

# **UNIFIED DLP SOLUTIONS FOR EMAIL SYSTEM**

**Project Id: TMP-2023-24-082**

Status Document - 01

Keshani S K - IT20299002

B.Sc. (Hons) Degree in Information Technology

Specialization in Cyber Security

Department of Information Technology

Sri Lanka Institute of Information Technology

Sri Lanka

December 2023

## Table of Contents

1. Microsoft Teams Details .....	3
1.1 Team Members .....	3
1.2 Team Meeting with Group Members .....	3
1.3 Team meeting with Supervisor.....	5
2. MS Teams Planner .....	6
2.2 Planner Board .....	6
2.2 Team Planner Charts.....	7
3. WhatsApp chat with Supervisor .....	8
4. WhatsApp Call and Chat with Team.....	10
5. Email and planned Teams Events .....	11
6. Gantt Chart .....	13
7. Work Breakdown Structure (WBS) .....	14
8. Gitlab Commits .....	15

# 1. Microsoft Teams Details

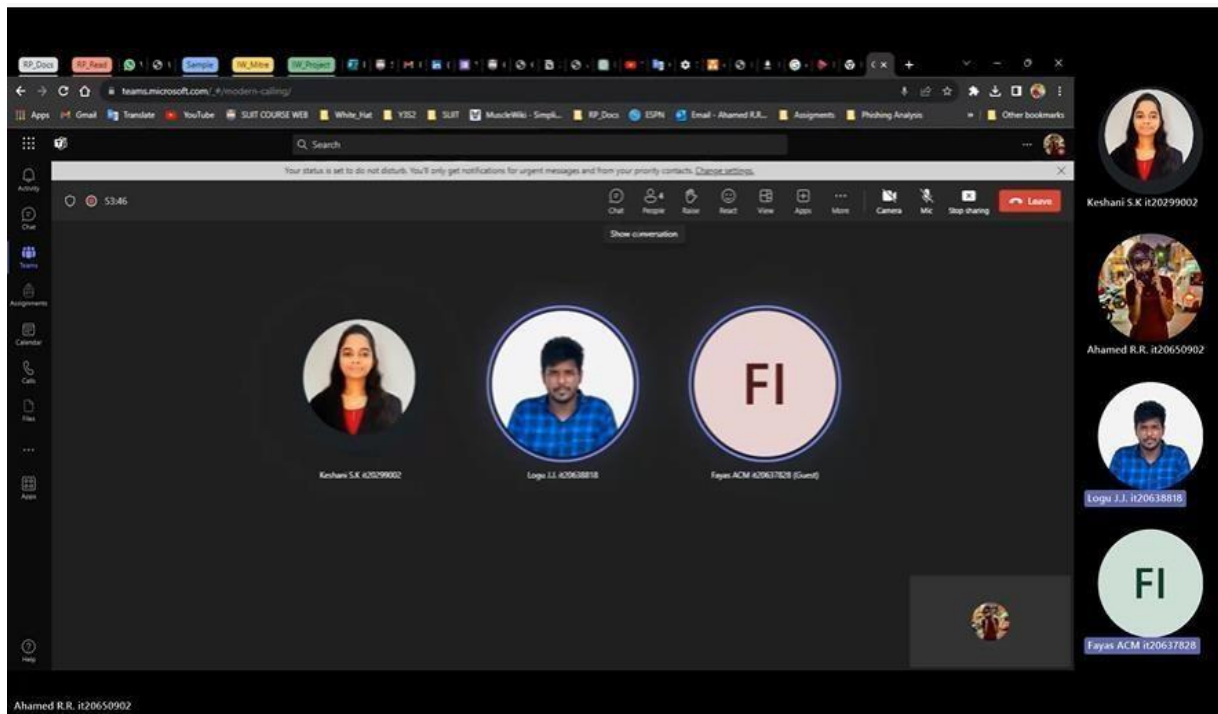
## 1.1 Team Members

The screenshot shows the 'Members' tab of a Microsoft Teams team. The team name is 'TMP-2023-24-082\_ Unified DLP Solutions for Email ...'. The 'Owners (2)' section lists Keshani S.K. (it20299002) as a Leader and Amila Nuwan as a Senior Lecturer. The 'Members and guests (3)' section lists Logu J.J. (it20638818), Ahamed R.R. (it206509...), and Fayas A.C.M. (it20637828). The interface includes a search bar, an 'Add member' button, and a table of members with columns for Name, Title, Location, Tags, and Role.

Name	Title	Location	Tags	Role
Keshani S.K. it20299002			Leader	Owner
Amila Nuwan	Senior Lecturer	Malabe		Owner
Logu J.J. it20638818				Member
Ahamed R.R. it206509...				Member
Fayas A.C.M. it20637828				Member

## 1.2 Team meeting with Group Members

The screenshot shows the 'Details' tab of a Microsoft Teams meeting titled 'PP1 Discussion with Supervisor'. The meeting is scheduled for 11/2/2023 from 8:00 PM to 10:00 PM. The meeting options include 'Show as: Busy', 'Category: None', and 'Time zone: (UTC+05:30) Sri Jayawardenepura'. The meeting is currently in a 'Tentative' state. The 'Tracking' section on the right shows the status of participants: Keshani S.K. (Organizer), Fayas A.C.M. (Accepted), Ahamed R.R. (Unknown), Logu J.J. (Unknown), and Amila Nuwan (Unknown). The meeting description includes the text: 'Please join all at 8pm today to discuss the PP1. Thank you..'



**Roohan Ahamed (Presenting)**

File Home Insert Draw Design Layout References Mailings Review View Help

Clipboard Font Paragraph Styles Editing Comments

**System Overview:**

```

graph TD
    Dataset[dataset] --> CreateCNN[Create multiple CNN models]
    CreateCNN --> FineTune[Fine tune each CNN model]
    FineTune --> TrainEnsemble[Train the ensemble model]
    TrainEnsemble --> ComputeAnomaly[Compute the anomaly predefined threshold to classify]
    ComputeAnomaly --> InputDesc[Input the description]
    InputDesc --> PredictLabel[Predict the label of the new document]
    PredictLabel --> DefineDoc[Define document for prediction]
    DefineDoc --> TrainLogit[Train the logistic regression model]
    TrainLogit --> LoadDataset[Load dataset]
    LoadDataset --> TrainLogit
    LoadDataset --> TrainRF[Train the Random Forest on]
    TrainRF --> CreateRF[Create a Random forest classifier]
    CreateRF --> PredictLabel
    
```

Roohan Ahamed Jethushan Logu

Muhammed Fayas kokila keshani

### 1.3 Team Meeting with Supervisor

The screenshot shows a Microsoft Teams meeting in progress. The main window displays four participants in a grid: Logu J.J. (top left), Fayas Abdul (top center, with a pink circle containing 'FA' over his face), Ahamed R.R. (top right), and Amila Nuwan (bottom center). Below the grid is a small video thumbnail of a fifth participant. The right sidebar shows the 'Participants' list with five people: Keshani S.K. (Organizer), Ahamed R.R., Amila Nuwan, Fayas Abdul (External), and Logu J.J. The meeting controls at the bottom are visible.

The screenshot shows a Microsoft Teams meeting where a presentation is being shared. The presentation is a document titled 'Proposal\_Report.docx' and is displayed in a split-screen view. The left pane shows a 'System Diagram' with a flowchart. The right pane shows '4. PROJECT REQUIREMENTS' with sub-sections: 4.1 Functional Requirements, 4.2 Non-Functional Requirements, and 4.3 Software Requirements. The right sidebar shows the 'Participants' list with five people: Keshani S.K. (Organizer), Ahamed R.R., Amila Nuwan, Fayas A.C.M., and Logu J.J. The meeting controls at the bottom are visible.

## 2. MS Teams Planner

### 2.2 Planner Board

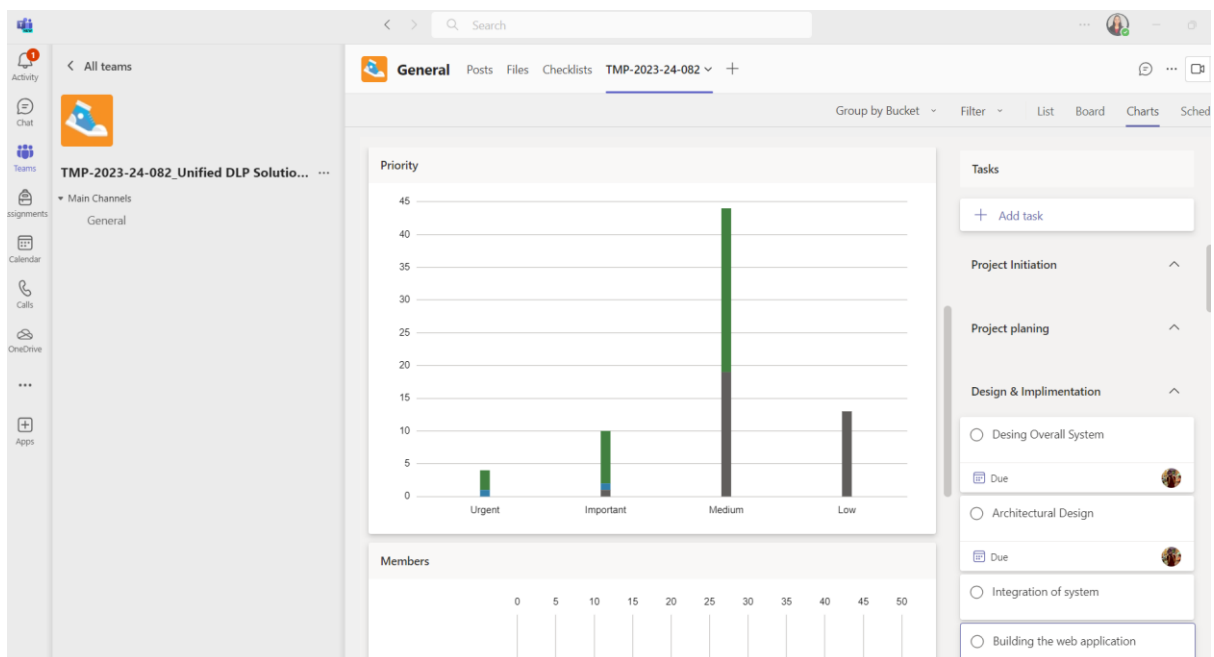
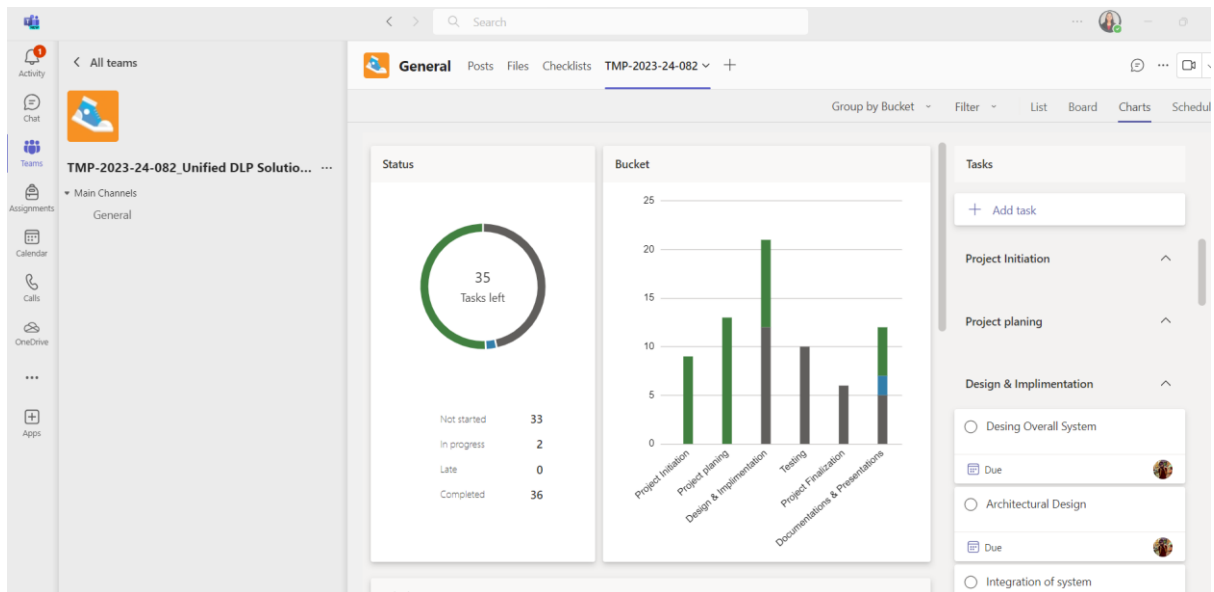
#### Completed Tasks

The top screenshot displays a list of tasks in the MS Teams Planner interface. The tasks are organized into a table with columns for Task title, Completed by, Priority, Completed date, and Bucket. The tasks are as follows:

Task title	Completed by	Priority	Completed	Bucket
Topic-Assessment	Ahamed R.R. it2065...		12/2023	Project Initiation
Proposal Presentation	Ahamed R.R. it2065...		12/2023	Project Initiation
Project proposal report	Ahamed R.R. it2065...		12/2023	Project Initiation
identifying research problems & research gaps within those components	Ahamed R.R. it2065...		12/2023	Project planning
Performing a background study about the topic	Fayas A.C.M. it2063...		11/2023	Project Initiation
Evaluation	Fayas A.C.M. it2063...		11/2023	Project Initiation
Designing an overall system Diagram	Fayas A.C.M. it2063...		11/2023	Project planning
creating a gannt chart	Fayas A.C.M. it2063...		11/2023	Project planning
assign classification labels	Keshani S.K it20299...		11/2023	Design & Implim...
creating text classification model	Keshani S.K it20299...		11/2023	Design & Implim...
text pre-processing	Keshani S.K it20299...		11/2023	Design & Implim...
Training the models	Keshani S.K it20299...		11/2023	Design & Implim...
Project charter	Keshani S.K it20299...		11/2023	Project Initiation
Identify individual research component	Keshani S.K it20299...		11/2023	Project planning
perform a background research about each research component	Keshani S.K it20299...		11/2023	Project planning
deciding the technologies to be used	Keshani S.K it20299...		11/2023	Project planning

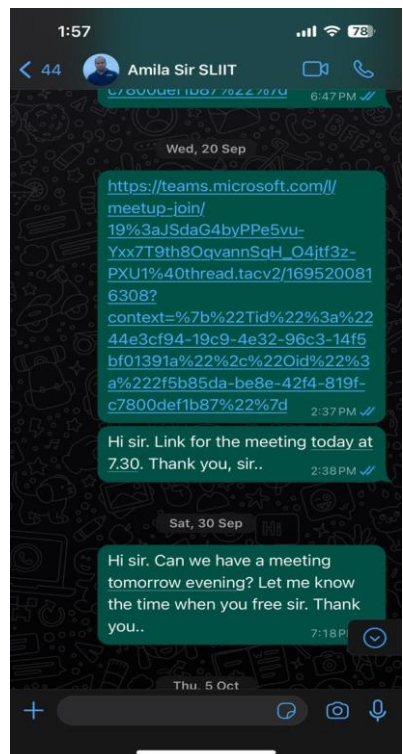
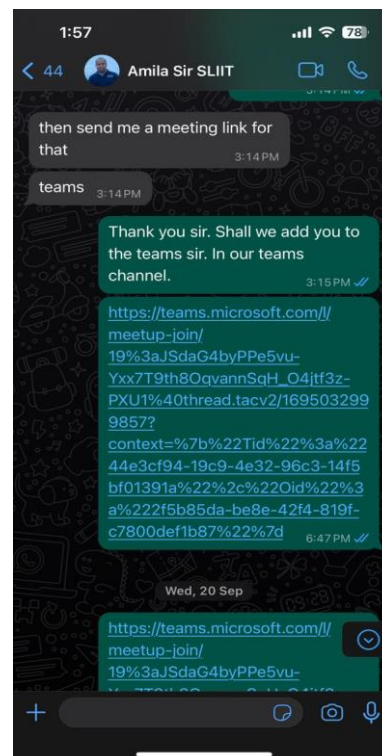
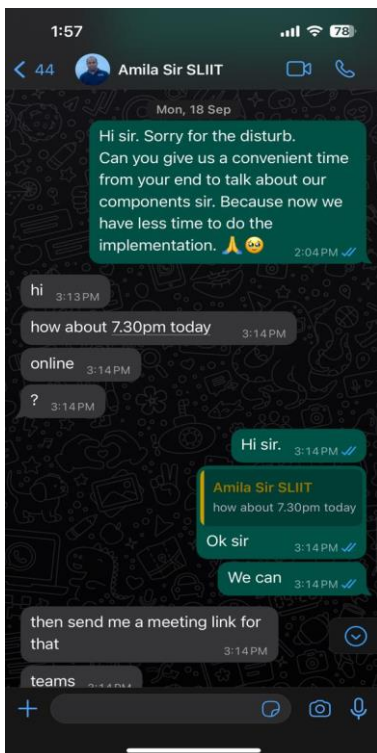
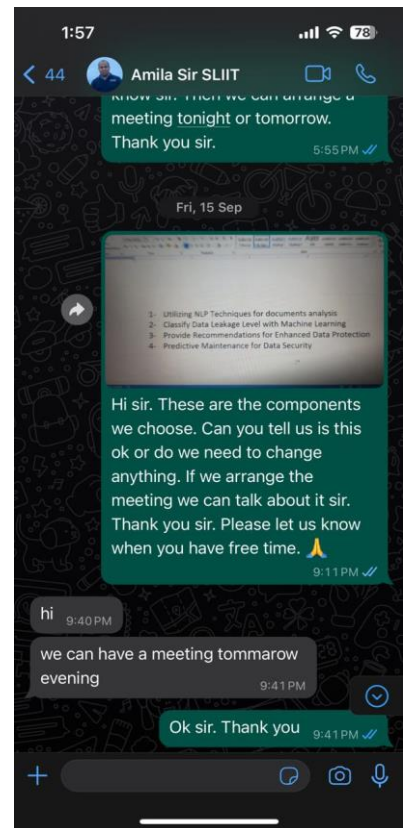
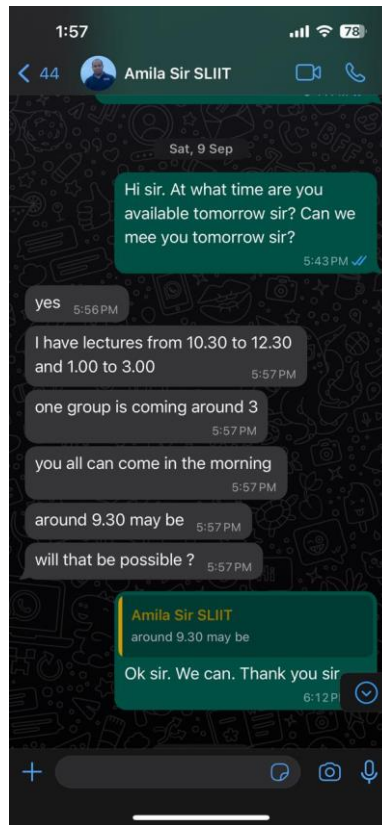
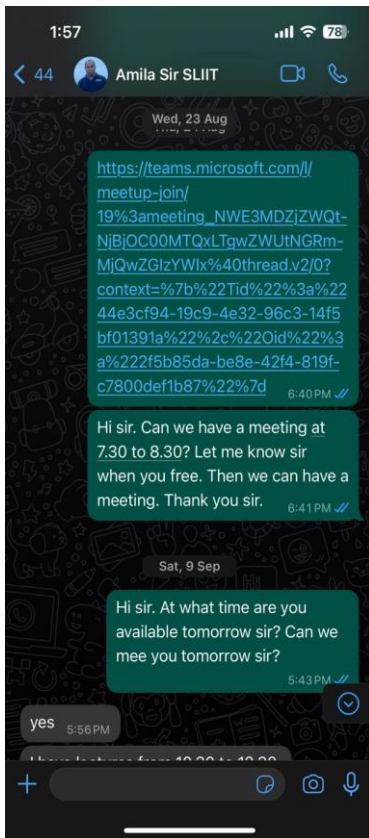
The bottom screenshot shows the same tasks in a 'Board' view, grouped by bucket. The buckets are Project Initiation, Project planning, Design & Implimentation, and Testing. The tasks are displayed as cards within each bucket, showing the task title, completed by, and due date.

## 2.2 Team Planner Charts

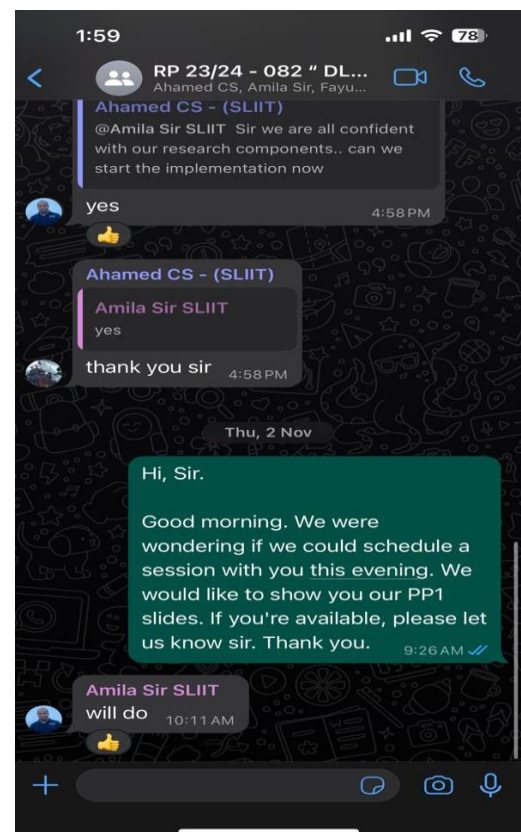
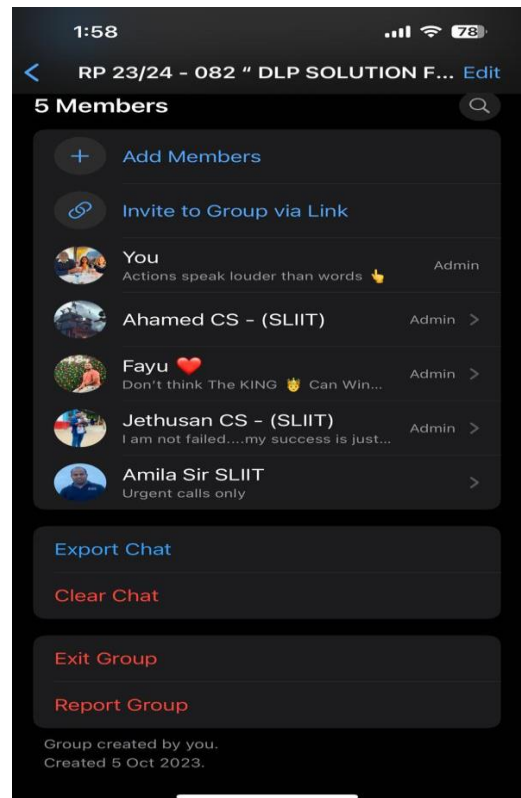
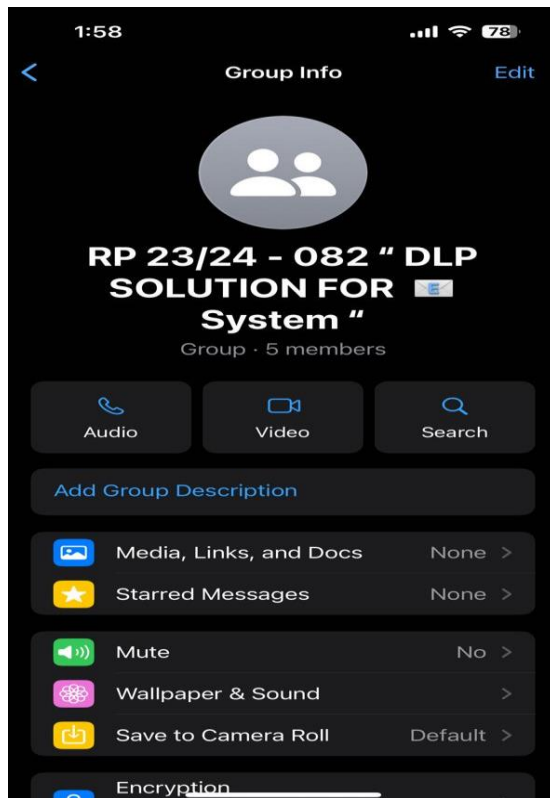




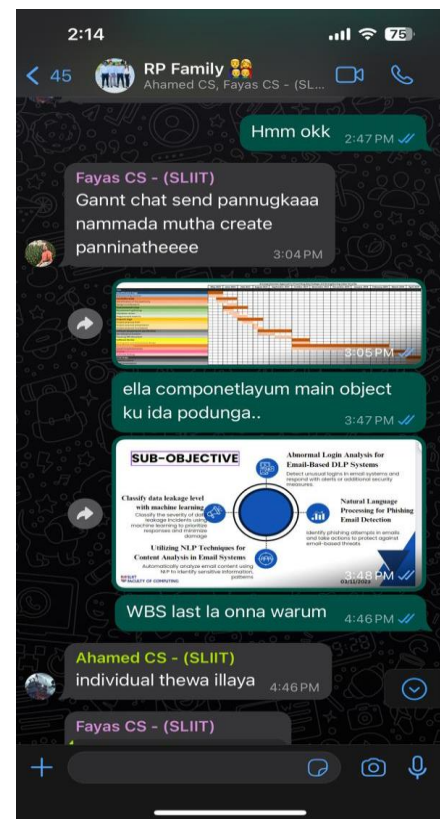
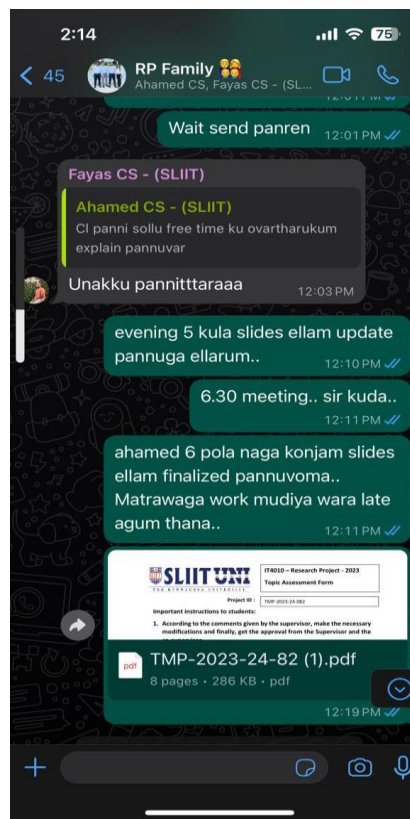
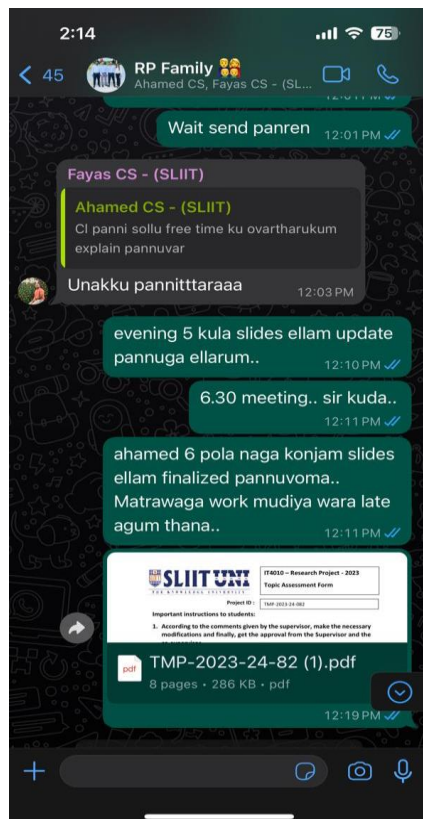
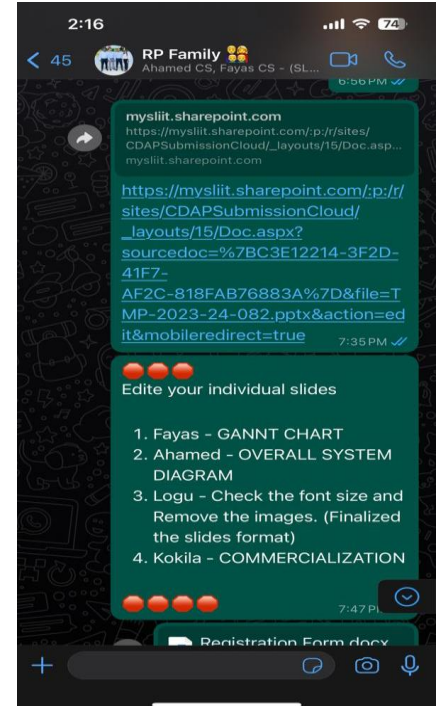
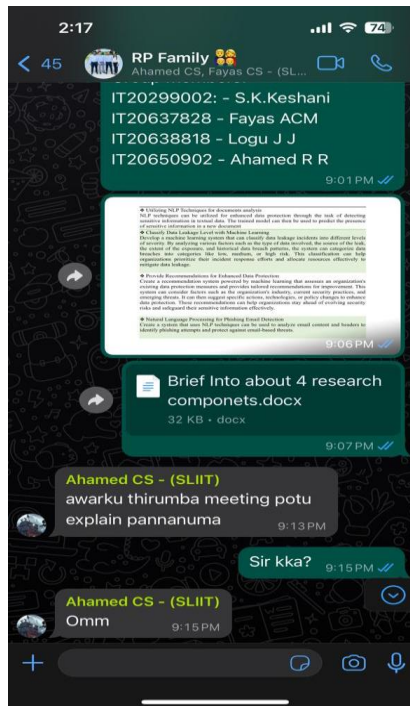
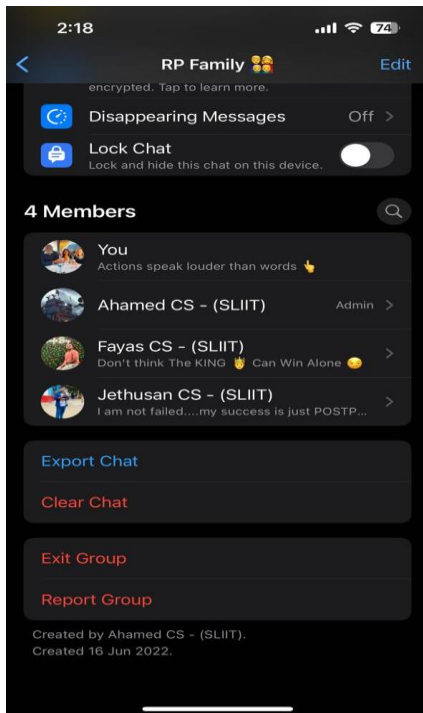
### 3. WhatsApp chat with Supervisor







## 4. WhatsApp Call and Chat with Team





## 5. Email and planned Teams Events

The screenshot shows a Microsoft Teams meeting invitation titled "Final Project Discussion". It indicates the event occurred 2 months ago on Thursday, 21/09/2023, from 21:00 to 23:00. The meeting is a Microsoft Teams Meeting. The organizer is Keshani S.K. (it20299002) on behalf of STRENGTHENING CYBERSECURITY <RP\_Fan>. A note says "Didn't respond 4". The interface has tabs for "Messages" and "Meeting Details". Under "Meeting Details", there is a warning: "[EXTERNAL EMAIL] This email has been received from an external source – please review before actioning, clicking on links, or opening attachments." Below this, it says "Please join all at 9 pm." and "Thank you..". At the bottom, it says "Microsoft Teams meeting". On the right, under "Organiser", it lists the meeting title "TMP-2023-24-082\_A COMPRE...", the date "Sent on Thursday, 21/09/2023 at 20:39", and a list of required attendees: Keshani S.K. (it20299002), Ahamed R.R. (it20650902), Fayas A.C.M. (it20637828), and Logu J.J. (it20638818).

Fig: Planned MS Teams Meeting

The screenshot shows an email from Keshani S.K. (it20299002) to Amila Nurwan. The email is dated Tuesday, 08/10/2023, at 22:13. The subject is "Brief Into about 4 research c...". The body of the email contains a warning: "[EXTERNAL EMAIL] This email has been received from an external source – please review before actioning, clicking on links, or opening attachments." Below this, it says "Dear Sir," and "I hope this message finds you well. I wanted to update you on our progress regarding the research project. We have initiated the implementation of the following four components:" followed by a list of four components: 01. Abnormal Login Analysis for Email-Based DLP Systems, 02. Natural Language Processing for Phishing Email Detection, 03. Utilizing NLP techniques for Content Analysis, and 04. Classify Data Leakage Level with Machine Learning. The email concludes with: "I've attached a detailed document outlining our approach, methodology, and progress on each of these components for your review. If you have any concerns or feedback regarding these components or if you believe there are additional aspects we should consider, please do not hesitate to share your insights. Your guidance is invaluable to us, and we aim to ensure the system we develop is both comprehensive and effective."

Fig: Email Communication with Supervisor

PP1 Discussion with Supervisor

Follow up

This event occurred 1 month ago (Thu 02/11/2023 20:00 - 22:00)

Microsoft Teams Meeting;  
Teams meeting

Keshani S.K it20299002 <it20299002@my.sliit.lk> on behalf of IMP-2023-24-082\_Unified DLP Solutions fo  
Accepted 1, Didn't respond 4

Messages

Meeting Details

Keshani S.K it20299002 <it20299002@my.sliit.lk> on behalf of  
TMP-2023-24-082\_Unified DLP Solutions for Email System <RP\_Family@sliit.lk> updated the meeting time  
Thu 02/11/2023 20:00 - 22:00 Thu 02/11/2023 19:00 - 21:00

T

Keshani S.K it20299002 <it20299002@my.sliit.lk> on behalf of TMP-2023-24-082\_Ur

To: Keshani S.K it20299002; Ahamed R.R. it20650902; Logu J.J. it20638818; Amila Nuwan; +2 others

Thu 02/11/2023 17:50

[EXTERNAL EMAIL] This email has been received from an external source – please review before actioning, clicking on links, or opening attachments.

Please join all at 8pm today to discuss the PP1.

Thank you..

Fig: PP1 Discussion Meeting with Supervisor

Research Project - Update the Power point Slides

Follow up

This event occurred 1 month ago (Wed 04/10/2023 19:00 - 20:00)

Microsoft Teams Meeting;  
teams meeting

TMP-2023-24-082\_A COMPREHENSIVE APP  
Keshani S.K it20299002 <it20299002@my.sliit.lk> on behalf of STRENGTHENING CYBERSECURITY <RP\_Fan  
Didn't respond 4

Messages

Meeting Details

[EXTERNAL EMAIL] This email has been received from an external source –  
please review before actioning, clicking on links, or opening attachments.

Please join at 7 pm.

Thank you..

Organiser

T

TMP-2023-24-082\_A COMPRE...

Sent on Wednesday, 04/10/2023 at 18:52

Keshani S.K it20299002  
Required

Fayaz A.C.M. it20637828  
Required

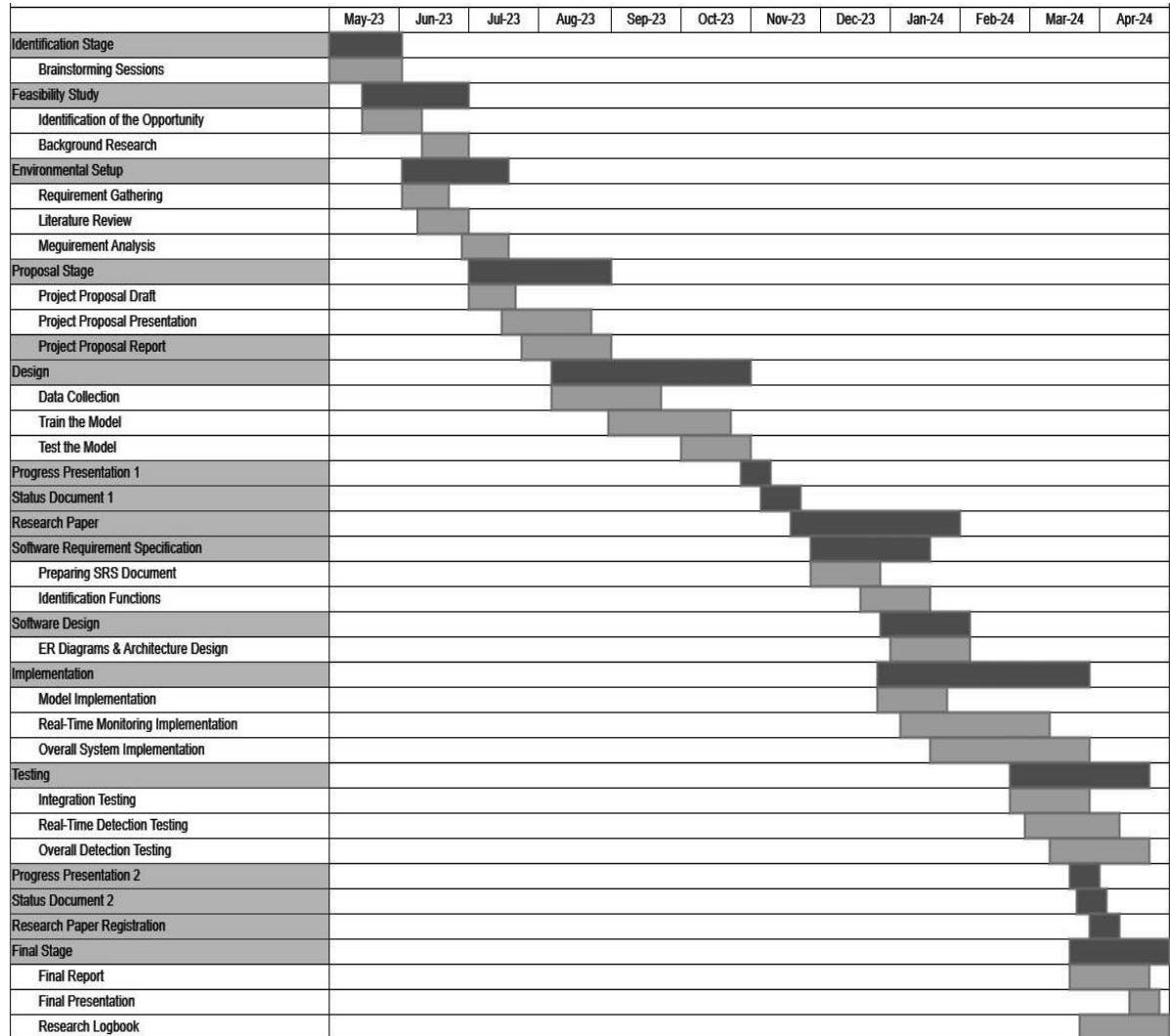
Ahamed R.R. it20650902  
Required

Logu J.J. it20638818  
Required

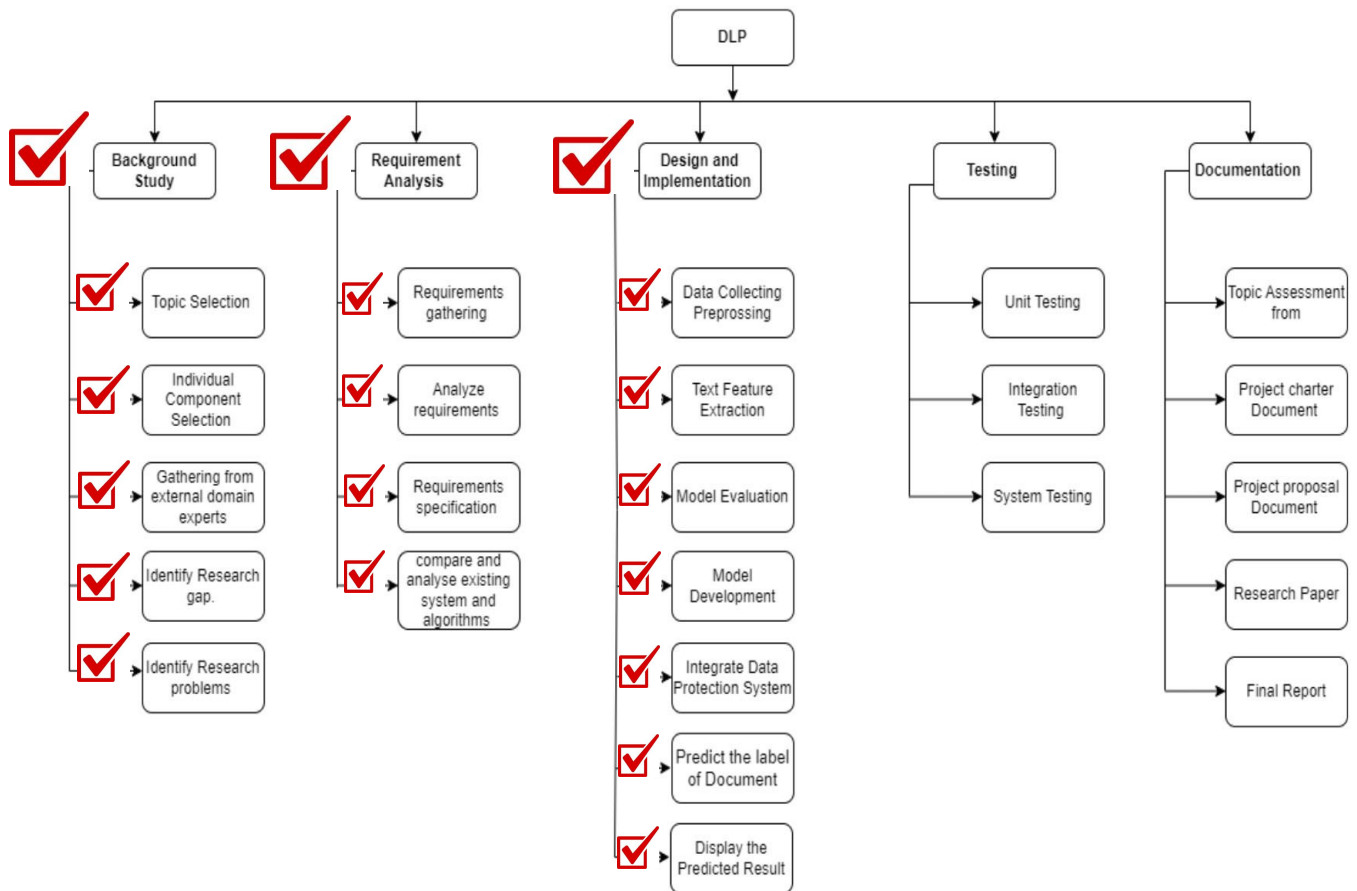
Microsoft Teams meeting

Fig: PP1 Slides Updating Meeting with Team Members

## 6. Gantt Chart



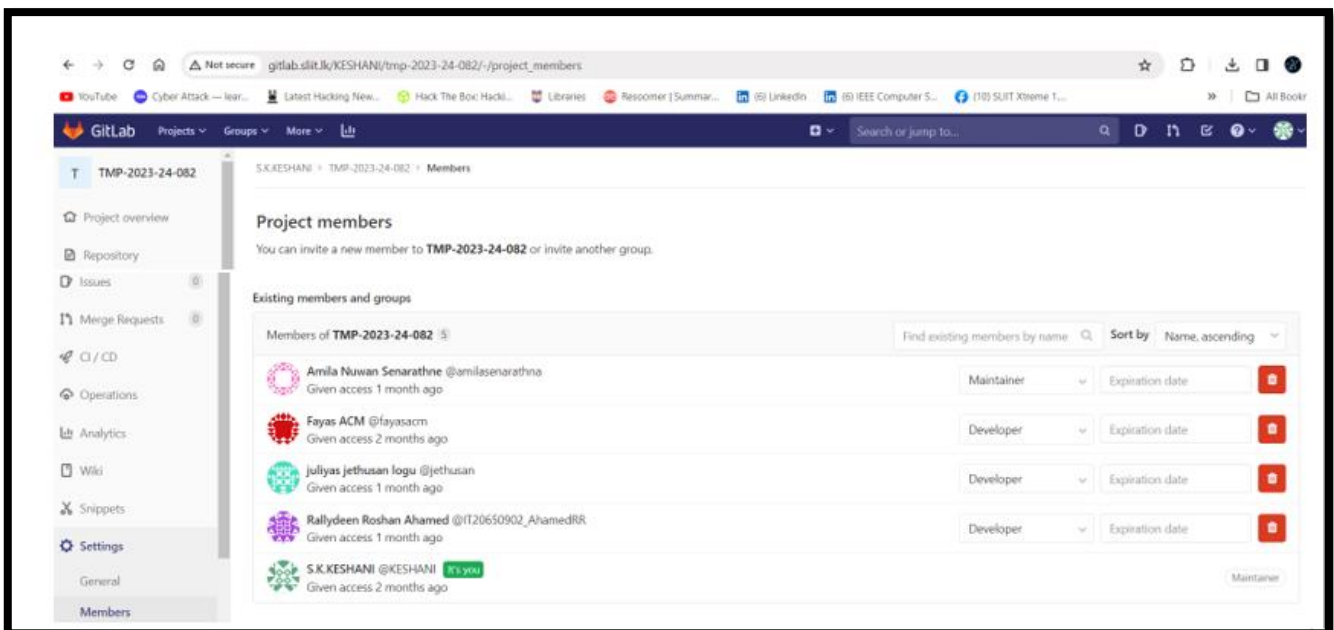
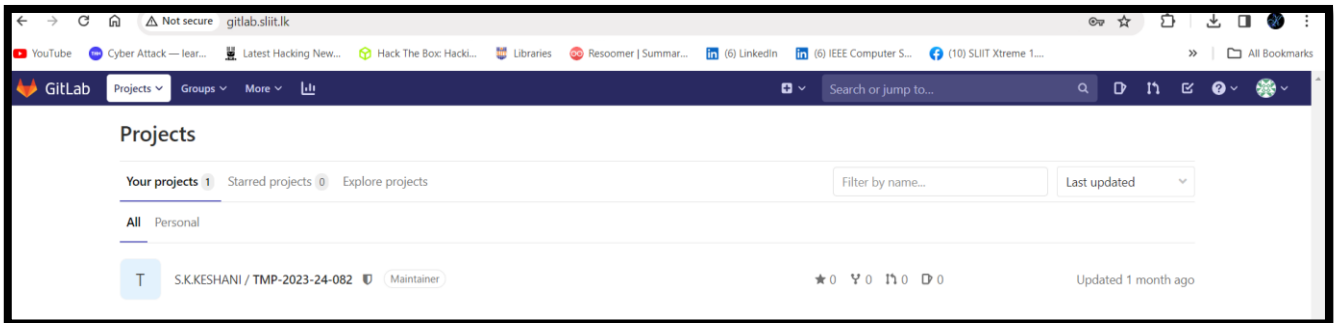
## 7. Work Breakdown Structure (WBS)



Completed


## 8. Gitlab Commits

Fig: Team Members and Supervisor





Your changes have been successfully committed.

 Update README.md  
S.K.KESHANI authored in 28 seconds

842b9fed

README.md 1.07 KB

</>

Edit

Web IDE

Replace

Delete

## TMP-2023-24-082

---

### Main Research Topic : Unified DLP for Email System

---

### Sub Research Component : Classify data leakage level with machine learning

---

### Induction and Background :-

---

1. Rank threats based on their severity.
2. Allocate resources efficiently to handle potential breaches.
3. Inform stakeholders about the gravity of a leakage incident.
4. Take proactive measures depending on the predicted severity of a leak.

---

### RESEARCH PROBLEM

---

1. Difficulty in identifying meaningful patterns and anomalies in large volumes of data.
2. Lack of effective tools and strategies for quick and accurate identification.
3. Need for minimizing the risk of data loss and reputational damage.

---

**Main Objectives:-** The core objective of classifying data leakage level using machine learning in a DLP system is to predict the potential impact of a data leak, enabling organizations to respond appropriately and minimize damage.

---

### Sub Objectives :-

---

1. Model Validation and Testing
2. Interactive Classification Dashboard
3. Real-time Alerting System
4. Data Masking and Auto-Redaction
5. Confidence Scoring