchain technology has potential in transforming on how we are keeping our data. Blockchain with its decentralized, persistency, anonymity and auditability allow it to be a good way to be able to store the user’s data. Blockchain technology still has ways to go but what it provides now has already start to change the way to store the data. Its architecture and key characteristics have already been implemented into various application and software. With the growing of blockchain technology, the more features will be able to be implemented (Zheng et al., 2017).

## 2.3 Similar System

### 2.3.1 IDrive

IDrive is an automated backup storage that provide the users with cloud storage and cloud backup (By Aleksander Hougen & Aleksandar Kochovski). Cloud storage and cloud backup is not the same as cloud storage is used for storing files and other data from the user on the cloud and it focuses on usability and accessibility. While on the other hand, cloud backup is a system or the way that cloud can back up the individual’s data and keep the data safe, it allows the easy restoration and business continuity (Ninja, 2023). This allow IDrive to be able to be one of the most used auto-backup system that is available for users to choose from.

Since IDrive uses both cloud backup and cloud storage, and each having their own space, they can backup many different types of data including disk image, NAS backup and server backup. With the available of many types of backups made possible, many people have opted using IDrive as their main way of backing up their data (By Aleksander Hougen & Aleksandar Kochovski). IDrive is expected to have around 4 million customers allowing them to backup multiple computers and their mobile device into the same account (*About IDrive®*, n.d.).

IDrive data centre which is in the United States have their own ways to be able to keep their data from their customer safe. There they have the data encryption, physical security, and network security. The data encryption that they are using are known to be the 256-bit AES encryption on their transfer and storage system (*FAQs on Data Security, Secure Access in Cloud - IDrive*, n.d.). This type of encryption is one of the most reliable and virtually impenetrable (Kananda, 2022). There is also physical security in placed for their data centre. This physical security includes smoke detection, fire suppression system, motion sensor, secured access, security camera and security breach alarms. The last security system that they have is their network security, they regular check their network and upgrade their network when needed and also in some time they would ask third parties’ company to review their network infrastructure (*256-bit AES Encryption for Complete Protection - IDrive®*, n.d.).

### 2.3.2 Storj

Storj is another backup system that is available to the public for use. Storj that was founded at Atlanta, Georgia, United States of America, provides their customer with backup services that sends their data to Storj distributed network with an end-to-end encryption by default. Storj is a decentralize cloud storage that also used to power of blockchain technology. With this, Storj can provide fast and secure way to be able to back up the data. Storj network is powered by many computers that are located worldwide, and each computer required the Storj digital currency for payment. Founded in 2014 and operates more than 10,000 actives computers, Storj has more than six petabytes of available storage for them to use. (Rosenberg, 2022).

Storj is among the first company to implement decentralized cloud market and cloud storage network. Storj also uses blockchain technology which allows the system to be more secure, decentralized, privacy and more transparent of them to the public and user. Storj can store the data by breaking the data into smaller data packets and share it among the networks. This allows the no specific company or organization to be able to have access to all the data that is stored as it is distributed across the network. The users are also in charge of their own data and Storj also enable the user to be able to download the data from many different computers in their network. This allows the users to be able to have access to their data and be able to download their data faster thus increase efficiency (Kriptomat, n.d.).

Storj has multiple security features that enable them to keep their client data safe and secure. Storj uses encryption specifically AES 256-bit encryption to be their default encryption. Since they are using blockchain technology and decentralized, they are also able to keep the data safe by ensuring that the smaller pieces of data are spread across their network. Even if the attackers manage to attack one of the nodes or also known as their computer in the network, they are not able to get all the data as it is split between multiple nodes (*Storage That’s so Secure, Even We Can’t See Your Data.*, n.d.).

### 2.3.3 Sia

Sia is decentralized storage system that is available for use. Users can use Sia to be able to backup their data into their storage system. Sia aim is to create a cloud storage marketplace that able to store the data and have Intermediaries, boundaries, vendor lock-in, spying, throttling, and walled gardens are not allowed. Sia team contain 13 different types of members including the president known as Luke and the chairman known as Eddie (*Sia - the Sia Foundation*, n.d.). Sia is now operated by a non-profit organization and still runs with the same team and foundation.

Sia store their data by encrypting and distributing the files across their decentralized network. This allows no other party or users to be able to denies access to the user data. The encryption used by Sia is Threefish encryption algorithm. This type of encryption is resistant to side-channel assaults and related-key attacks (*Should I Use IPFS or SIA? - Filebase*, n.d.). Block cypher Threefish is a big, flexible one. It is defined for 256 bits, 512 bits, and 1024 bits, respectively. The tweak value is 128 bits for all block sizes, and the key is the same size as the block (*Schneier on Security: Threefish*, n.d.).

Sia average storage in the cloud cost 90% cheaper than other cloud storage provider. Sia storage has been around since the year 2014. Sia working with many other programmers to be able to create an open-source programme that is now providing and building innovative creation. All Sia programmer also produce and create other projects such as Filebase, VUP, Pixeldrain, Skynet, Arzen. With this, Sia programmers have sown that they have the capabilities to be able to create this application (*Sia - Decentralized Data Storage*, n.d.).

### 2.3.3 Comparison

Table 2: Similar System

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Function/Features** | **IDrive** | **Storj** | **Sia** | **Proposed System** |
| **Type of Backup** | Auto-backup once in 24 hours | Can be auto backup or manual backup | Can be auto backup or manual backup | Automated backup with the choice how often the individual wants to have their data being backup. |
| **Storage used** | Cloud Backup and Cloud Storage | Data is encrypted, divided into smaller bits, and then distributed globally through a distributed network of independent computers. | Data encrypted and distributing the files across their decentralized network | Ethereum based blockchain network. |
| **Algorithm** | Encryption AES 256-bit | Encryption AES 256-bit | Threefish encryption algorithm | Encryption SHA 256-bit |
| **Security Measure** | physical security for data centre and network security by third parties. | decentralized network of independent computer storing the data | distributing the files across their decentralized network | private and public key infrastructure, multi-signature authentication. |

## 2.4 Summary

There are many different types of auto-backup system out there from different types of backup system whether it is manual, automated or both combine, or whether different type of storage used, or the security features, all of the backup system provides the users with the same function as they are helping the user to be able to be able to copy the data and keep it safe and allowing the user to be able to download the data when needed. As the three types of storage system that is shown above, they have their similarities and differences such as one uses the blockchain technology and one uses the cloud technology. The blockchain technology is a new technology and not everyone is using it yet. Since the type of technology used is also different, it is also to expect that the security features that is in place will also be different as the security feature that should be implemented has to work and do the proper thing for the system and allowing the system to be more secure. Thus, making each backup system the same yet it is not the same as users might have different used for their system. Some users need to back up the network or database, but some backup system does not allow it. It is important that the users do their research before choosing a backup system (Young, 2023).

# Chapter 3: Technical Research

## 3.1 Programming Language

### 3.1.1 Python



Figure 2: Python Logo

One of the possible programming languages to be used in this project. Python is a high-level programming language that is develop by a Dutch programmer named Van Rossum, in the year 1989. He first started making the python programming language because he wants to create a language that is to read, write and maintain by other programmers. Python first public released in the year 1991 on the February 20. The first release of python is known as the version 0.9.0. Python 2 is released in the year 2000. Some Python 2 design decisions made the language more challenging to maintain and progress as the language matured. Python 3 was developed in response and released on December 3, 2008. Python 3 was a significant improvement since it introduced backward-incompatible changes to increase consistency and remove extraneous functionality. Python 2 has reached its end of life and is no longer actively supported; nonetheless, Python 3 is not fully backward compatible with Python 2, necessitating certain code tweaks when migrating between versions. By the 2010s, Python had risen to become one of the most widely used programming languages, and its popularity has only increased since then. The fact that numerous IT companies, startups, and organisations chose to use Python for their projects has helped to foster the language's vibrant community and environment (Ivey, 2023b).

Python is a good code to be used to be able to code or implement blockchain with due to multiple reasons. One of the reasons is that programmers that decides to Its simplicity comes from a variety of characteristics. For instance, in Python, white spaces signify code blocks. Thus, adding keywords or curly brackets is not a concern for developers. Python allows developers to create blockchains quickly and with little code. Python makes life easier for developers because it is a scripting language that doesn't require compilation. Python additionally provides the pre-compiling option, which is advantageous for developers working with blockchain. Python also offers free coding tools designed specifically for creating blockchains. Python also have premade libraries and framework to be able to assist the programmer to make the programmer to be able to code more effectively (T. K. Sharma, 2022).

### 3.1.2 Java



Figure 3: Java Logo

Java is a high-level programming language that is well known across many different programmers. In the year of 1991, James Gosling started the development of java programming language as an object-oriented programming language. Gosling wanted to create a virtual machine and language that used a recognised notation like that of C but was more precise and straightforward than C/C++. Java 1.0 saw its first public execution in 1995. 'Write Once, Run Anywhere' was promised for well-known platforms with no runtime fees. With security limiting network and file access, it was incredibly secure and adaptable. The significant web browsers quickly integrated it into their default settings in a stable "applet" setup. Sun released a significant amount of Java under the GNU General Public Licence as free and open-source software (GPL) on November 13, 2006. Sun finished the process on May 8 by making the Java core code entirely free and open source, with the exception of a minor section of the code for which Sun lacked copyright (UNext Editorial Team, 2023).

Java is a language that is also capable of coding and implementing blockchain technology. Features of Java includes object-oriented programming language, simple, architecture-neutral, secure, independent platform. There are many features that makes Java a unique programming language but reasons that Java is one of the languages chosen to be able to use to do this project is due to multiple reasons such as difficulty of java is less then C or C++, it is independent from the operational system, easy memory allocation process, large library’s collection, effortless memory cleaning. Some of the blockchain platforms that uses Java as their programming language are IBM Blockchain, Ethereum, Neo’s contracts, BitcoinJ, Hyperedger’s contracts (T. K. Sharma, 2022).

### 3.1.3 Programming Language Chosen

Python and Java like stated above, both of this programming language are capable of be chosen to be used to be able to code this project. Each of the programming language has their own advantages and disadvantages when coding this project as discussed above. Python and Java is both famous programming language that are widely used by may programmers to be able to develop a software. These coding languages are because they are easy to learn and understand and have big capability on how and what they can provide to the programmer in just one programming language (Coding Ninjas, n.d.).

In this project the programming language that will be used is Python. Python is chosen over Java due to the ease of use and readability, rapid prototyping, rich ecosystem for blockchain, extensive data processing and analysis, strong community support, flexibility, and versatility. Python is also chosen due to ability when it comes to Rapid Application Development. Java is still a good programming language that is able to be used to code this project, but due to all of the reasons stated above, Python will be more suitable to be used to be able to code this project and system (T. K. Sharma, 2022).

## 3.2 Interactive Development Environment (IDE) Chosen

An integrated development environment (IDE) is a piece of software that aids in the productive creation of software code by programmers. To write code, the programmer can use any text editor. Nevertheless, most integrated development environments (IDEs) offer features beyond text editing. They offer a centralised interface for popular developer tools, greatly enhancing the effectiveness of the software development process. As opposed to painstakingly integrating and configuring various tools, developers can start programming new apps right away. Few of the reasons that IDE was preferred over text editor code is because code editing automation, syntax highlighting, intelligent code completion, refactoring support, local build automation, compilation, testing, debugging.

A logo with colorful squares

Description automatically generated

Figure 4: PyCharm Logo

PyCharm will be chosen as the IDE that will be used in this project. PyCharm is one of the most famous Python IDE that is available for programmers or users to use. Features that are available in PyCharm to enable the programmers to be able to use PyCharm or prefer it because it has intelligence code editor, error detection, code navigation, refactoring, live changes can be viewed directly in a web browser, autocompletion features and debugging tools. Finally supports many libraries that help them to be able to be versatile. The advantages of using PyCharm are simplicity of installation and use, the abundance of plugins, and the productivity shortcut. Another useful feature is the colorization and auto-completion options. With PyCharm, software development is generally completed significantly more quickly. The procedure is improved by the error highlighting. The Python developer community is relatively large, which enables speedy problem-solving.

## 3.3 Operating System Chosen

The operating system is important as it is where all the task is going to happen. There are a few operating system that can be used such as Windows and Linux. As for this project, the operating system that will be chosen is Windows, specifically Windows 11. Windows currently have 1.6 billion active devices making it the most used operating system ahead of another operating system. Windows operating system has changed over the years, by each Windows update, the operating system would most likely be better than the previous generation of Windows. As of June 2022, Windows operating system holds 76.33% of all the operating system market followed by macOS with 14.64% then Linux at 2.42%. This shows that Windows has dominated the operating system market (Wise, 2023).

The reason of using Windows to do this project is because of few reasons. One of the main reasons of using Windows is due to the accessibility as the computer that will be coding the project has Windows as the main operating system. Windows also supports numerous software development and tools to enable the programmer to be able to develop the software better and freely. As a result of its user-friendliness, software integration of tools, community support, and familiarity with Windows OS, Windows 10 is the best alternative for developing this project.

## 3.4 Summary

The summary for this chapter is that the programming language that the programmer will used to be able to develop this software application is by using the Python programming language. The reason of using the Python programming language is due to its simplicity and it can integrate blockchain technology. The programmer will also use PyCharm as their IDE to be able to develop this software application. Lastly the operating system that will be chosen is the Windows operating system which will give the programmers access to the tools freely.

# Chapter 4: System Development Methodology

## 4.1 Overview

A conceptual model for project management known as the systems development life cycle (SDLC) details the phases of an information system development project, from the early phase of a feasibility study to the ongoing maintenance of the finished application. Both technical and non-technical systems can use SDLC. A system is typically an IT technology, including hardware and software. Standard SDLC participants include system and software engineers, development teams, and end users. Project and programme managers are also frequently present (Gillis, 2019).

There are multiple different types of system development life cycle models. The models include the waterfall method, rapid application development, joint application development, fountain model, spiral model, build and fix, synchronize-and-stabilize and the most common method is agile software development (Gillis, 2019). For this Chapter we will go through two types of models and compare them and selecting one of the two models. The two models that will be selected is the Waterfall and Agile methodology.

## 4.2 Waterfall Methodology

### 4.2.1 Background of Waterfall Methodology

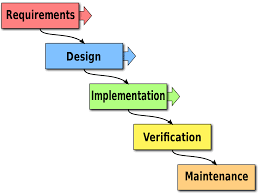


Figure 5: Waterfall Methodology Diagram

Established by Winston w. Royce in the year 1970, the waterfall methodology contains 5 phases or stages within it. These phases are Requirements, Design, Implementation, Verification and Maintenance. Using a linear approach, the waterfall methodology is a popular project management technique. Before proceeding to the next phase in a Waterfall workflow, each stage must be finished. Even though there are many different project management approaches, Waterfall is best suited for projects when the goals are known from the start (Hoory, 2022).

### 4.2.2 Overview of Waterfall Methodology

As stated above, waterfall methodology consists of five different stages which are requirements, design, implementation, verification, and maintenance. In this part, each area will be discussed further and be explain what does each of the phase’s role in the waterfall methodology.

First stage of waterfall methodology is the requirements or also known as analysis phase. In this stage, the developer would need to outline the big picture of the project. It is important that the developers outline the outcome as in the waterfall methodology, it is important that the end results are clearly establish before the project has even started. Not only the results, but the developers should also outline and document the project requirements. This phase is important as this is the beginning phase and without proper outline, the project does not have a clear goal and might be an issue in the later phases. It is required for the developers to write with the best details possible as it will be track again in the later stages of the project (*Agile Vs. Waterfall Methodology Explained | PM Tips for Non-PMs*, n.d.).

The second stage of the project is also known as the design phase. In this phase, the main idea to find solution to be able to achieve the final product. It is important that in this phase is to create an overall plan for the project and how are the developer going to be achieving the goals that has been documented in the first stage. The one way to be able to achieve this phase properly is by creating a plan and this plan can include sitemap, wireframes, architecture plans, user flows, detailed project brief. Those plan that has been stated is from multiple types of industry as waterfall methodology can be used in many different types of industry. It is important that this part is done properly as this part will be then use by the programmer to be able to bring the project into life and make the project successful (*Agile Vs. Waterfall Methodology Explained | PM Tips for Non-PMs*, n.d.).

The third stage of waterfall methodology is the implementation phase. This stage is also crucial as this stage is responsible to be able to create the project itself with all the requirements and plans that has been given to the programmer. This stage is when the programmer will use the designs and plans that has been given to them by the developers in stage 2 then by using the technology to be able to implement them. This stage is also the stage where data may be needed to be collected and the design will be gone through to see if it is possible to be implemented (Hoory, 2022).

The next stage is the verification stage. This stage is the stage where the project is given out and be tested. In this stage, the project will also see if they are able to fully reach the requirements set in the previous stages. If the requirements are not met, they are able to go back and check why it does not match and make changes if necessary. In this stage the project will also go through various testing to allow the programmer to know the capability of the project itself. It is important that in this stage it should not be neglected as this is the last stage before the project will be set to the public (Hoory, 2022).

The last phase is the maintenance phase. Once the project has been validated and verify, the work of the programmers is not done as they still have to maintain the project. The maintaining of the project can include upgrades or updates that will be needed in the future. In this phase it is the only phase that will not have the need to look at the task, milestones and hand-off that happen at the previous phase. But this phase still remains as one of the important phases in the methodology as this ensures the smooth running of the system that has been released (Hoory, 2022).

## 4.3 Agile Methodology

### 4.3.1 Background of Agile Methodology

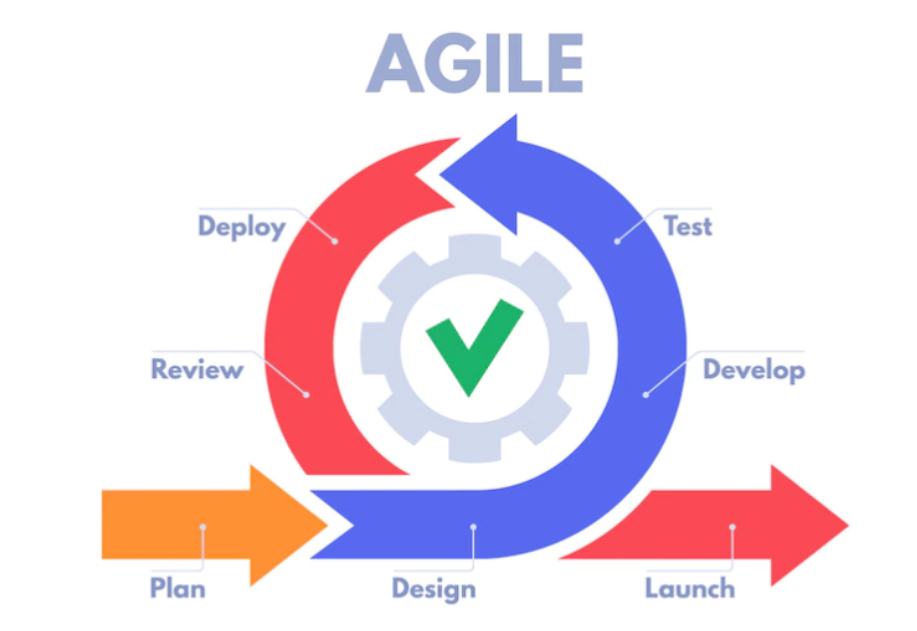


Figure 6: Agile Methodology Diagram

Agile Methodology came after the waterfall methodology. This happens as when programmers started creating web applications. The early work was largely carried out at startups, where teams were smaller, collocated, and frequently did not have typical computer science degrees. Financial and competitive incentives existed to hasten the release of new websites, applications, and capabilities. Platforms and development tools quickly evolved in response. So, in the year 2001, Kent Beck, Martin Fowler, Ron Jeffries, Ken Schwaber and Jeff Sutherland form a team and came up with the Agile Manifesto. Agile Manifesto is a document that contain the groups beliefs and thought on how software development should be. They emphasised working together rather than documenting everything, self-organization over strict management procedures, and the capacity to deal with continual change as opposed to being forced into a rigorous waterfall development approach. With this, the Agile methodology was created (Sacolick, 2022).

### 4.3.2 Overview of Agile Methodology

Agile methodology contains different stages within it. There are 6 stages in the agile methodology, these includes concept, inception or requirements identification, iteration or development, release, maintenance, and retirement. In this part of the report, we will go through each stage in details on how the stages function or used in helping to the develop the project.

The first stage of Agile methodology is the concept stage. This is the stage where the stakeholder of the project would the objective and scope of the project. It is important to take their views as they are the person who is sponsoring the project and they would want the project to go how the visualize it. This is the stage where it is necessary to outline and document the key requirements to the project. It is also important to determine the time needed to be able to complete the project. In this stage, it is only required that the minimum requirements are documented as they would be added in the later stage. This will assist in determining whether or not the product is practicable (Nehra, 2022).

The next stage is the inception or requirements identification. In this stage it is necessary to create the software development team. As the product owner will be responsible to choose the programmers that will help in making the project come alive. This will include both front-end and back-end developers. It is important to choose the programmers that will best suit the project and teams so they would be able to work with the provided resources and tools. After the team is chosen, the team can then start designing a prototype and create an architecture for the project so that they can present it. With this they can see if the prototype and architecture will fulfil the requirements needed from the stakeholder (Nehra, 2022).

The third stage of this methodology is iteration or development stage. This stage is going to happen when the stakeholder has approved the design. The team then will continue the project by developing the project into a system that can be fully used with all the requirements needed. The system will also go through various test and reviews until the final product is agreed on. It is important that the team always must go through the requirements that has been documented as the stakeholder will expect that the requirements has been reach (Nehra, 2022).

The next stage in this release. In this stage the product will be given to the QA to be able to conduct the necessary test to ensure that the product is clean and has no bugs that will affect the system. The functionality and the quality of the system will be tested until the QA team members are satisfy. If there are any bugs found within the system, the development team will be address immediately as they would address the problem rapidly. This stage also may get the user feedback. This will further improve the product before releasing it into the public for use (Nehra, 2022).

The fifth stage is the maintenance stage. This stage is like the maintenance stage in waterfall methodology. The developers will and needs to the system working without problems. This also means that they would have to provide support towards the users that is using the system. They would also have to provide update to the system when it is needed to allow the system to be more convenient to the users (Nehra, 2022).

The last stage of the agile methodology is the retirement stage. There will always be a time where there is either a better system available or the current system is outdated and provides inconsistent for the organization. When this time comes, it is normal to turn off the system and retire it. If there is another system in placed to replace the old system, the developers will be needed to transfer the data to the new system. The old system will then be discontinued, and the remaining activities will be closed and finally the system would not be functioning (Nehra, 2022).

Above describe the 6 stages within the agile methodology. It is different from the waterfall methodology. According to what the developers need, they would adopt different types of methodology. No matter what the methodology is, it is important to have the system being able to fully work reaching the requirements in the end (Nehra, 2022).

## 4.4 Comparison of System Development Methodology

### 4.4.1 Comparison Table

Table 3: Methodology Comparison Table

|  |  |  |
| --- | --- | --- |
|  | **Waterfall Methodology** | **Agile Methodology** |
| Life Cycle | Linear | Iterative |
| Approach | Plan-Driven | Value-Driven |
| Team size | Large team | According to what is needed |
| Change Management | Inflexible | Flexible |
| Delivery of Product | End of the SDLC | Early Stages by providing prototypes. |
| Risk | High | Low |
| Size of Project | Large | Small |
| Development Pace | Slow | Fast |

### 4.4.2 Selected System Development Methodology and Justification

While waterfall methodology is known to be the famous methodology and mostly used, in this project the agile methodology would be used in placed of the waterfall methodology. This is due to multiple reason. As the table shown above the difference between the two methodologies, agile is more suitable in this project due to the life cycle, approach, team size, change management, delivery of product, risk, size of product, development pace.

The life cycle of agile is iterative making it possible to keep improving the project while the programmer starts to make the system unlike the waterfall methodology where it is linear, and it is hard to make changes when necessary. The approach of agile is value-driven where waterfall in the other hand is plan-driven making it more flexible. Team size also matters in this project as the programmer that is creating this project is small. Risk is another important due to it being a small and fast project that needs to be done, it is better to have low risk. The size of this project and the pace also suites the agile methodology as it is small and fast. Overall, the agile methodology is more suitable to be able to carry out this project.

## 4.5 Application of Agile Methodology on this Project

To be able to properly apply agile into this project, this section will be discussing on how agile will help to develop the project. It will cover each stage used in completing the project.

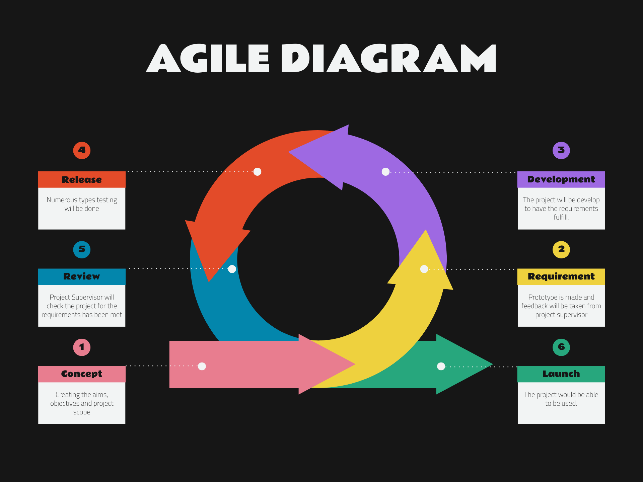


Figure 7: Agile Diagram for the Project

**Concept**

After selecting and deciding on the title, the research to further understand the topic will be done. This has been shown in this report in chapter 2 where the researcher has conducted the literature review to understand on how and what this system needs and provide. After the research, the aim, objective, and project scope has also been documented in chapter 1.

**Inception or Requirement Identification**

After the research has been done and the aims, objective and project scope has been done and the programmer now knows what he is task with, the programmer now starts to think of ideas that can be implemented, and a prototype is shown this to the project supervisor and get the feedback. With this, the programmer can then make adjustment to be able to meet the requirements and achieve the changes that has been suggested by the project supervisor. A questionnaire will also be sent out to further understand what the users needs to make the proper adjustments.

**Iteration and Development**

Once finish with the second stage, the programmer will then proceed with developing the software. This will then require the programmer to look at the deliverables and complete the deliverables that has been documented. With the requirements set in place, and the programmer understands better on the needs and the prototype feedback, the programmer can then develop the project. This also enable the programmer to complete the development at a faster pace.

**Release**

Once the development is finish, the project will then have to go through testing to make sure that it is according to the needs and requirements that has been documented. The types of testing that may be done to the products are system testing, unit testing, integration testing (Kinsbruner, n.d.). The product will also be tested to see the functionality and I there is any bugs found. If there is any problem, the programmer would be notified, and the problem will require fixing. If not, the project will be able to be finalize.

# Chapter 5: Research Methods

## 5.1 Introduction

### 5.1.1Research

Research is an important part of this project to get feedbacks and to understand the market better. Research can be done in many ways, but there are 2 main types of research known as qualitative and quantitative research. Each of this research has its own advantages and disadvantages (Kinsbruner, n.d.).

Quantitative research is not about how many data that can be gathered, but the ability to understand the feeling of the people that participate. This allows the researcher to be able to gather more insight into the project and understands the needs of the people. Interviews is a good example of qualitative type of research. The advantages of using this type of research are it allows the researcher to understand and capture the feeling of the person that is helping with the research, enables the further explain in detail as not everything ca be understood if it is written down, a more flexible approach, enables the generation of ideas that might be able to help. The disadvantages of using this research are we are not sure if the results will be a true reflection that is needed for the topic, the results might favour an outcome thus making it not consistent, bias might play a role thus also making the results unreliable and the questions that is asked is only the question that the researcher can think of (Vaughan, 2021).

In the other hand the other type of research is also known as quantitative research. Quantitative research gather data that are observable by using  [statistical](https://www.liveabout.com/market-research-101-data-analysis-2296676), computational, or mathematical techniques. The example of quantitative research is survey and questionnaires. Like qualitative, it also has both advantages and disadvantages. The advantages of using this method are the data is more reliable and less open to argument, straightforward analysis and prestige as the questions is not very technical and easy to understand to most people. The disadvantages of using this research method are false focus on numbers, difficulty setting up research model and it can mislead (DeVault, 2020).

### 5.1.2 Sampling Methods

Sampling methods is to draw valid conclusion from a group of people while on the other hand, sample is the chosen group of people to participate in the research to be able to collect their data. There are two different types of sampling method which are probability and non-probability (McCombes, 2023).

Probability sampling method is when the researcher used a random sampling technique to be able to create the sample. This allows everyone that are taking part in the sampling method to have an equal chance of getting chosen to have them participate in the sample. There are five different types of probability sample which are simple random sampling, stratified sampling, systematic sampling, cluster random sampling, muti-stage random sampling. Each of these types has their own advantages and disadvantages. Simple random sampling allows the creation of sample which is represents the population the most while when doing it on the large scale, it might be tedious, and it will consume a lot of time. Cluster sampling is easy and convenient to be used, while if the unit is homogenous, it might not work well. Stratified sampling creates layers to able to highly represent the population in layers, same as simple sampling, the disadvantage is when doing it in large scale it is tedious and time consuming. Systemic sampling able to create a sample that I representative of the population without the need of random number generator, while it is not as random as simple sampling method (*Probability Sampling: Definition, Types, Advantages and Disadvantages - Statistics How To*, 2021).

Non-probability sampling method is the opposite of probability sampling as the chances of getting selected cannot be calculated. Advantages of using this sampling method is that it is easy to be used and when the population is small, it is possible to use this sampling method. While using this, it doesn’t allow the sample to fully represent the population. There are eight types of non-probability samples which are convenience sampling, haphazard sampling, purposive sampling, expert sampling, heterogeneity sampling, modal instance sampling, quota sampling, snowball sampling (*Non-Probability Sampling: Definition, Types - Statistics How To*, 2023).

The researcher should find out and understand the difference of each sampling method to allow the researcher to conduct better and fair samples. This require the researcher to be able to know which sampling method and which type of sampling the researcher is going to use. This is important as each of the sampling methods and the types has its own advantages and disadvantages and not all types of the methods can be used in the project.

### 5.1.3 Data Collection Method

After understanding both the research types and sampling method, the researcher has chosen the research types and sampling method to be able to use in this project. As for the research method, the researcher has chosen to use the quantitative type. This is due to the researcher wants to understand the population more and what they need and understand about the project. As the sampling method, the researcher has gone for probability sampling. In more details, the researcher is going for survey to be able to get the sample. This means that probability sampling is chosen as the survey taken would reflect on most of the population.

## 5.2 Design

The select way of getting the information as said above is by survey. The survey will be design in google forms and will be send out to the public. The survey will also be strictly confidential as their data will only be used for this report and project. Survey that is provided is also not required for the individual to answer and it is voluntary. Survey is chosen over other methods like interview as the researcher wants to understand the population better as in general.

The main aim for the survey that was given out is to get feedback from the population about backup system and their idea of blockchain technology. These are the chosen as their main topic as backup system is the main aim of this project. As it is the main idea of this protect is require that the researcher enables and takes in all the information that he can to understand the population better. This would include also due to one of the problems is not everyone understand the use of backup system and not everyone is using the backup system. This is why, the survey about backup system and their understanding and how much the population knows is important. As for the blockchain technology, as one of the deliverables of this topic is about blockchain technology, it is important that the researcher also have information about blockchain technology and how much the population understand about this new emerging technology and their views on this technology. Not only that, but the survey also includes questions that would require the population to be able to give their suggestion on how to make the application better.

A screenshot of a survey

Description automatically generated

Figure 8: Background about the Research

Above shows the declared information about the survey conducted. As it is required for the researchers to declared and explain why they are conducting the survey and explain that the data collected is confidential and it is not mandatory to answer the survey. It will also describe who is conducting the survey and the title of the project to allow the individual who answer the survey to know what the survey is about.

A screenshot of a survey

Description automatically generated

Figure 9: Demographic Question

Above shows Section A of the survey which will understand the demographic of the individuals who answer. There will be two question is this are one being what the induvial age group is and the other one would be what the current employment status of the individual is. This is important as different demographic profile may result in the difference in the survey and cause the data to be inconsistent.

A screenshot of a computer

Description automatically generated

Figure 10: Section B Background

Section B is followed the previous demographic profile question. Section B will start with a simple explanation of backup system and blockchain technology. This allows the individual who participate in the survey to understand what backup system is and blockchain technology.

**Question 1:**

A white background with black text

Description automatically generated

Figure 11: Question 1

**Objective:** This question main objective is just to be able to know if the individual has already a backup system available. As from the knowledge above, not everyone has a backup system working. The participants will have 3 option to answer which are yes, no, and not sure. The participants are required to answer this question.

**Question 2:**

**A screenshot of a computer

Description automatically generated**

Figure 12: Question 2

**Objective:** This question is asked to understand what the participants is view on backup system and do they think that it is necessary to have a backup system in this era. This question is based on linear scale from 1 to 5. 1 being unlikely or not important to 5 being likely or important.

**Question 3**

A screenshot of a computer

Description automatically generated

Figure 13: Question 3

**Objective:** This question is related to question 1. In this question, if those who answer yes, this question is there to know what backup system the participants have. There are a few backups system that has been written down, but those with other backup system can juts choose the other option and write down their name of their backup system.

**Question 4:**

**A screenshot of a computer error

Description automatically generated**

Figure 14: Question 4

**Objective:** This question is to understand if the participant understands the use and how does the backup system work. As from the review above, not everyone understands how a backup system works. This question is required to answer by all participants as it is important to understand how many people understand what a backup system is.

**Question 5:**

**A white background with black text

Description automatically generated**

Figure 15: Question 5

**Objective:** This question is to understand how many participants understand their backup system security wise. As there are many different type of security features, some are better than others, so it is a survey to know if the participants just simply get a backup system or understand what their backup system security features before getting it.

**Question 6:**

A screenshot of a computer

Description automatically generated

Figure 16: Question 6

**Objective:** This is a follow up for the previous question. This is for those who knows the security features and this question is asking those participants on what security features does their backup system have.

**Question 7:**

**A screenshot of a computer

Description automatically generated**

Figure 17: Question 7

**Objective:** This question is to follow up with the previous question on asking about their backup system. This question is to understand where the backup system makes a copy of the user data. This is important to understand how many people prefer which storage system.

**Question 8:**

**A screenshot of a computer

Description automatically generated**

Figure 18: Question 8

**Objective:**  The question above aim is to understand what the public view of the population about backup system. This question as shown will be using the linear scale to understand with 1 being not important and 5 being very important.

**Question 9:**

**A screenshot of a computer error

Description automatically generated**

Figure 19: Question 9

**Objective:** The question above is shown to understand the population on the favourable amount of time before the backup system automatically backup the data.

**Question 10:**

**A close-up of a computer screen

Description automatically generated**

Figure 20: Question 10

**Objective:** This question is to get the opinion of the population if most of them understand what is blockchain technology or not. This is crucial as the rest of the question are mostly related to blockchain technology.

**Question 11:**

**A close-up of a computer screen

Description automatically generated**

Figure 21: Question 11

**Objective:** The question is specific question to understand the population view on storing data into blockchain technology. As stated above, blockchain provides many security features but it is still important to understand the population view on this matter.

**Question 12:**

**A screenshot of a computer

Description automatically generated**

Figure 22: Question 12

**Objective:** This question is to understand what the population view of the security features that can benefit a backup system. It is important to understand as there are a lot of security features that can be implemented thanks to blockchain technology.

**Question 13:**

**A screenshot of a computer message

Description automatically generated**

Figure 23: Question 13

**Objective:** This question is to understand the population on which platform is the most wanted to have a backup system created. This is important as is it needed to understand the population before starting the project to be able to attract the interest in the population.

**Question 14:**

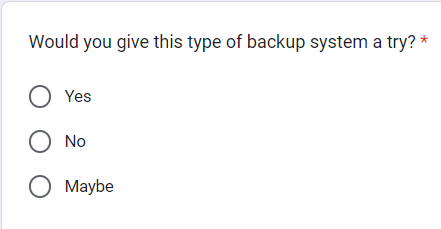
****

Figure 24: Question 14

**Objective:** The question above is to understand the population if they are willing to give the backup system a try after completing the development process. It come with three choices which is required for the individual that participated in the survey.

## 5.3 Summary

In conclusion, doing the survey which a quantitative method of research allowed the researcher to be able to understand more about the needs and understanding of the population. The question that is available in the survey also helped the researcher to be able to understand and made modification accordingly to the answers provided. This is allowed since quantitative method allows the data collect to be able to represent most of the population. With this, the survey is important and will be conducted and collecting 80 data from individual to be able to see the results clearly.

# Chapter 6: Requirements Validation

## 6.1 Analysis of data

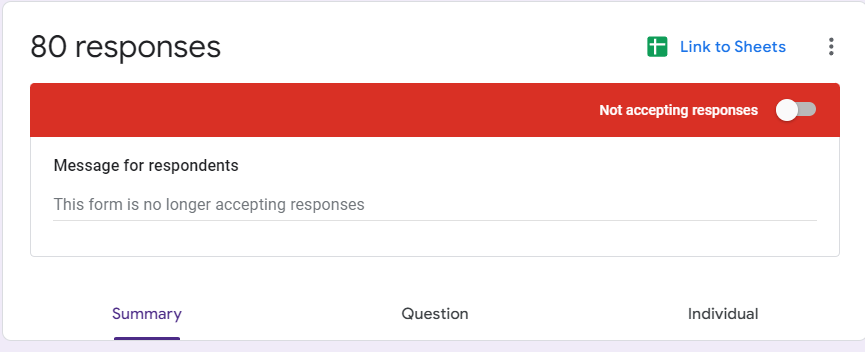


Figure 25: Responses

As stated before, the aim is to be able to get 80 respondents to be able to participate in this survey. The researcher manages to be able to get the amount of individual and this allows the data collected to be more precise and accurate.

### 6.1.1 Demographic Data

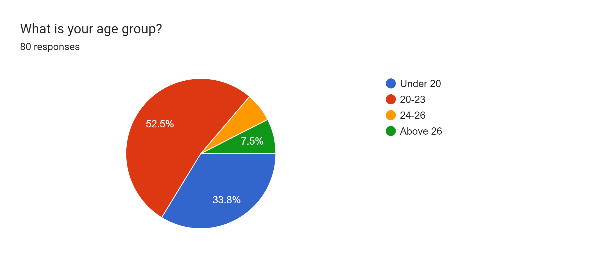


Figure 26: Demographic Question 1

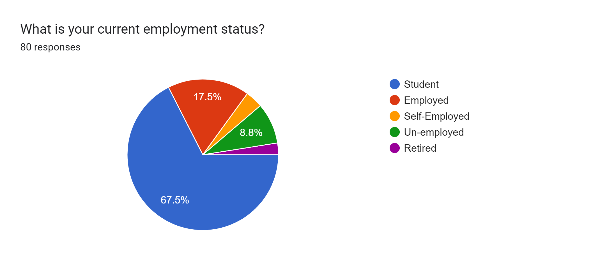


Figure 27: Demographic Question 2

These are the two demographics question in the survey. One being what is the age group and the other one being what is the individual employment status. With the first question, most of the population is age between the age group of 20-23 followed by the group age of below 20 while the rest of the group age are making little percentage of the population. This can show that the researcher has reached out to younger and active individual. The reason that this also happen is that the research is conducted by mostly university student that is currently studying.

As for the second question, the employment status of most of the population is student followed by employed. As stated before, the research is done within a university and most of the individual that answered is studying, the results shown is reflected to the population that answered the survey. There are also some other individuals that answered the survey as it is not strictly to students and other individual such as employed or working individuals also can answer the question.

### 6.1.3 Analysis of Data

**Question 1**

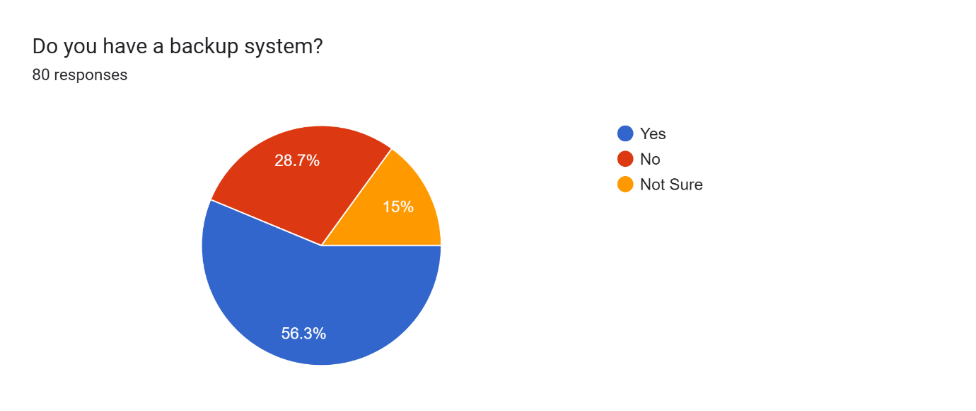


Figure 28: Data of Question 1

From the pe chart above, most of the population has a backup system working and running. This can be clearly seen by the percentage of yes which is at 56.3%. With this, the researcher can conclude that individuals in the current era adopt and uses backup system.

**Question 2**

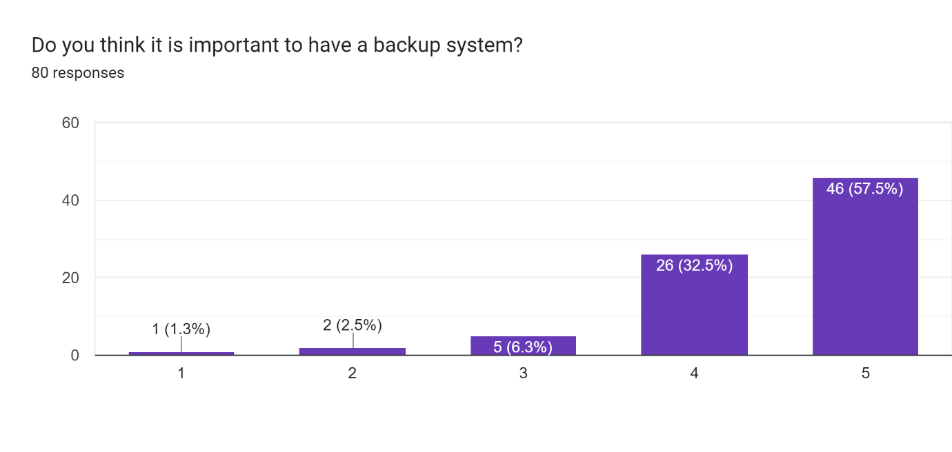


Figure 29: Data of Question 2

The first question in the survey is regarding about the importance of a backup system and the view of the individual if it is important or not. The bar graph shown above can be clearly seen that number 4 and 5 is most chosen by the population. From the answers, the population clearly stated that on average most people feel that a backup system is important. Which allows the researcher to be able to conclude that a backup system in the eye of most population is an important system to have for an individual.

**Question 3**

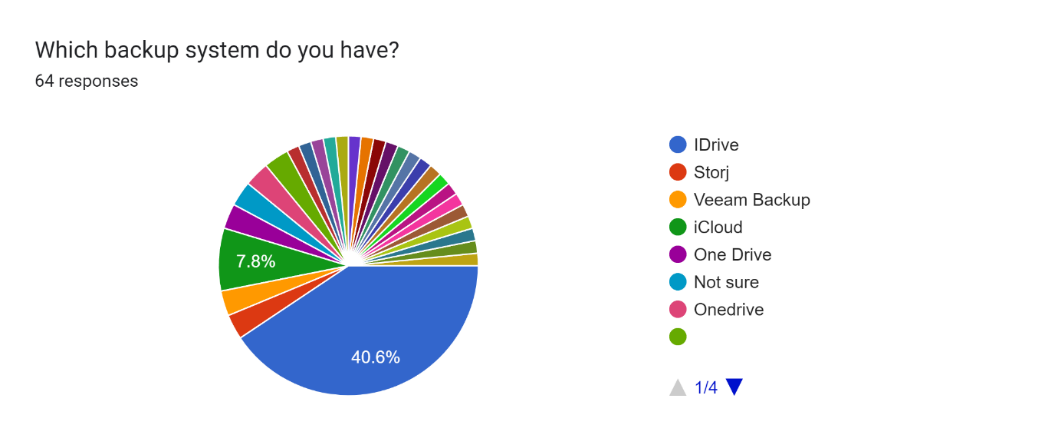


Figure 30: Data of Question 3

From the pie chart above, it is clearly visible that there are many different types of backup system that the current individual has and using. But from the chart above, there is 2 types of backup system that are more prominent which are IDrive and iCloud. While IDrive still has the most people using at 40.6% of the population using while iCloud has 7.8%. Ultimately, the results shown that there are many different types of backup system that the population uses to be able to back up their data.

**Question 4**

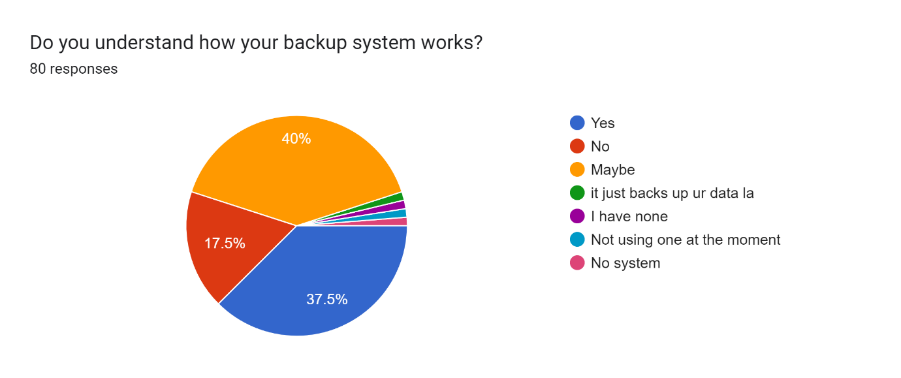
****

Figure 31: Data of Question 4

From the pir chart above, it is shown that most of the population does not have the full understanding on how the backup system works. This can be shown by the 40% of the population choosing “maybe”. While there are 37.5% of people understanding the backup system, it is also shown that 17.5% of the population do not know how their backup system works.

**Question 5**

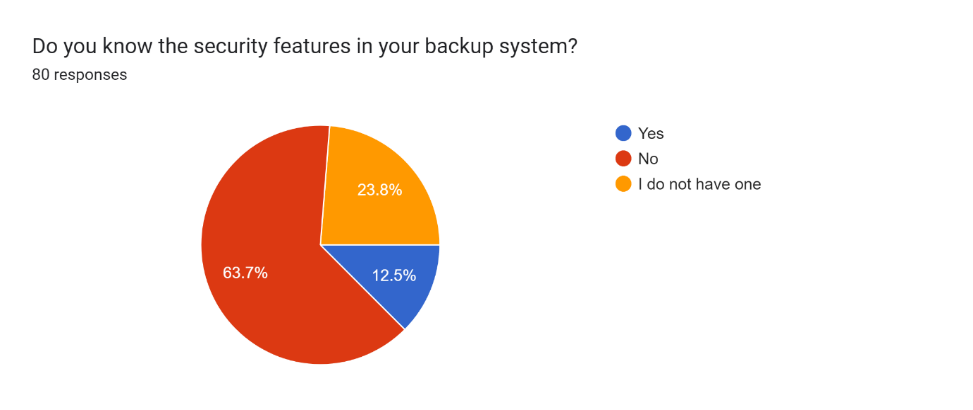
****

Figure 32: Data of Question 5

This question has a definite answer from the pie chart shown above. 63.7% of the population do not know or understood the security features that is available in their backup system. Only 12.5% of the people understood their security that their backup system offers to protect their data.

**Question 6**

**A screenshot of a computer

Description automatically generated**

Figure 33: Data of Question 6

The question above requires the individual who answered yes in the previous question to state out their features that is available. From this we can see that those who answer yes not everyone properly knows and understand what security features that their backup system provided.

**Question 7**

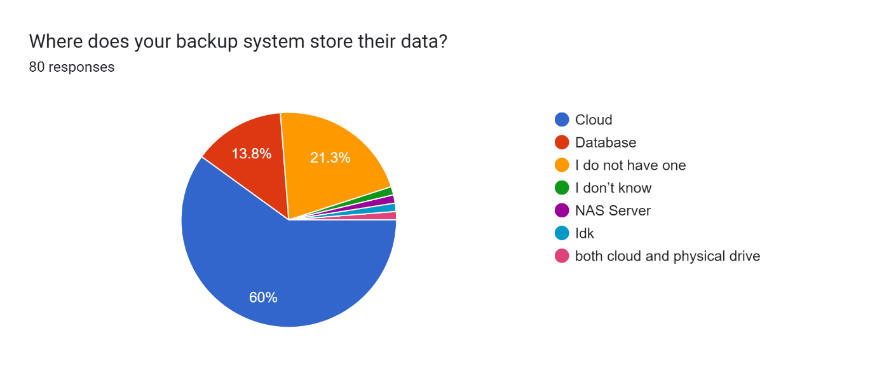


Figure 34: Data of Question 7

From the pie chart above, we can see that most of the population that uses the backup system, have their data stored at a cloud storage. There is 60% of the population uses vloud and only 13.8% of people backup their data using database. This shows that cloud is the preferred storage method.

**Question 8**

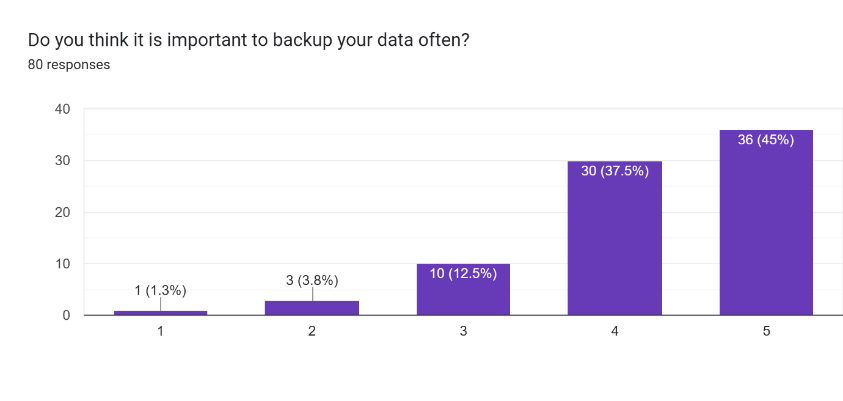


Figure 35: Data of Question 8

From the bar graph above, it is clear to see that the average people voted that it is important to backup the data. This can be seen by most of the amount individual selecting number 4 or 5 as their selection choice. Which number 5 being that backing up the data often is important.

**Question 9**

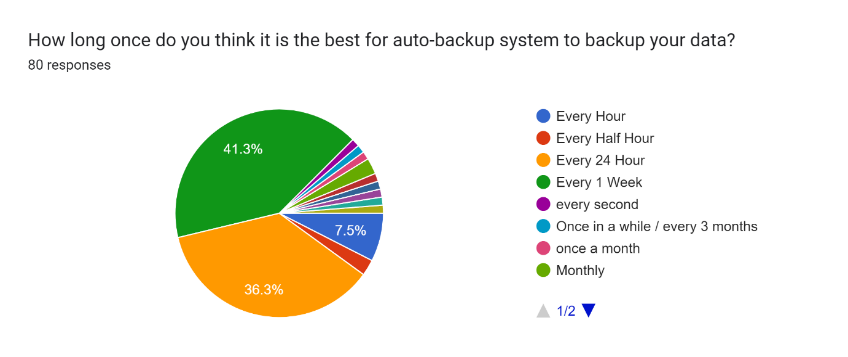


Figure 36: Data of Question 9

From the pie chart above, it is a close result between every 1 week and every 24 hours. In the end with 41.3% of the population every week is more favourable. This then can be concluded that from the population, every 1 week of backup is more preferred. This can be for multiple reason as since most of the individual that answered are students, their data does not have much information that needs to be backup instantly.

**Question 10**

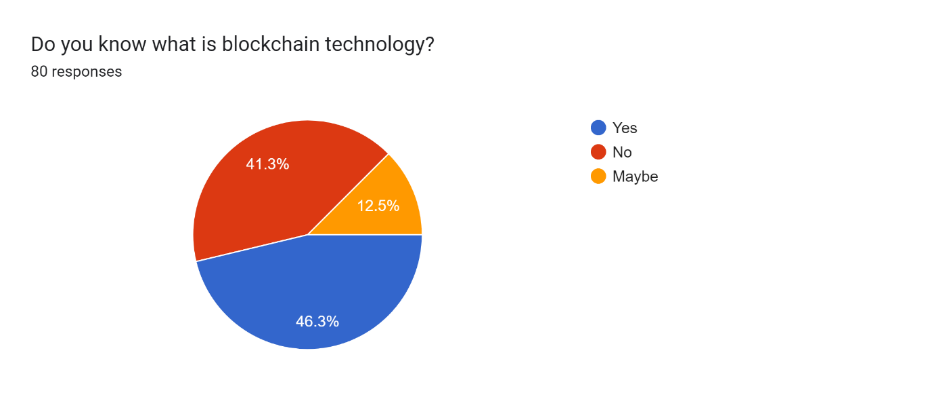


Figure 37: Data of Question 10

From the data above, 46.3% of the population understand fully about blockchain. As for the other remaining percentage, either they do not know what is blockchain or they have only a little understanding of blockchain. This allows the researcher to understand that not a lot of individuals understand what is blockchain technology.

**Question 11**

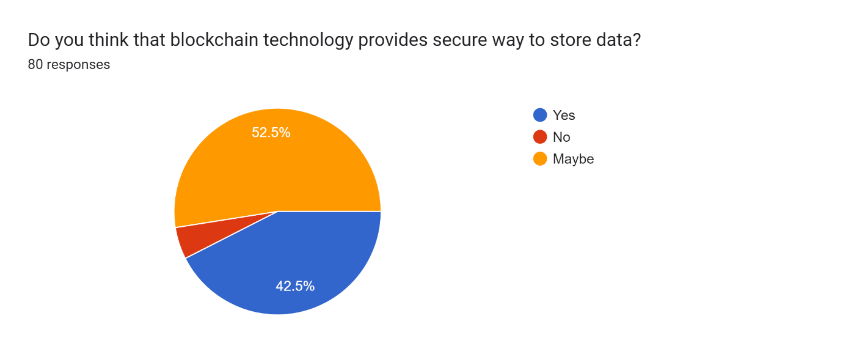


Figure 38: Data of Question 11

From this chart shown above, most of the responds indicate that they are not sure about blockchain technology on providing a safe and secure way to protect their data. The answer also might have an affect due to from the previous question, it is known that most of the population do not really understand or do not understand about blockchain technology.

**Question 12**

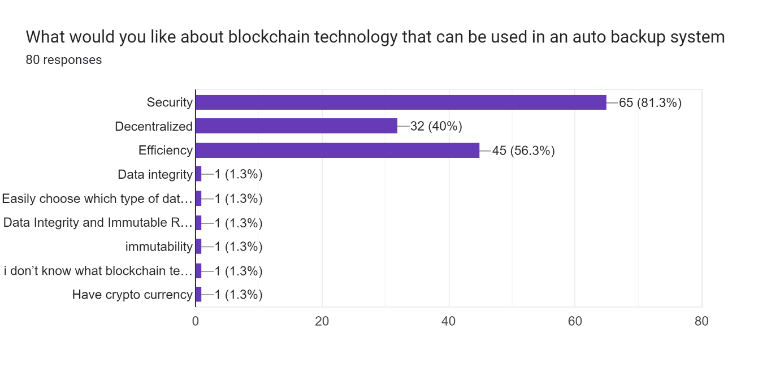
****

Figure 39: Data of Question 12

For this question, it is to be used to understand what the population view on the used of blockchain technology in the backup system. From the graph above, with 65 votes, security is the reason the population thought that can be used in the backup system. It is also a notable mention that the decentralized and efficiency of blockchain technology is also voted as this may help the backup system to be more favourable.

**Question 13**

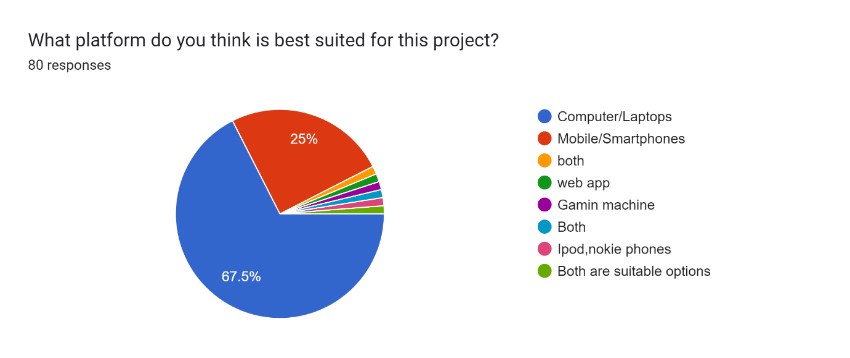
****

Figure 40: Data of Question 13

From the chart above, there are many options that has been suggested. The two most dominant option are computer/laptops and mobile/smartphones, with computer/laptop still have the greater percentage. This can be de to the population stores more important data in their computer, which will require the individual to choose computer/laptops over other devices or platform.

**Question 14**

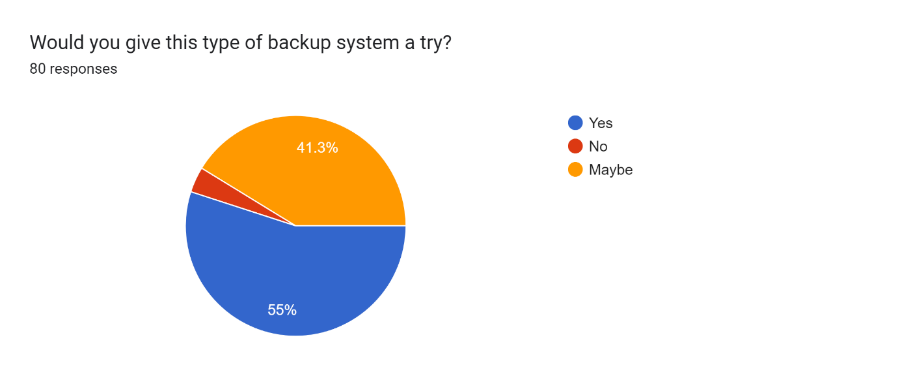
****

Figure 41: Data of Question 14

One of the more important questions to this project is to understand if the population if willing to be able to give this project a try. From the pie chart most of the population are willing to give the project a try with the percentage of 55%. 41.3% od the population are not sure if they are willing to give the project a try as they might see the product and decide afterwards as this type of project is a new type of project. From this question, the researcher can conclude that in general, the population are willing to give the project and the product a try.

## 6.2 Summary

After conducting the survey to understand the population better, the researcher can come into conclusion and understand about his project better. From the question asked, the population thinks that a backup system is needed and important, and it is important to have a backup system. The survey also asked about the security feature that might allow the population to be able to select a particular backup system. Blockchain technology is also covered in the survey and in general, the population do not really understand blockchain technology. The survey can be conducted better in a way that the target would be individual has better understanding behind blockchain technology and backup system itself. Since the research method is quantitative, it is one of the variables that has to take into the matter. After conducting the survey, the researcher think that it is a better choice to be able to do a hybrid research method as it may be possible to get a more accurate data.

# Chapter 7: Conclusion

In summary of the report that has been conducted, there are a lot that has been achieve in the report. Covering the aims, objective and why the researcher have chosen making an auto-backup system with AES 256-bit encryption all have been explain in the report above. The researcher has also conducted intensive research about this topic found in chapter 2 of this report. In the research that the researcher has conducted, topic such as backup system, cyber security, ransomware and blockchain can be found and has made the researcher to understand more about each respective topic. There is also 3 similar system that has been compared to further understand the project better.

As the research and understanding the population better, the quantitative research method was opted. As stated before, it is better that in this research the hybrid method was used instead to understand the population and get valuable insight from specialist that understand backup system and blockchain better. As this would be benefit as more understanding can help the researcher to understand the project better. The results from the survey conducted itself has open the researcher on how the population would understand the project and has gotten valuable insight on the project that will be created.

There are still a lot of places that can be research on and have a further understanding of such as blockchain technology security features and study of the trends in cyber security of ransomware. Overall, the research that has been done by the researcher has provided sufficient information and data to be able to carry out the research about the project.

# References

Cadenas, C. (2023). How Many Computers Are In The World? [Up-to-Date Stats]. *Windows Report - Error-free Tech Life*. https://windowsreport.com/how-many-computers-are-in-the-world-2022/#:~:text=Forrester%20states%20that%20most%20of,403.1%20million%20units%20in%202023.

Mathieu. (2023, April 18). *Why is computer data backup important?* Bocasay. https://www.bocasay.com/importance-data-back-up/

*What is Backup? (Data Backup) Comprehensive Guide - Acronis*. (n.d.). Acronis. https://www.acronis.com/en-us/blog/posts/data-backup/#:~:text=Backup%20is%20the%20process%20of,deleted%20them%20from%20your%20system.

Evangelist, S. M. P. (2022). Backup Evolution: Traditional vs Modern Backup. *bluexp.netapp.com*. https://bluexp.netapp.com/blog/cbs-blg-backup-evolution-traditional-vs-modern-backup#h\_h2

Gondek, C. (n.d.). *How Long Has Blockchain Technology Been Around?* OriginStamp. https://originstamp.com/blog/how-long-has-blockchain-technology-been-around/#:~:text=Blockchain%20technology%20is%20a%20revolutionary,widely%20adopted%20across%20various%20industries.

*What is Blockchain Technology - IBM Blockchain  | IBM*. (n.d.). https://www.ibm.com/topics/blockchain

Yev. (2022). The State of Backups: Who’s Most at Risk. *Backblaze Blog | Cloud Storage & Cloud Backup*. https://www.backblaze.com/blog/the-state-of-backups-whos-most-at-risk/

SpidersC. (n.d.). *Reddit - Dive into anything*. https://www.reddit.com/r/Windows10/comments/amxgdm/does\_anyone\_else\_think\_onedrive\_is\_annoying/

Imber, D. (2023, June 9). The Latest Cyber Crime Statistics (updated June 2023) | AAG IT Support. *AAG IT Services*. https://aag-it.com/the-latest-cyber-crime-statistics/

Zesium\_Admin. (2022, December 13). *How is Blockchain Boom Affecting the Software Industry - Zesium*. Zesium. https://zesium.com/how-is-blockchain-boom-affecting-software-industry/

Sahu, M. (n.d.). Top 12 Commerce Project Topics & Ideas in 2023 [For Freshers]. *upGrad blog*. https://www.upgrad.com/blog/how-to-use-blockchain-to-store-data/

Simply Explained. (2017, November 13). *How does a blockchain work - Simply Explained* [Video]. YouTube. https://www.youtube.com/watch?v=SSo\_EIwHSd4

Nonprofit Technology Collaboration & Baylor. (2013). Back Up Data. *Back up Data*, 1–9.

Chervenak, A., Vellanki, V., & Kurmas, Z. (1998). Protecting File Systems: A survey of Backup techniques. *ResearchGate*. https://www.researchgate.net/publication/2431989\_Protecting\_File\_Systems\_A\_Survey\_of\_Backup\_Techniques

Yarrapothu, S. (n.d.). Effectiveness of Backup and Disaster Recovery in Cloud: A Comparative study on Tape and Cloud based Backup and Disaster Recovery. *Effectiveness of Backup and Disaster Recovery in Cloud*, 8–34.

Ruggiero, P., & A. Heckathorn, M. (2012). Data Backup Options. *Data Backup Options*. https://www.cisa.gov/sites/default/files/publications/data\_backup\_options.pdf

Abdumalikov, G. (2022). INTERNATIONAL JOURNAL ON HUMAN COMPUTING STUDIES. *Profound Importance of Cyber Security in the Field of Business*, *4*(2), 43–46. https://media.neliti.com/media/publications/408257-profound-importance-of-cyber-security-in-a2668cec.pdf

Supayah, & Ibrahim, J. (2016). AN OVERVIEW OF CYBER SECURITY IN MALAYSIA. *Kuwait Chapter of Arabian Journal of Business and Management Review*, *6*(4), 12–20.

Ghundare, S., Patil, A., & Lad, R. (2020). Importance of Cyber Security. *International Journal of Engineering Research & Technology (IJERT)*, *8*(05), 1–3.

Sharma, A. K., & Sharma, R. (2023). Cyber Security and Indian Cyber Laws. *Cyber Security and Indian Cyber Laws*.

Kalakuntla, R., Vanamala, A., & Kolipyaka, R. R. (2019). Cyber Security. *Holistica*, *10*(2), 115–128. https://doi.org/10.2478/hjbpa-2019-0020

Simoiu, C., Gates, C., Bonneau, J., & Goel, S. (2019). “I was told to buy a software or lose my computer. I ignored it”: A study of ransomware. *“I Was Told to Buy a Software or Lose My Computer. I Ignored It”: A Study of Ransomware*.

Barracuda. (2020). 2023 ransomware insights: The prevalence and impact of ransomware attacks around the world. *Market Report*.

Shah, N., & Farik, M. (2017). Ransomware-Threats, Vulnerabilities And Recommendations. *ResearchGate*. https://www.researchgate.net/publication/317913092\_Ransomware-Threats\_Vulnerabilities\_And\_Recommendations

Humayun, M., Zaman, N., Alsayat, A., & Ponnusamy, V. (2021). Internet of things and ransomware: Evolution, mitigation and prevention. *Egyptian Informatics Journal*, *22*(1), 105–117. https://doi.org/10.1016/j.eij.2020.05.003

Alessandrini, A. (n.d.). RANSOMWARE Hostage Rescue Manual. *Know Be4*, 1–14.

Fuchs, P. (n.d.). Blockchain. *Mercer*, 1–22.

Mgcub, A. S. . R. S. (n.d.). *An introduction to blockchain technology and its application in libraries*. DigitalCommons@University of Nebraska - Lincoln. https://digitalcommons.unl.edu/libphilprac/6630

Saha, S. (2019). A Study on Blockchain Technology. *papers.ssrn.com*. https://doi.org/10.2139/ssrn.3477373

Crosby, M., Nachiappan, Pattanayak, P., Verma, S., & Kalyanaraman, V. (2015). BlockChain Technology Beyond Bitcoin. *Berkeley University of California*, 1–25.

Zheng, Z., Xie, S., Dai, H., & Wang, H. (2017). An overview of blockchain technology: architecture, consensus, and future trends. *ResearchGate*. https://doi.org/10.1109/BigDataCongress.2017.85

Ninja, T. (2023, July 21). The Difference Between Cloud Storage & Cloud Backup | NinjaOne. *NinjaOne*. https://www.ninjaone.com/blog/cloud-storage-vs-cloud-backup-know-the-difference/#:~:text=Cloud%20backup%20securely%20backs%20up,and%20share%20with%20other%20devices.

By Aleksander Hougen (Managing Editor) &  Aleksandar Kochovski (Editor)— Last Updated: 24 May’23 2023-05-24T19:28:58+00:00. (2023, May 24). *IDrive Review 2023 [Backup features, speed & pricing]*. Cloudwards. https://www.cloudwards.net/review/idrive/#:~:text=IDrive%20works%20by%20encrypting%20your,you%20can%20decrypt%20the%20data.

*About IDrive®*. (n.d.). https://www.idrive.com/online-backup-company#:~:text=The%20company's%20core%20services%20include,account%20for%20the%20same%20cost.

Kananda, V. (2022, July 8). Why You Should Use AES 256 Encryption to Secure Your Data. *Ipswitch*. https://www.ipswitch.com/blog/use-aes-256-encryption-secure-data#:~:text=Advanced%20Encryption%20Standard%20(AES)%20256,a%20lot%20easier%20from%20here.

*256-bit AES encryption for complete protection - IDrive®*. (n.d.). https://www.idrive.com/enterprise/security#:~:text=Our%20physical%20security%20features%20include,surveillance%20and%20security%20breach%20alarms.

Rosenberg, E. (2022). Storj (STORJ) Cryptocurrency Definition. *Investopedia*. https://www.investopedia.com/storj-cryptocurrency-definition-5224160

Kriptomat. (n.d.). *What is Storj & how does it work? Who created STORJ?* https://kriptomat.io/cryptocurrencies/storj/what-is-storj/#:~:text=Storj%20is%20among%20the%20first,decentralization%2C%20privacy%2C%20and%20transparency.

*Storage that’s so secure, even we can’t see your data.* (n.d.). https://www.storj.io/benefits/security#:~:text=Secure%20by%20Default,Storage%20Nodes%20across%20the%20globe.

Young, S. (2023, March 7). *Pick the Best Data Backup Solution for your Business - PennComp*. PennComp. https://penncomp.com/data-backup-solution-for-your-business/

*Sia - The Sia Foundation*. (n.d.). Sia. https://sia.tech/about-sia-foundation

*Should I use IPFS or SIA? - Filebase*. (n.d.). https://docs.filebase.com/what-is-filebase/storage-networks/should-i-use-ipfs-or-sia

*Schneier on Security: Threefish*. (n.d.). https://www.schneier.com/academic/skein/threefish/

*Sia - Decentralized data storage*. (n.d.). Sia. https://sia.tech/

Sharma, T. K. (2022). Java or Python: Which suits blockchain better? *Blockchain Council*. https://www.blockchain-council.org/blockchain/java-or-python-which-suits-blockchain-better/#:~:text=Developers%20can%20use%20Python%20to,need%20to%20write%20much%20code.&text=Python%20simplifies%20developers'%20lives%20as,developers%20to%20work%20in%20blockchain.

Ivey, A. (2023b, July 24). History of Python programming language. *Cointelegraph*. https://cointelegraph.com/news/history-of-python-programming-language#:~:text=Birth%20of%20Python,to%20read%2C%20write%20and%20maintain.

UNext Editorial Team. (2023, March 17). History of Java Programming Language| UNEXT | UNEXT | UNEXT. *UNext*. https://u-next.com/blogs/java/history-of-java/#:~:text=Java%20was%20developed%20in%20the,Embedded%20Systems%2C%20and%20corporate%20solutions.

Coding Ninjas. (n.d.). *Coding Ninjas Studio*. https://www.codingninjas.com/studio/library/java-vs-python-and-c

Wise, J. (2023). Windows users: How many people use Windows in 2023? - EarthWeb. *EarthWeb*. https://earthweb.com/windows-users/#:~:text=Windows%20runs%20on%201.6%20billion%20active%20devices.

Gillis, A. S. (2019). systems development life cycle (SDLC). *Software Quality*. https://www.techtarget.com/searchsoftwarequality/definition/systems-development-life-cycle

Hoory, L. (2022, March 25). What is waterfall methodology? Here’s how it can help your project management strategy. *Forbes Advisor*. https://www.forbes.com/advisor/business/what-is-waterfall-methodology/

*Agile vs. Waterfall Methodology Explained | PM Tips for Non-PMs*. (n.d.). [Video]. TeamGantt. https://www.teamgantt.com/waterfall-agile-guide/waterfall-methodology

Sacolick, I. (2022, April 8). *A brief history of the agile methodology*. InfoWorld. https://www.infoworld.com/article/3655646/a-brief-history-of-the-agile-methodology.html

Nehra, M. (2022, May 11). 6 stages of the Agile Development Lifecycle. *decipherzone.com*. https://www.decipherzone.com/blog-detail/agile-development-lifecycle

Kinsbruner, E. (n.d.). *The complete guide to different types of testing | Perfecto by Perforce*. Perfecto by Perforce. https://www.perfecto.io/resources/types-of-testing

Indeed Editorial Team. (2023). Types of research methods (With best practices and examples). *Indeed.com*. https://www.indeed.com/career-advice/career-development/types-of-research-methods#:~:text=Professionals%20use%20research%20methods%20while,using%20numbers%20to%20measure%20data.

Vaughan, T. (2021, August 5). 10 Advantages and Disadvantages of Qualitative research. *Poppulo*. https://www.poppulo.com/blog/10-advantages-and-disadvantages-of-qualitative-research

DeVault, G. (2020). Advantages and Disadvantages of quantitative research. *LiveAbout*. https://www.liveabout.com/quantitative-research-advantages-and-disadvantages-2296728

McCombes, S. (2023). Sampling Methods | Types, Techniques & Examples. *Scribbr*. https://www.scribbr.com/methodology/sampling-methods/

*Probability Sampling: Definition,Types, Advantages and Disadvantages - Statistics How To*. (2021, May 31). Statistics How To. https://www.statisticshowto.com/probability-and-statistics/sampling-in-statistics/probability-sampling/

*Non-Probability Sampling: Definition, Types - Statistics How to*. (2023, March 3). Statistics How To. https://www.statisticshowto.com/probability-and-statistics/sampling-in-statistics/non-probability-sampling/#:~:text=Advantages%20and%20disadvantages&text=It's%20also%20easy%20to%20use,you%20are%20representing%20the%20population.

# Appendices

## Log Sheets

### Log Sheet 1

A project log sheet with text and images

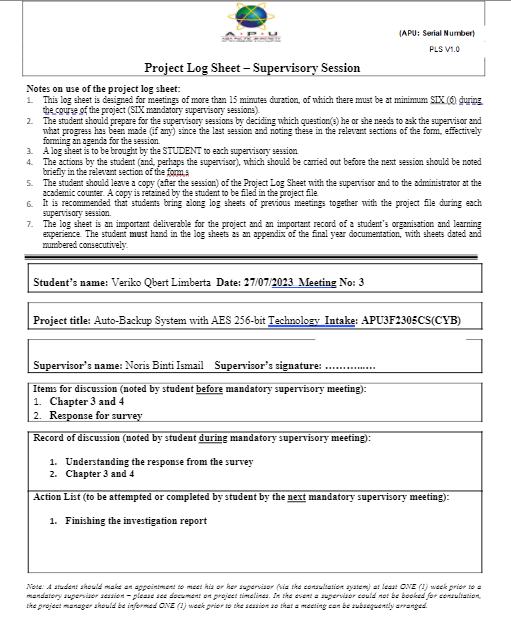
Description automatically generated

### Log Sheet 2

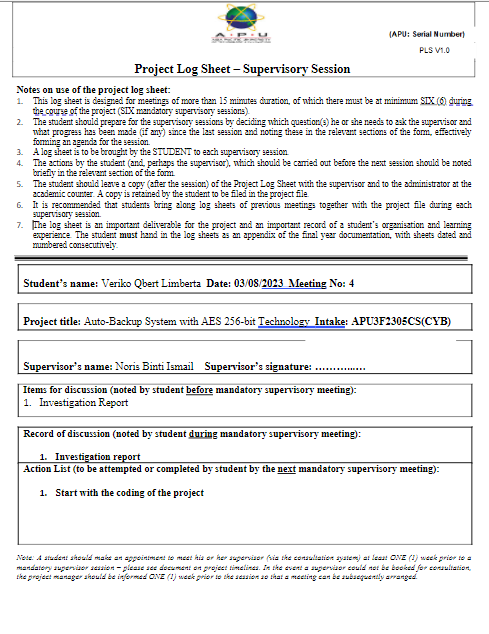
A document with text and images

Description automatically generated

### Log Sheet 3



### Log Sheet 4



## PPF

A close-up of a document

Description automatically generated

A close-up of a paper

Description automatically generated

A close-up of a document

Description automatically generated

A screenshot of a document

Description automatically generated

A close-up of a document

Description automatically generated

A document with text on it

Description automatically generated

A screenshot of a document

Description automatically generated

## Ethics Form

A close-up of a form

Description automatically generated

A close-up of a form

Description automatically generated

A document with a black text

Description automatically generated with medium confidence

A document with a letter

Description automatically generated with medium confidence

## Gantt Chart Semester 1

A screenshot of a computer

Description automatically generated