

# DO NOT LEAK (DNL)

---

A Second Year Project Report

Submitted to the Faculty

of the

Bennett University

By

Abhinav Robinson

E18CSE006

Ashok Kumar

E18CSE029

Bhavya Batra

E18CSE035

Srishti Verma

E18CSE181

Department of Computer Science Engineering

2020

Greater Noida-201310, Uttar Pradesh, India

## **TABLE OF CONTENTS**

### **LIST OF TABLES**

### **LIST OF FIGURES**

#### **1. INTRODUCTION**

1.1. Problem Statement

#### **2. Background Research.**

2.1. Proposed System..

2.2. Goals and Objectives.

#### **3. Project Planning.**

3.1. Project Setup.

3.2. Stakeholders.

3.3. Project Resources.

3.4. Assumptions.

#### **4. SYSTEM ANALYSIS AND DESIGN**

4.1. Overall Description.

4.2. Users and Roles.

4.3. User Stories (Requirements)

4.4. Design diagrams/ UML diagrams/ Flow Charts/ E-R diagrams.

4.4.1. Use Case Diagrams.

4.4.2. Class Diagram..

4.4.3. Activity Diagrams.

4.4.4. Sequence Diagram

## **LIST OF TABLES**

Table 1: Goal and Objectives.

Table 2: Decision Description

Table 3: Stakeholders

Table 4: Project Resources

Table 5: Assumptions

Table 6: User and Roles

Table 7: User Stories

Table 8 - 10: Sprints 1 - 3

## **LIST OF FIGURES**

Figure 1: Use-case diagram

Figure 2-3: Admin and End-User Use-case Diagram

Figure 4: Class Diagram

Figure 5-6: Admin and End-User Activity Diagram

Figure 7: Sequence Diagram

Figure 8: Data Architecture

## 1. INTRODUCTION:

Data Leakage/Loss is one of the major issues that we face on a regular basis. We transfer data on a regular basis from one user to the other or from one entity to the other, but there is always a threat that the data might go into the wrong hands. As today's world is filled with loop-holes that anyone can misuses.

Data Leakage Investigation as it implies is, to detect the point at which there is loss of data and to provide a strong proof that it's being lost from that point and try to find a patch for it. This service is built to monitor a specific website or a server for the main aspect.

### 1.1. Problem Statement

In the course of doing business, we sometimes handover sensitive data to trusted third parties. While the sharing is in process, there is a strong possibility of the data being transferred, leaked. Through this project our team aims to minimize the said leakage.

## 2. Background Research

The main idea of the project came up after the incident of **"GOD OF BENNETT"**, who used to change the attendance of the students. This is due to some weak points in the server, which can be accessed by unauthorized personnel. These are the main resources from which we worked on the project.

<https://github.com/ronakm/Data-Leakage-Detection>

<https://www.ijser.org/researchpaper/Data-Leakage-and-Detection-of-Guilty-Agent.pdf>

### 2.1. Proposed System

This project aims to find the weak holes of a server or database, by analysing the data flow and using probability. We hope to catch anyone, who is using the data for unfruitful purposes and where the data is being lost.

We will also provide algorithms that do a rerun of the data tracks every time there is a leak detected in order to see where the data goes further. We make sure to provide watermarks, and also fake data that appears realistic to the leaker and try to track with these agents.

## 2.2. Goals and Objectives

Table 1: Goal and Objectives

#	Goal or Objective
1	Make the system User Friendly.
2	Data allocation strategies with high probability of leakage identification.
3	Detection of when the data has been leaked and by whom
4	Meet the changing needs and desires of clients and consumers.
5	Practice high ethical standards.
6	Provide better customer service.

## 3. Project Planning

### 3.1. Project Setup

#	Decision Description
1	Web Based Experience (Progressive Web Apps) Vs. Traditional Apps
2	Code should be maintainable and scalable and Agile Dev. Model Strategies.
3	Getting Verified by Security Standards to ensure safety and maintain Trust.
4	The program should be platform independent and easy to deploy.

### 3.2. Stakeholders

Stakeholder	Role
Business Partners	Investors.
Customers	End Users, who secure their data with our service
Faculty	Instructor/Mentor
Dev Team	Development / Bug Fixing / Testing
Legal Team	To maintain and prevent any legal issues that arise, due to the data privacy policies.

### 3.3. Project Resources

Resource	Resource Description	Quantity
Database Server	A database server provided by the sponsoring company.	1
Capstone Team	Our team of students who will be the primary developers of the project.	4
Dr. Anurag Gauswami	The mentor who will be able to provide us with technical assistance.	1
Mac Workstation	Macbook for developing IOS/MacOS side app.	1
Windows Workstation	Windows machines for developing the windows app.	3

### 3.4. Assumptions of where there may be a data leakage

	<b>Assumption</b>
Security Illiterate	Employee with little or no knowledge of security.
Testers	Employees that connect a variety of devices to their PC's.
Bored Employees	Employees who use the company IT for other purposes.
Interns	The internship students, who have access to data that they may use for other purposes, or use it for open-source projects.

## 4. SYSTEM ANALYSIS AND DESIGN

### 4.1. Overall Description.

Data leakage is an uncontrolled or unauthorized transmission of classified information to the outside. It poses a serious problem to companies. It is essential to discover data leakage as soon as possible, as it poses a serious threat to companies on the scale of going bankrupt. The purpose of this PROJECT is to design and implement a data leakage detection system based on special information retrieval models and methods, and also try removing the dark data. Data is to be stored in .pdf or as an image file.

### 4.2. Users and Roles

User	Description
Developer	A seasoned developer who is tasked with initial front-end, back-end and ultimately generating a firm process for applying these techniques to future server data.
System Admin	A developer who is tasked with managing system working and essential management roles such as viewing auto-generated reports, system logs and user queries.
Security Admin	A developer tasked with the job to (1) explore new security systems (2) discover flaws in old models and patch them out (3) view system logs to check if anything seems out of place (hacked/modified data)
Remote Developer	Considering the current world scenario, all of our team are working from home.
End Users	The end users, who use the data. We keep a track on the data that pertains to the company and used by the end users.



#### 4.3. User Stories

ID	Feature name	Story points
1	Register Systems	6
2	Systems Logging and Reports	6
3	User Data Management	3
4	Data Encryption Systems	15
5	User Query System (Feedback and Helpline)	3
6	Front-End Development (Web page + WebApp)	3
	<b>TOTAL</b>	36

## SPRINT 1

Estimated User Story Points: 5

Actual Completed User Story Points: 7

ID	Added	Description	Status	Story Points	Actual Equivalent Story Points	% Completed
100	Onset	<i>As a End User,</i>  <i>I want to be able to register online,</i>  <i>So that I can register quickly and access the service</i>	C	2	3	100%
101	Onset	<i>As a End User,</i>  <i>I want to be able to log-in online,</i>  <i>So that I can quickly access the service if i am already registered</i>	C	2	3	100%
102	Onset	<i>As a System Admin,</i>  <i>I want to be able to log-in as a admin</i>  <i>So that I can keep track of existing users and new users.</i>	C	1	1	100%

Acceptance Criteria		Verification
110	A user cannot submit a form without completing all the mandatory fields	Create a test case to verify non-empty fields.
111	Information from the form shall be stored in the registration database after form submission	Create a test case to verify information is stored in the database.
112	Payment shall be accepted via credit card	Create a test case to verify the credit card payment method from the bank.
113	An acknowledgment email shall be sent to the user after submitting the form.	Create test cases to verify sending of acknowledgement email after successful payment.

ID	Tasks	Resource
1	<i>Create a registration page with all required fields (FName, LName, organization, Address details, email, credit card details, Username, Password) and register button at the bottom, also have the same for Admin</i>	<b>Team member 1</b>
2	<i>Create a registration page with all required fields (Username and Password) and Login button at the bottom.</i>	<b>Team member 2</b>
3	<i>Develop a backend functionality that checks required fields are non-empty when the user clicks on register button.</i>	<b>Team member 3</b>
4	<i>Built a functionality which verifies payment from bank based on credit card details at 1.</i>	<b>Team member 4</b>
5	<i>Built a functionality which stores registration data in the database according to the specifications and sends acknowledgement email to the registered email else display payment failure message.</i>	<b>Team member 4</b>

## SPRINT 2

**Estimated User Story Points: 10**

**Actual Completed User Story Points: 14**

ID	Added	Description	Status	Story Points	Actual Equivalent Story Points	% Completed
200	Onset	<b>As a End User,</b>  <i>I want to be able to download the app from the website</i>  <b>So that I can launch the app locally</b>	<b>C</b>	<b>2</b>	<b>4</b>	<b>100%</b>
201	Onset	<b>As a End User,</b>  <i>I want to be able to query from the website</i>  <b>So that I can ask any query that I may face regarding the usage of the software</b>	<b>C</b>	<b>2</b>	<b>2</b>	<b>100%</b>
202	Onset	<b>As a End User,</b>  <i>I want to use the service</i>  <b>So that I can encrypt my data</b>	<b>C</b>	<b>4</b>	<b>6</b>	<b>100%</b>

203	Onset	<p><b>As a System Admin,</b></p> <p><i><b>I want to be able to view who is registered</b></i></p> <p><b>So that I can keep track of existing users and new users.</b></p>	C	2	2	100%
Acceptance Criteria			Verification			
210	Website successfully identifies the client platform and prepares application download.		Create Test cases that verify downloads for each Major OS platform.			
211	A Query section exists for customers to successfully interact with Support Staff.		Create a Usability test for the Query Section on the Website.			
212	Customer successfully launches the app and is able to Encrypt/Secure their data connection.		<p>Usability test to determine if customers can use the platform successfully.</p> <p>Create Test Cases for major bottleneck scenarios for encryption of data.</p>			
213	System successfully generated a log of all users registered by the form.		Create test cases for creating system logs in for all types of registrations.			

ID	Tasks	Resource
1	Developed a system which automatically detects the host OS platform and prepares the download for them, with a click of the Download Button.	Team member 1
2	Developed the Back-End for the Query System.	Team member 2
3	Integrated and Developed the UI for the Query System in accordance to modern UI/UX standards.	Team member 2
4	Developed the essentials of the Encryption System that secured the User Data Connection.	Team member 3
5	Developed a system which logs Error/Lost data information and registered User data.	Team member 4

### SPRINT 3

**Estimated User Story Points: 15**

**Actual Completed User Story Points: 15**

**Main User Story: SECURING USER DATA AND OUR SERVERS**

***“As a End-User, I want to use this service, so that I can secure my data connection.”***

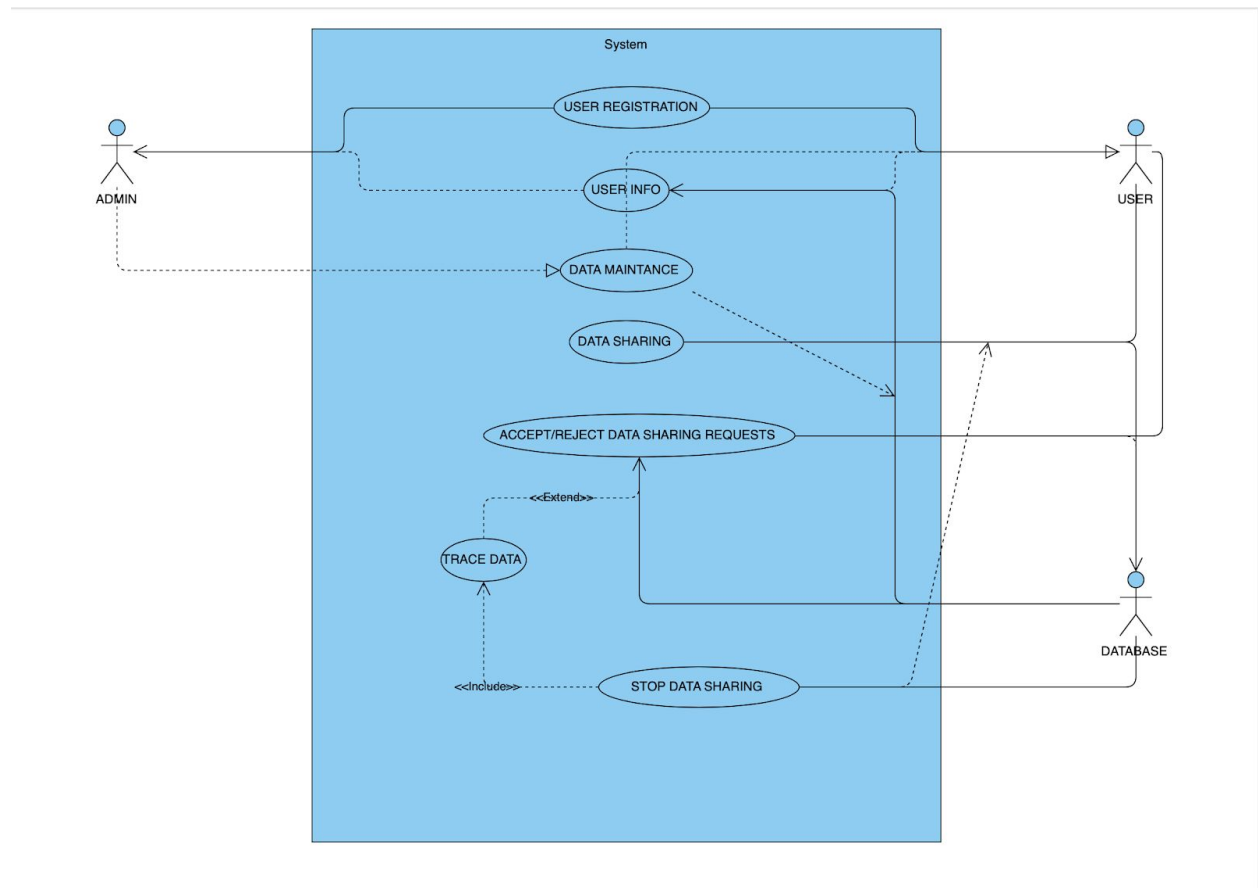
ID	Added	Description	Status	Story Points	Actual Equivalent Story Points	% Completed
300	Onset	<i>As a End User, I want to use this service,  So that I can secure my data connection.</i>	C	15	15	100
Acceptance Criteria			Verification			
310	User Downstream Data must be encrypted from our servers.		Create test cases to verify search results by quiz name.			
311	User Upstream Data must be encrypted from our servers.		Create test cases to verify search results by quiz topics.			
312	Our servers must have a Firewall to protect them from Cyber Attacks.		Create test cases to verify search results by creation and last used date.			



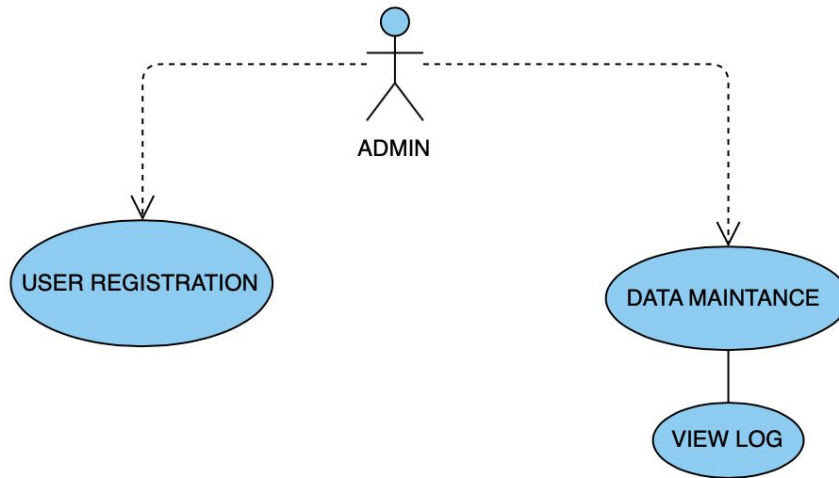
ID	Tasks	Resource
1	User Downstream Data Encryption setup (AES-256)	Team member 1
2	User Upstream Encryption setup (AES-256)	Team member 2
3	Incorporation of a third-party security vendor.	Team member 3

## 4.4. Design diagrams/ UML diagrams/ Flow Charts/ E-R diagrams

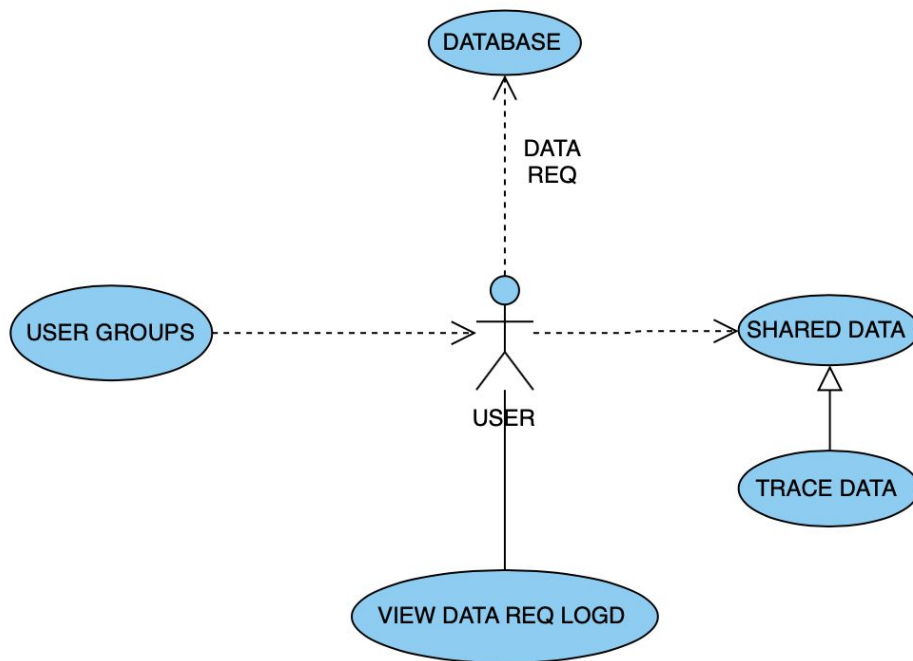
### 4.4.1. Use Case Diagrams



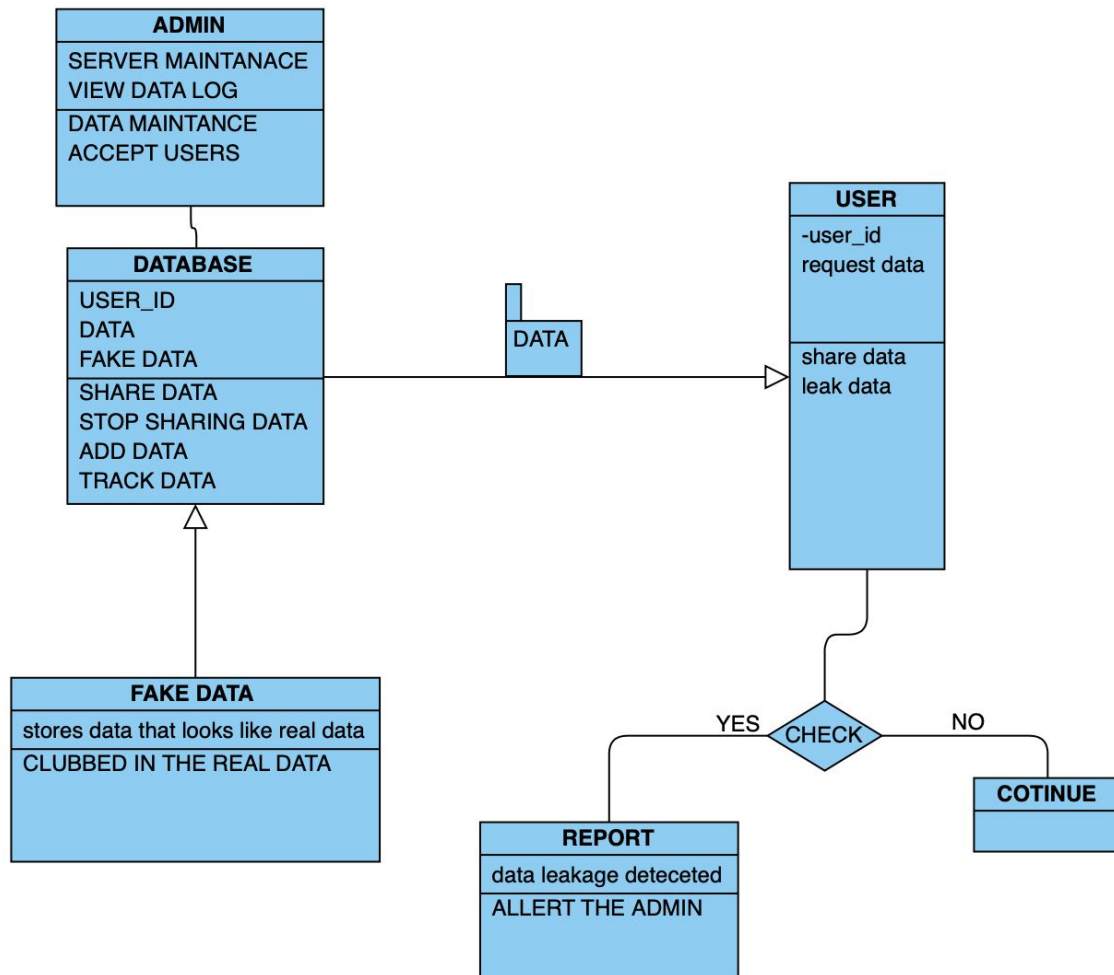
Admin use case diagram



USER USE CASE DIAGRAM



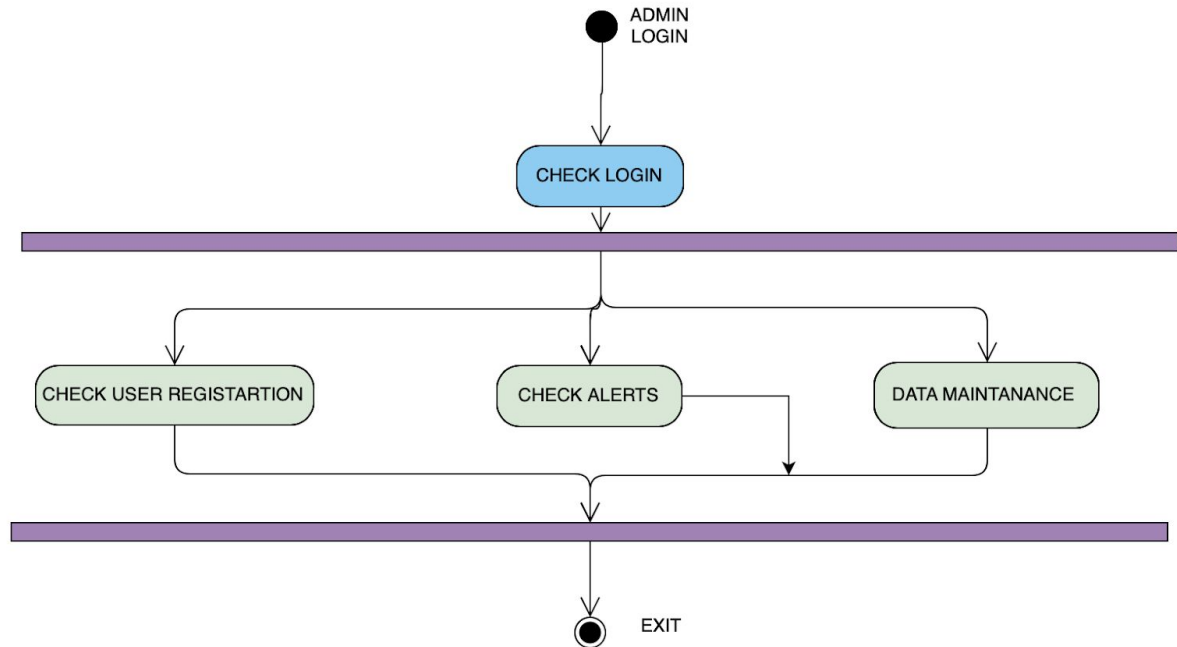
#### 4.4.2. Class Diagram



CHECK = the probability calculation of the data leakage and which user.

#### 4.4.3. Activity Diagrams

##### ADMIN ACTIVITY DIAGRAM



##### USER ACTIVITY DIAGRAM

