

Project: Vault Guard (Back up security)

Name: Edwin ngila kyalo

ADM.NO 240166

DIPLOMA IN CYBER SECURITY AND ETHICAL HACKING INSTITUTE OF SOFTWARE TECHNOLOGIES

2024

# DECLARATION

I, Edwin ngila kyalo of AdmNo. 240166, certify that this project is my own work, based on my personal study and research done. I acknowledge all the resources and materials used to come up with this project whether it was articles, class notes, reports, and any other kind of documentation.

I also certify and confirm that this project has never been submitted anywhere for academic audit and that it has not been copied in whole from any source or otherwise plagiarized from any source or persons.

I confirm that I have identified and declared all possible conflicts that I may have.

Signed and confirmed by: -

Candidate Name:

Candidate signature:

Date signed:

Supervisors Name:

Supervisors Signature:

Date signed:

# ACKNOWLEDGMENT

I extend my heartfelt gratitude to my beloved family and friends for their unwavering support throughout this project. Additionally, I am deeply thankful for the invaluable resources provided by the Institute, without which this endeavor would not have been possible. Your collective contributions have played an indispensable role in the success of this project.

# ABSTRACT

Vault Guard is an advanced online backup system designed to offer users seamless data storage solutions. Through this platform, users can effortlessly create an accounts and securely upload their data to an online Vault, where file paths are meticulously stored in a relational database. Once uploaded, the data undergoes a strict scanning process using cutting-edge third-party APIs to detect and eliminate any potential viruses. In the event of a virus detection, users are promptly notified, ensuring that infected files are not uploaded. Following this, robust encryption measures are applied, with encryption keys securely stored within the database. The encrypted data is then stored on a server equipped with well-organized file management systems. Users enjoy flexibility, being able to upload a diverse array of file types, including MP3 files, Word documents, PDFs, photos, and videos. Furthermore, authenticated users have the privilege to download their files from the online vault or seamlessly upload new ones. The project embodies the principles of Confidentiality, Integrity, and Availability (CIA), according to *Washington University in St. Louis (2024)* “A comprehensive information security strategy includes policies and security controls that minimize threats to these three crucial components” incompliance to the quote, data remains confidential through authentication protocols, files are safeguarded against tampering through encryption, and data remains readily accessible on the internet to conclude that the project upholds the cybersecurity core principles.

# Abbreviation and Definition

**Back up** - refers to the process of making copies of data or files to prevent loss in case the original data is lost, corrupted, or accidentally deleted.

**PDF-** stands for Portable Document Format. It is a file format developed by Adobe.

**MP3-** stands for MPEG Audio Layer III. It is a digital audio encoding format that compresses audio

Table of Contents

[DECLARATION i](#_Toc170119630)

[ACKNOWLEDGMENT ii](#_Toc170119631)

[ABSTRACT iii](#_Toc170119632)

[Abbreviation and Definition iv](#_Toc170119633)

[1 Chapter One: Introduction 1](#_Toc170119634)

[1.1 Background Of The Study 1](#_Toc170119635)

[1.2 Objectives 2](#_Toc170119636)

[1.3 Problem Statement 2](#_Toc170119637)

[1.4 Scope of Project 2](#_Toc170119638)

[2 Chapter Two: Literature Review 3](#_Toc170119639)

[2.1 Introduction 3](#_Toc170119640)

[2.2 what is backup security? 3](#_Toc170119641)

[2.3 Types of Backups 3](#_Toc170119642)

[2.4 Best practices as a backup site to consider 4](#_Toc170119643)

[2.5 Disadvantages of using physical data backup methods 4](#_Toc170119644)

[2.6 Viruses corrupting data 5](#_Toc170119645)

[2.7 Data analysis and validation 5](#_Toc170119646)

[2.8 Evaluation 5](#_Toc170119647)

[3 Chapter Three: Project methodology 6](#_Toc170119648)

[3.1 Introduction 6](#_Toc170119649)

[3.2 Representation of the methodology 6](#_Toc170119650)

[3.2.1 Planning 7](#_Toc170119651)

[3.2.2 Design 7](#_Toc170119652)

[3.2.3 Develop 7](#_Toc170119653)

[3.2.4 Test 7](#_Toc170119654)

[3.2.5 Deploy 7](#_Toc170119655)

[3.2.6 Review 7](#_Toc170119656)

[3.2.7 Advantages of using Agile 8](#_Toc170119657)

[3.2.8 Disadvantages of using Agile 8](#_Toc170119658)

[4 Chapter Four: Project Implementation 9](#_Toc170119659)

[4.1 Project duration 9](#_Toc170119660)

[4.2 Implimentation of security features 9](#_Toc170119661)

[4.2.1 Authentication security 10](#_Toc170119662)

[4.2.2 API security 11](#_Toc170119663)

[5 Chapter Five 12](#_Toc170119664)

[6 Chapter Six 13](#_Toc170119665)

[7 Chapter Seven 14](#_Toc170119666)

[8 REFERENCES 15](#_Toc170119667)

[9 Appendix 16](#_Toc170119668)

# Chapter One: Introduction

Vault Guard is an online backup system that enables users to create an account and upload data to the online Vault. Subsequently, the data undergoes scanning for viruses via third-party APIs, followed by encryption and storage on a server with well-organized files. Users have the flexibility to upload various file types, including MP3 files, Word documents, PDFs, photos, and videos. From the online vault, authenticated users can download their files or even upload them to the online vault.

## Background Of The Study

According to *True List, 2024* studies show that 62% of people who own a computer have lost data at some point, while 76% have deleted something unintentionally together with a survey conducted by Shulmistra, D. (2023, December 12) states that “*76% of respondent’s IT leaders said that they experienced a severe loss of critical data in the past year, and 45% of them lost data permanently*”, This shows that 45% of the 76% did not have a backup for their data, concluding that data loss is a big challenge for people and organizations in the world hence for the need of Vault Guard.

In today's technological world, data is widely recognized as a critical asset. Some individuals go so far as to consider it invaluable, while companies such as secureprivacy.ai regard data as "the new gold". Given its significance to many, protecting data is paramount. Hence, initiatives like Vault Guard has to be developed to enable individuals to create copies of their crucial files and securely store them in an accessible manner.

Top of Form

## Objectives

The objective of the project is to offer users an alternative method for data backup, distinct from the traditional approach of storing data on hard disks and flash drives, which are susceptible to loss.

* Protect user data. Though authentication of users before accessing the stored data and encrypting data before it is stored in the server.
* Making data readily accessible to the user when the user wants access.
* Providing users, a quick recovery of data in case of a data lose.
* Cost effective providing users a cost free way of storing their information.

## Problem Statement

Form the back ground of the study we learn that 76% of people in the world experience loss of data. This is a big problem that people are carently facing. People are complaining a lot about losing valuable data in their devices and the traditional methods of backing up data to an external hard drive is also not working out for them because most of the time the data gets corrupted or even the hardware it’s self gets lost or even damaged.

## Scope of Project

This project is about providing a storage solution to users who are looking for a way they can back up their data just in case something happens to the original copy, in addition the user can access the data from any ware so long as the user knows the password to their account and has access to the internet. Also users have no worry about their data been leaked because the data will be encrypted and stored in the server and the key is stored in the data base whenever the user whats the data back the key is retrieved back to and the data is then decrypted and the user can view the data or even download it back to their devices.

# Chapter Two: Literature Review

## Introduction

In literature review, we are going to cover documented work done by other scholars under the topic of data back up together with the security measures that are put into place in help keeping user data secure, reliable and accessible to the intended users. Note that the literature review presented below is according to what the project aims to achieve and resolve.

## what is backup security?

*“"Backup security" is a term used to describe making extra copies of data, files, or even entire databases to keep them safe from loss or damage*.” According to *ReasonLabs* *in 2023*, having backups is a smart way to protect your important information. It is not just big companies that need to do this; everyone should make sure they have a backup of their digital data. By creating and keeping a copy in a secure place, you can be sure that if something goes wrong like your computer crashes or you accidentally delete a file you will still have access to your information. This habit can save you a lot of trouble and stress in the long run.

## Types of Backups

There are three types of data backup that are commonly known and used which are: -

* **Full back up:** The most basic and comprehensive backup method, where all data is sent to another location.
* **Incremental backup:** Backs up all files that have changed since the last backup occurred.
* **Differential backup:** Backs up only copies of all files that have changed since the last full backup.

For the project we are going to impliment Fullback up where the users have the ability to upload a file to the site and can also be able to retrieve the file as it is.

## Best practices as a backup site to consider

The following are some of the considerations that have been stated by *Crump, G. (2019, July 10)* by the article that he wrote on best practices that companies should consider when performing data backup.

* **Robust Security Measures –** Implement strong security measures to protect user data. This includes encryption at rest and in transit, multi-factor authentication (MFA), and secure user authentication mechanisms.
* **User-Friendly Interface –** Design a user-friendly interface that is easy to navigate and understand. Users should be able to quickly access and manage their data without complicated steps. Provide clear instructions and helpful resources.
* **Collaboration Features -** Include collaboration features that allow multiple users to work on the same documents or projects. This can be especially useful for business users who require shared access and teamwork capabilities.
* **Scalability and Performance -** Design the system to scale as user demand grows. Performance should remain consistent regardless of the number of users or the amount of data stored. Implement measures to ensure low latency and quick response times.

## Disadvantages of using physical data backup methods

According to an article written by *Open Learning. Retrieved May 7, 2024,* shows that physical devices are the main method used by majority of users with pc to back up their data compered to virtual methods of data storage which is widely enlarged by many of the companies

The following are some of the highlighted disadvantages of physical storage: -

* Hard disks are fragile and easily damaged if dropped or exposed to extremely high temperatures or magnetic fields.
* If small hard disks are used once to make a backup then archived, the replacement cost is much higher than for tape or optical media.
* SSDs are more expensive than equivalent capacity magnetic disks.
* SSDs and memory cards are only useful for short term storage of up to 5 years

## Viruses corrupting data

According to an article written by *Corporate Finance Institute. Retrieved May 7, 2024* Computer viruses pose a significant threat to data stored on hard drives and within corporate network systems. These malicious programs can steal, corrupt, encrypt, or delete critical information. Moreover, they have the potential to spread across an entire organization's network, disrupting the operation of computer hardware.

## Data analysis and validation

Form the data that I have collected and analyzed in this sections indicates that people have not really adapted to the new way of utilizing available cloud storage as a method of back up, and if they are the service providers have not implimented the needed concerns to make the user experience the best and the user data more secure and easy to access. As a storage or rather a backup service platform user that are using the services are entitled to their data and can access it at any time they need and their data should be secured and not be distributed.

## Evaluation

The things that should be considered in the next step of the documentation are: -

* The implimentation of data encryption to the user uploaded files.
* Conducting virus checks to files been uploaded to the cloud.
* Implimentation of user authentication

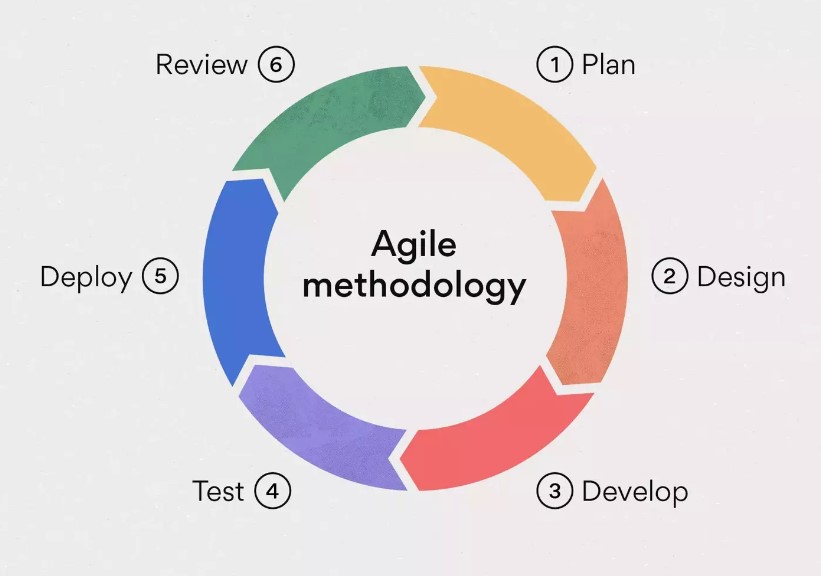
# Chapter Three: Project methodology

## Introduction

Project methodology outlines the approach and steps taken to develop a system. There are various methodologies available, such as Agile, Waterfall, and Scrum, each with its own unique framework. For this project, the Agile methodology has been chosen as the most suitable.

Agile divides the project into manageable units called sprints, allowing tasks to be distributed among different team members. This approach is especially effective when working with large teams, as it promotes collaboration and flexibility. Agile emphasizes meeting customer needs and prioritizes the software's functionality over exhaustive documentation. This methodology allows for iterative development and encourages regular feedback from stakeholders, ensuring that the final product aligns with user expectations.

## Representation of the methodology



### Planning

In this sage the project manager is needed to speaks to the customer and know the customer requirements. The manager is also tasked to review the initial working system and come up with a documentation that is going to guide the developers during the development stage of the Agile life circle. At this stage the project manager is needed to come up with a project schedule showing the time that it will take to came up with the end product. The project manager is given the task of identifying both the front end and back end team and briefing them.

### Design

In this stage the team is tasked to review the documentation given by the project manager which has the requirements given by the customer and design the system around the requirements. In this stage the developers also come up with sketches for the system.

### Develop

In this stage the project manager is tasked on selecting a team of front end and back end developers which are then divided up into scrum teams in both sides, in the front side and the back side. Then the project manager then gives each scrum team a sprint to work on and each sprint is given a date to which the work should be submitted.

### Test

After the development stage the system is presented as a beater version and given to the testing team and the team takes the system through harsh conditions in order to know the holes the system has, so as they can patch up the system and test the system.

### Deploy

In this stage the system is presented to the customer and some members of the development team is tasked to educate the customer on how to use the system and desensitize them from the initial system. At this point the developers also help the customers to migrate to the new system.

### Review

Because the Agile methodology is more flexible the review stage is open at all times during the development process of the system, which means that in this stage the customer has the ability to

him in ideas at any time and also after the project has been delivered to them. This stage is also when the developer also updates the system per the user requirements.

### Advantages of using Agile

The advantages of using Agile methodology are: -

1. Agile is more flexible it can start from any point of the software life cycle.
2. Fast time delivery to the market due to the distribution of tasks.
3. Delivering of a quality product to the customer.
4. There is a risk redaction.
5. Competition among developers of the company.
6. Higher customer satisfaction because the customer is involved in the development process.
7. Continues delivery and continues improvements to the system.

### Disadvantages of using Agile

The disadvantages of using Agile methodology are: -

1. It is less predictable.
2. More time and commitment is needed.
3. It is demanding for both the customers and developers.
4. The system entropy is likely to affect a system developed through the methodology.
5. Little or no documentation of the end product because the system methodology focuses highly on customer satisfaction.

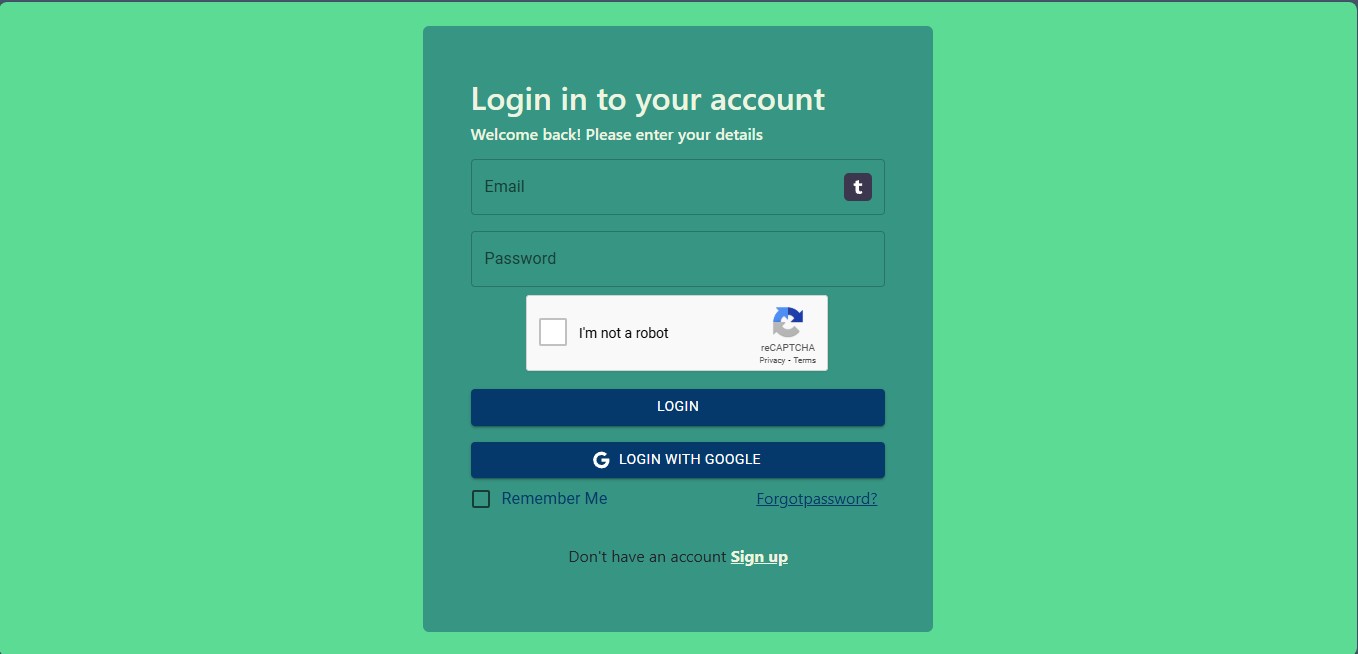
# Chapter Four: Project Implementation

For the application, I decided to implement ReactJS, a JavaScript library, for the frontend. For the backend services, I chose Firebase, which is a Backend as a Service (BaaS) provided by Google Cloud Services. Firebase comes with a NoSQL database and cloud storage. Since the project focuses on user data storage, we have heavily utilized both of these products along with Firebase's user authentication features.

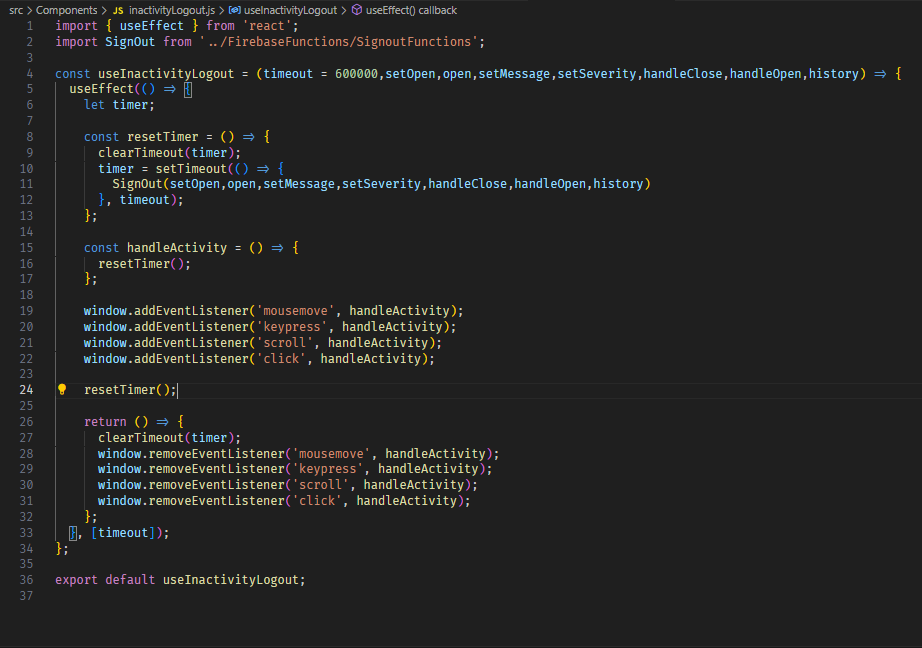
## Implimentation of security features

### Authentication security

To enhance the authentication process for our application, we have implemented a login page for user access. Additionally, we have integrated Google reCAPTCHA to secure the application against brute force attacks.

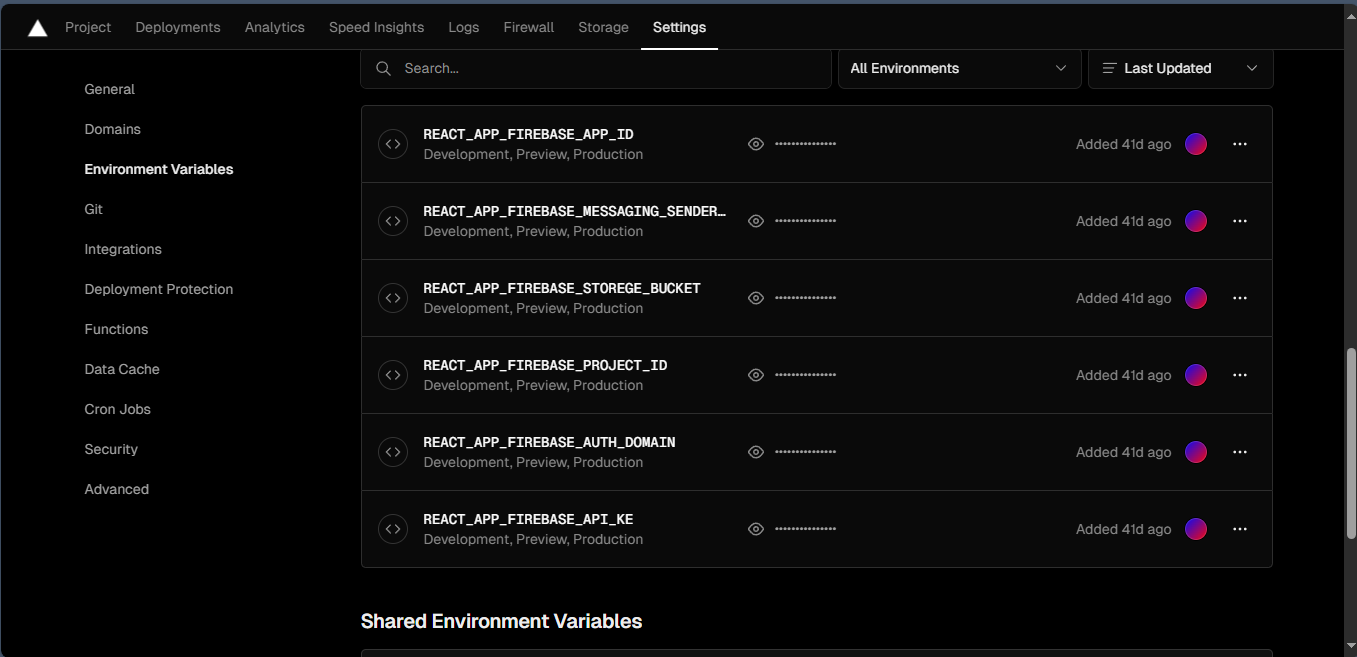


Implimentation of logout due to inactivity. This is a security implimentation that I have done for the application that logs out the user after 10 minutes of inactivity the code bellow shows how it is set to listen if the user’s mouse is not moving, keypress, scroll and click motions

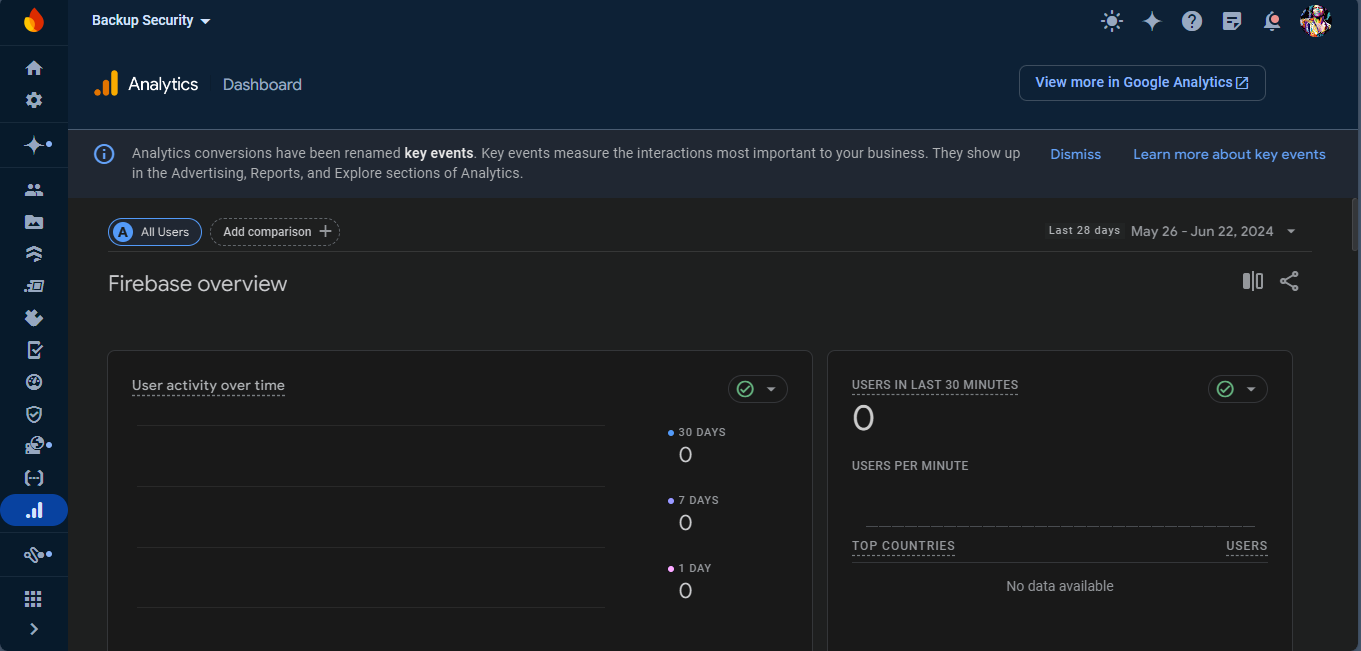


### API security

During the deployment stage, we ensured that the application's environment variables are hidden from attackers on Vercel. Instead of deploying the application along with the environment variable keys, which could be easily accessed by anyone visiting the site, we have securely managed these variables within Vercel's environment settings.



Additionally, we have implimented the use of google analytics that can help track anomalies within the system in case things are not adding up for the admin

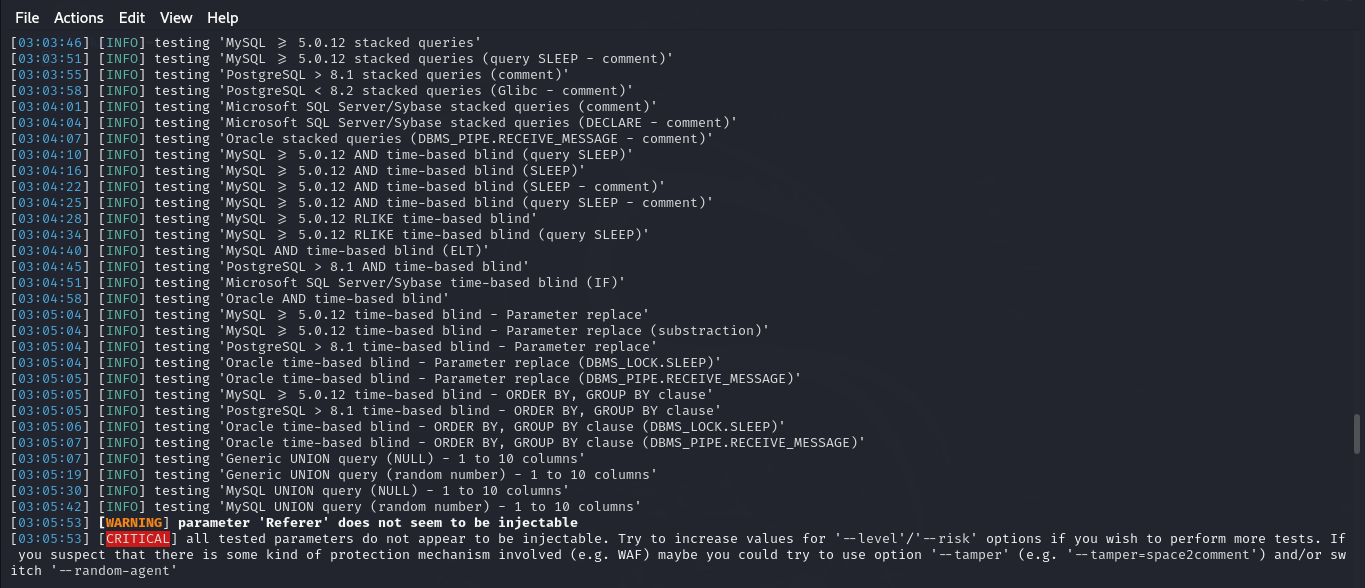


### Database security

In the process of development, we were able to impliment the use a non-SQL database to prevent hackers from compromising the security of the database. I conducted a test using SQLmap to test if the site can be compromised using SQL injections. The following was the outcome of the test.

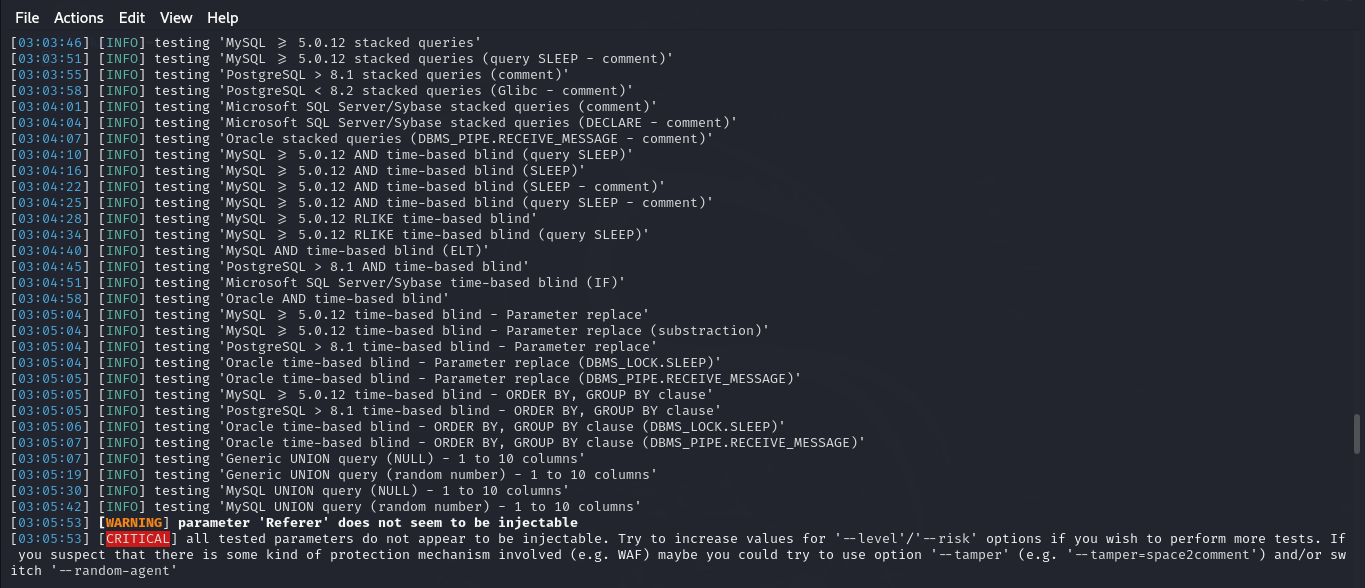
Test one when the breach level is 3





Test one when the breach level is 5





# REFERENCES

* *Confidentiality, Integrity, and availability: The CIA triad | Office of Information Security | Washington University in St. Louis*. (n.d.). <https://informationsecurity.wustl.edu/items/confidentiality-integrity-and-availability-the-cia-triad/>
* *True List. (2024, February 17). 11 Data loss statistics to know in 2024 | True List.* <https://truelist.co/blog/data-loss-statistics/#:~:text=While%20there%20are%20no%20exact,76%25%20have%20deleted%20something%20unintentionally>.
* *Shulmistra, D. (2023, December 12). 15 Data loss statistics All businesses should know.* <https://www.linkedin.com/pulse/15-data-loss-statistics-all-businesses-should-know-dale-shulmistra-y0lje#:~:text=Severe%20data%20loss%20affects%20three,of%20them%20lost%20data%20permanently>
* *Unlocking the significance of data in the digital era | Power, value, and protection. (2024, March 15).* https://secureprivacy.ai/. <https://secureprivacy.ai/blog/significance-of-data-in-digital-era#:~:text=Data%20is%20of%20paramount%20significance,customer%20experience%2C%20and%20monetize%20data>.
* *Crump, G. (2019, July 10). The 7 critical backup strategy best practices to keep data safe. SearchDataBackup.* https://www.techtarget.com/searchdatabackup/feature/The-7-critical-backup-strategy-best-practices-to-keep-data-safe

# Appendix

## Appedix I

### Project duration

Below is a Gantt graph that shows the development stages of the application and shows the time line the app took to get to production.

