

Logbook – Attack Defense and resilience of R25 -039



Project ID: R25 - 039

**Project Title: Data-Privacy Focused Federated Learning
Framework for Industrial IoT**

Student Details:

Names:

Nanayakkara Y.D.T. D

Student IDs:

IT21826368

Supervisor: Mr. Amila Seneratha

Co-Supervisor: Mr. Tharaniyawarma Kumaralingam

Date of Submission: 2025

Contents

1. Group Details.....	3
2. Project Details.....	3
3. Communication Methods	4
4. Meetings With Supervisors	6
5. Task Details	7
6. System Details	8
6.1 System completion status.....	8
6.2 System Design	10
I. System Architecture.....	10
II. Module Architecture	11
6.3 System Testing.....	12
6.4 System Codes.....	13
7. GitHub Upload	14
8. Documentation.....	14
8.1 Proposal.....	14
8.2 Presentation 1	15
8.3 Presentation 2.....	15
8.4 Final Presentation.....	16
8.5 Final Product.....	16
8.6 Research Paper.....	17
III. Conference Appetence	17
9. CDAP upload.....	18
10. Website	19
10.1 Development	19
10.2 Finalize.....	22

1. Group Details

Student Details:

Names:	Student IDs:	Research Component
Nanayakkara Y.D.T. D	IT21826368	Attack Defense and Resilience
Mendis H.R.M	IT21822612	Privacy Preservation
Weerasinghe K.M	IT21831904	Secure Aggrigation
Dissanayaka K.D.A.R. A	IT21828348	Secure Communicaiton and Protocol Enforcement

2. Project Details

Topic - Data-Privacy Focused Federated Learning Framework for Industrial IoT

Aim – To develop a product that going to full fill the research

Deliverables – Federated Learning Framework designed for industrial internet of things

This project was initiated to develop a secure and private **Federated Learning (FL) framework** specifically for **Industrial IoT (IIoT)** environments.

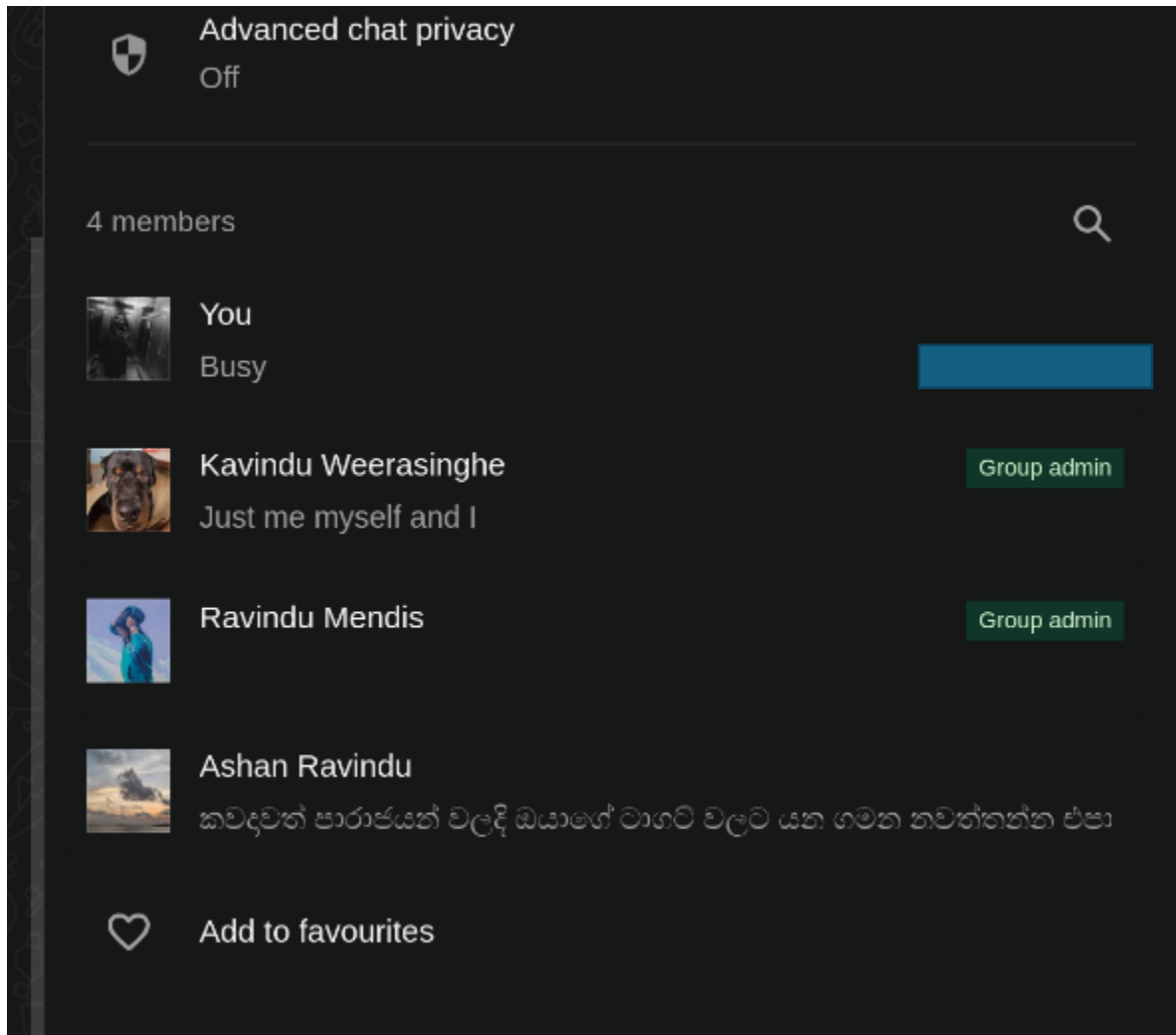
The Challenge: Traditional AI methods require centralizing sensitive factory data, which poses major **privacy risks** and clashes with the distributed nature of industrial operations. Existing FL solutions are insufficient because they fail to simultaneously provide robust security, data privacy, and efficient operation on **resource-limited IIoT devices**.

The Solution: The developed framework is a multi-layered system that provides **end-to-end protection**.

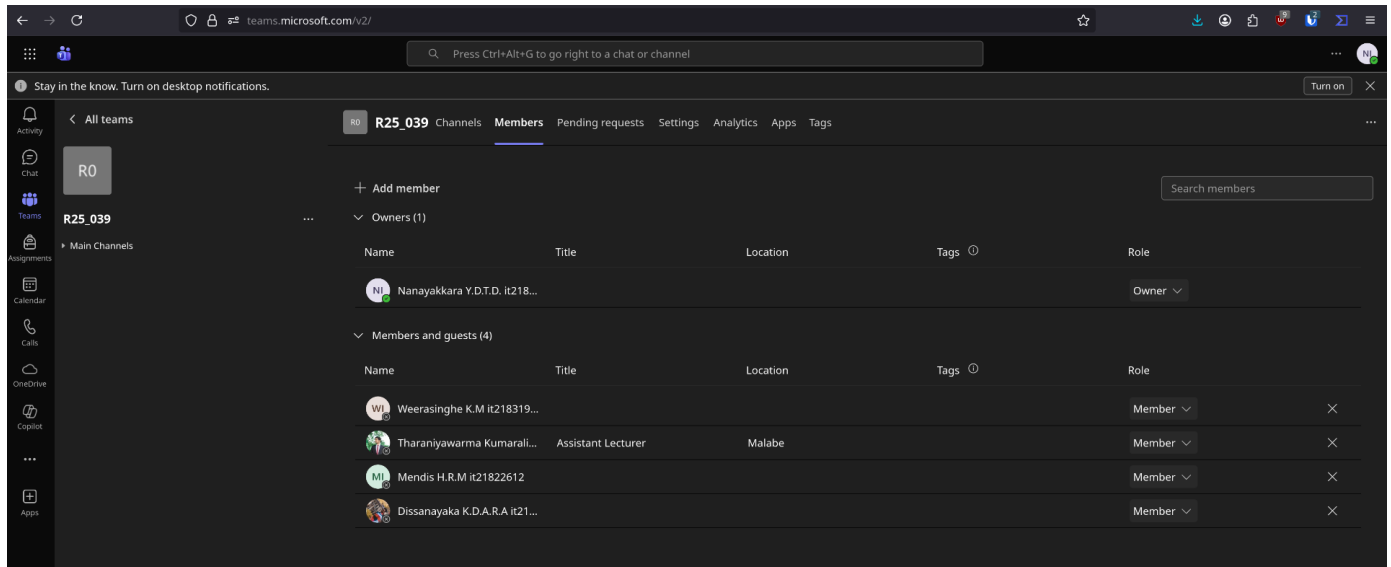
- It uses techniques like **Differential Privacy (DP)** and **Homomorphic Encryption (HE)** to guarantee data confidentiality.
- It implements a robust protocol that uses **client/server validation** to actively block cyber threats such as **Model Poisoning and Byzantine Attacks**.
- The system is optimized for **efficiency** to reduce overhead on IIoT devices.

3. Communication Methods

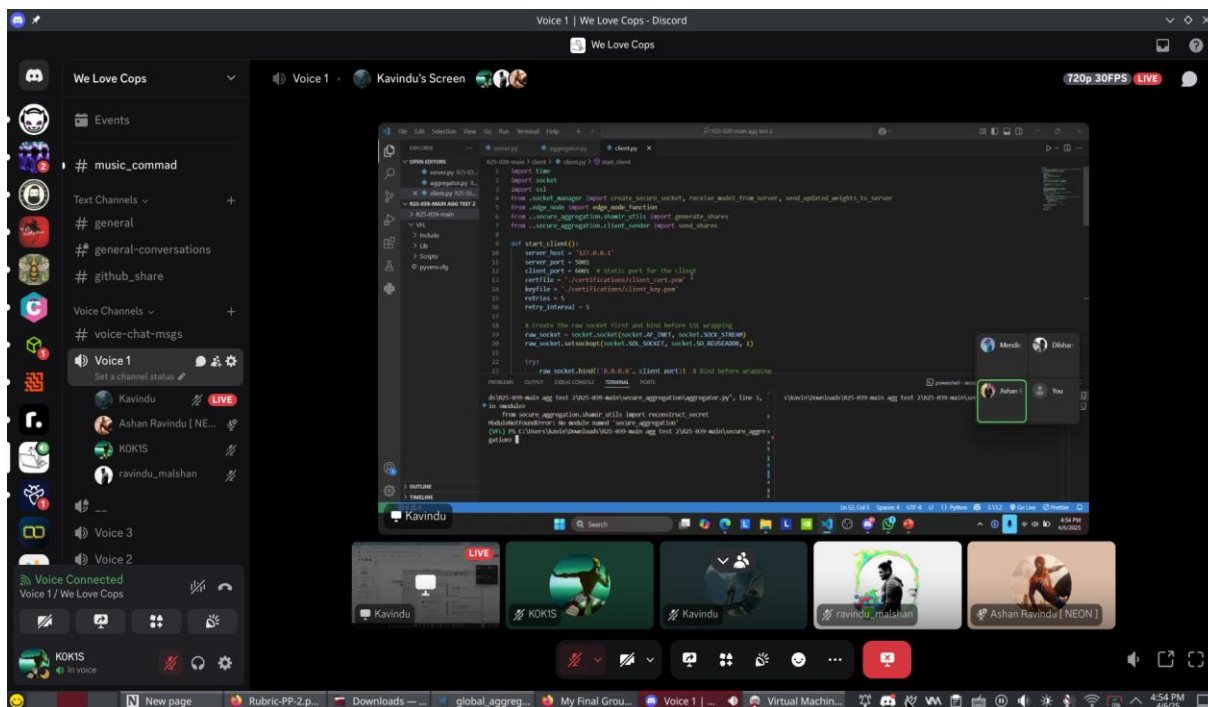
WhatsApp Group – Team



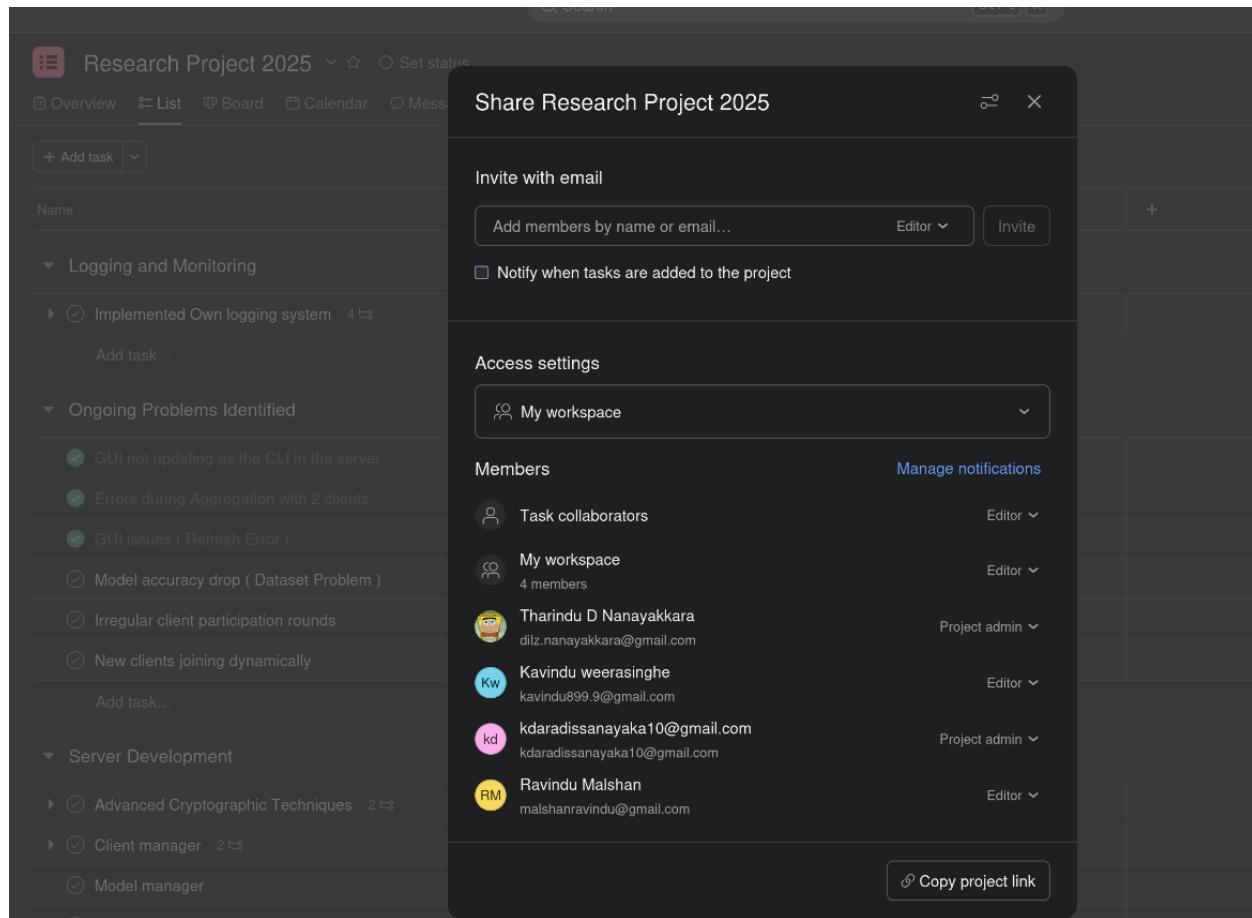
Microsoft Teams - All



Group Meetings – Discord – Team



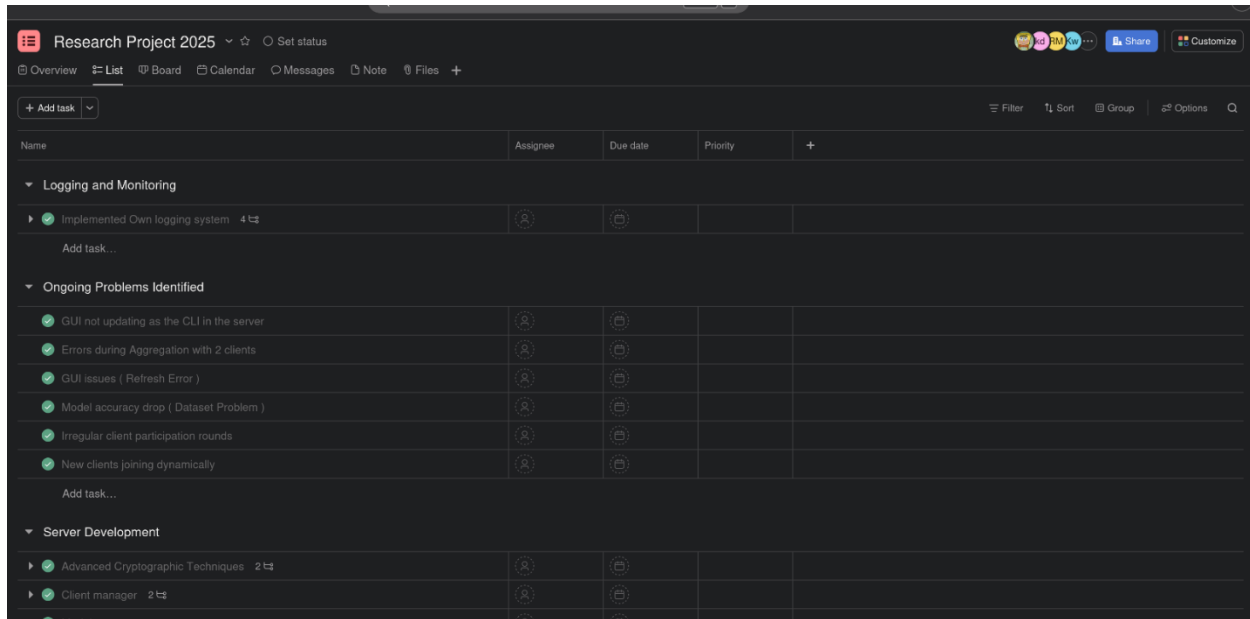
Asana – Task Assigning – Team



4. Meetings With Supervisors

All the meetings were conducted in person and only WhatsApp calls were taken to organize the meeting

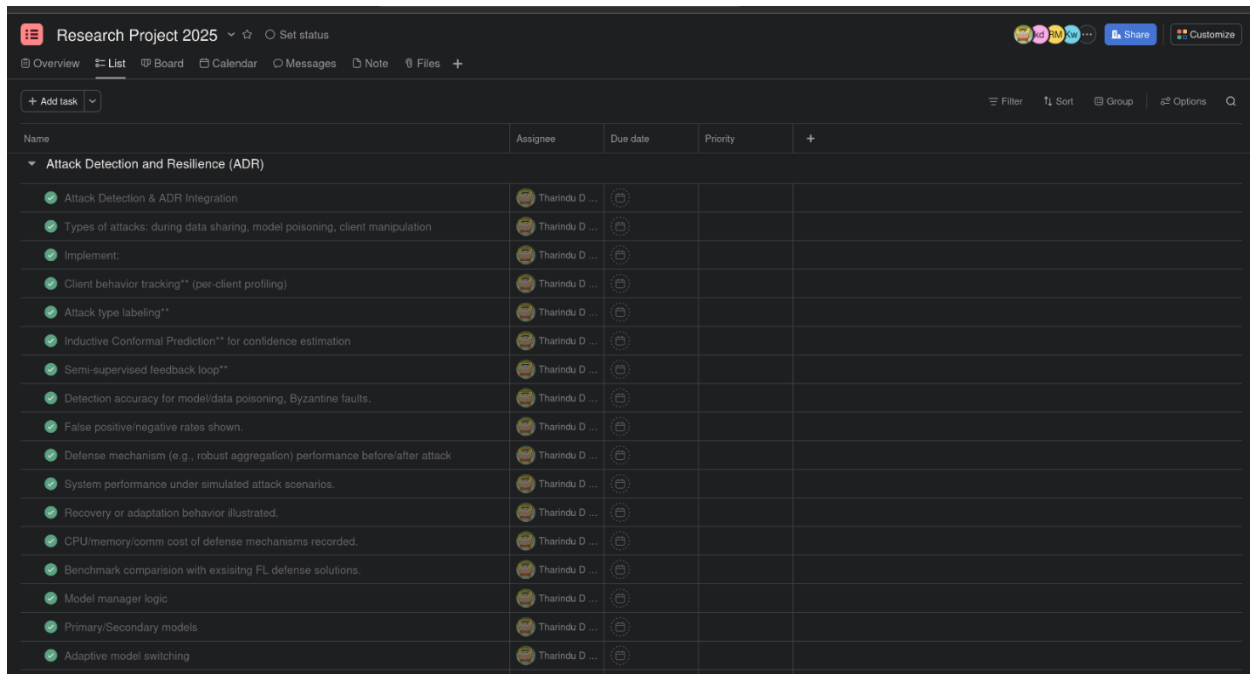
5. Task Details



Name	Assignee	Due date	Priority	
Logging and Monitoring				
Implemented Own logging system 4				
Add task...				
Ongoing Problems Identified				
GUI not updating as the CLI in the server				
Errors during Aggregation with 2 clients				
GUI issues (Refresh Error)				
Model accuracy drop (Dataset Problem)				
Irregular client participation rounds				
New clients joining dynamically				
Add task...				
Server Development				
Advanced Cryptographic Techniques 2				
Client manager 2				

Attack defense and resilience module was mine.

Personal task Assigning and Completion



Name	Assignee	Due date	Priority	
Attack Detection and Resilience (ADR)				
Attack Detection & ADR Integration	Tharindu D ...			
Types of attacks: during data sharing, model poisoning, client manipulation	Tharindu D ...			
Implement.	Tharindu D ...			
Client behavior tracking** (per-client profiling)	Tharindu D ...			
Attack type labeling**	Tharindu D ...			
Inductive Conformal Prediction** for confidence estimation	Tharindu D ...			
Semi-supervised feedback loop**	Tharindu D ...			
Detection accuracy for model/data poisoning, Byzantine faults.	Tharindu D ...			
False positive/negative rates shown.	Tharindu D ...			
Defense mechanism (e.g., robust aggregation) performance before/after attack	Tharindu D ...			
System performance under simulated attack scenarios.	Tharindu D ...			
Recovery or adaptation behavior illustrated.	Tharindu D ...			
CPU/memory/comm cost of defense mechanisms recorded.	Tharindu D ...			
Benchmark comparison with existing FL defense solutions.	Tharindu D ...			
Model manager logic	Tharindu D ...			
Primary/Secondary models	Tharindu D ...			
Adaptive model switching	Tharindu D ...			

6. System Details

6.1 System completion status

Finished

ADRM TUI

Attack Defense And Resilience Module | Privacy Preserving Module | Secure Aggrigation Module | Server Communication And Protocol Enforcement Module

Attack Defense And Resilience Module Details

```

Status:
Blocked Clients Count:
Champion Is Trained:
Challenger Is Trained:
Challenger Training Buffer Size:
Performance:

```

```

running_ml_mode
0
False
False
17
Champion: 0.0
Challenger: 0.0
History: []

```

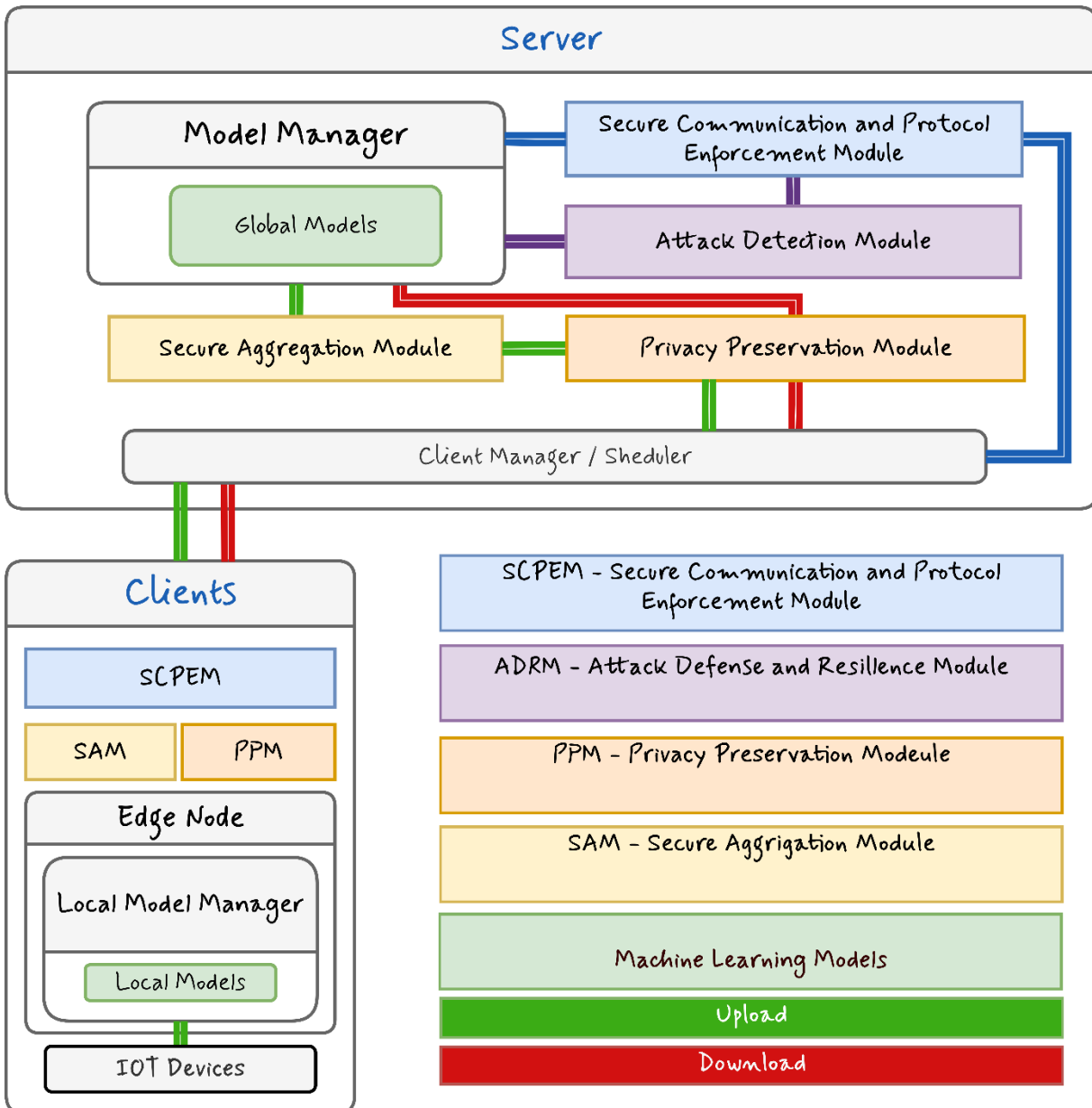
Client Health Status (Total: 4, Active: 3, Blocked: 1)

Client ID	Status	Reputation	Details
client_1	Connected	100	unknown
client_2	Connected	100	unknown
client_3	Blocked	75	Flagged as a statistical outlier compared to peers in the same round.
client_4	Connected	100	unknown

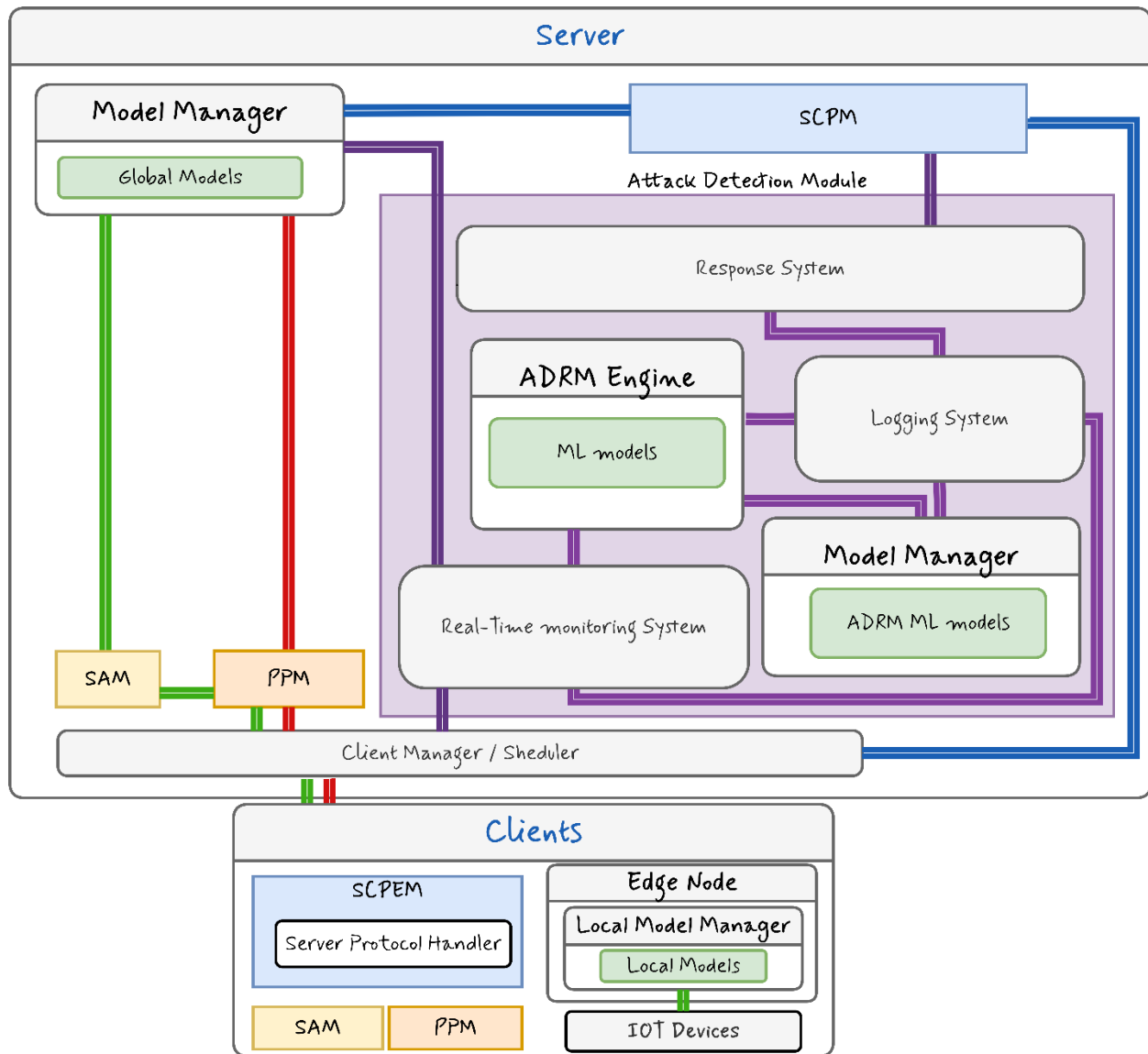
Name	Size	Modified
challenger_model.pkl	259.7 KiB	2 minutes ago

6.2 System Design

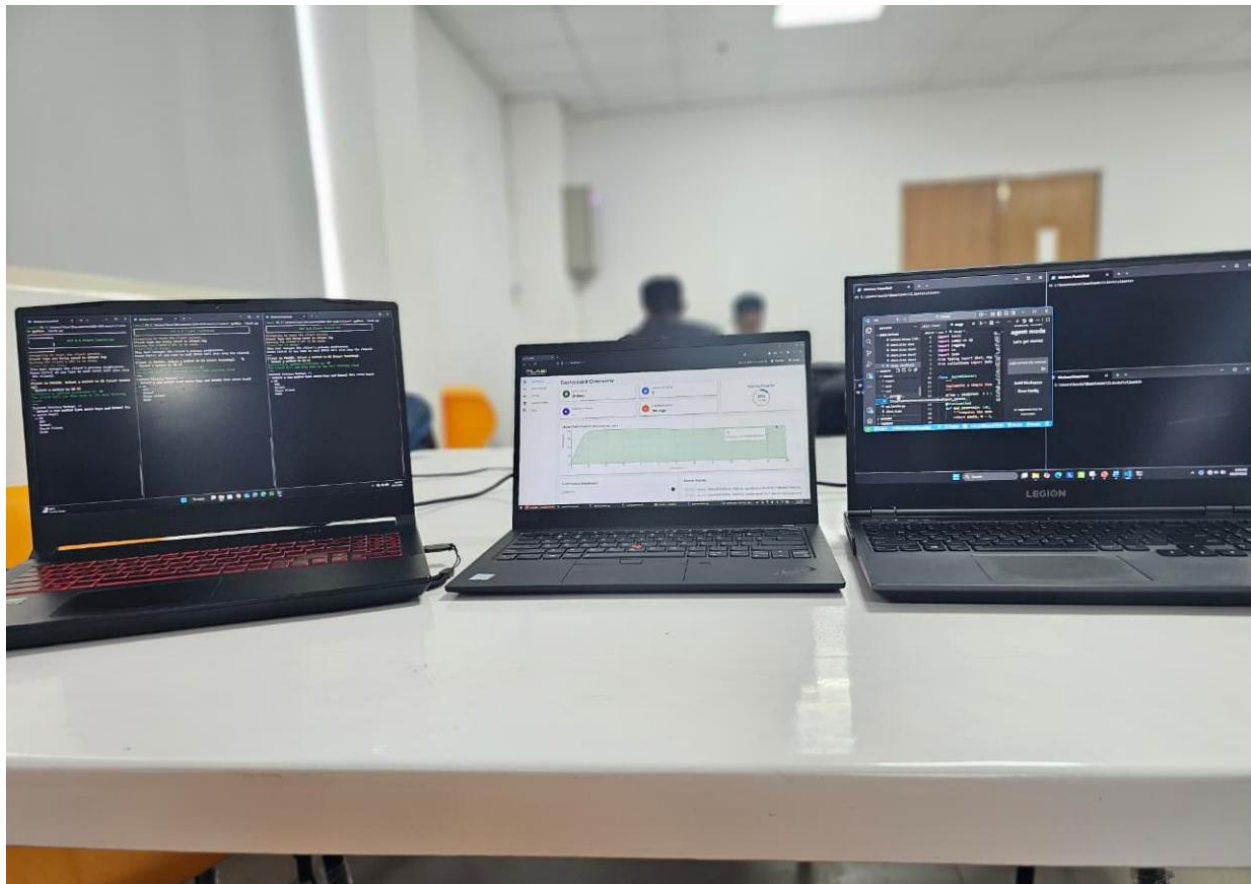
I. System Architecture



II. Module Architecture

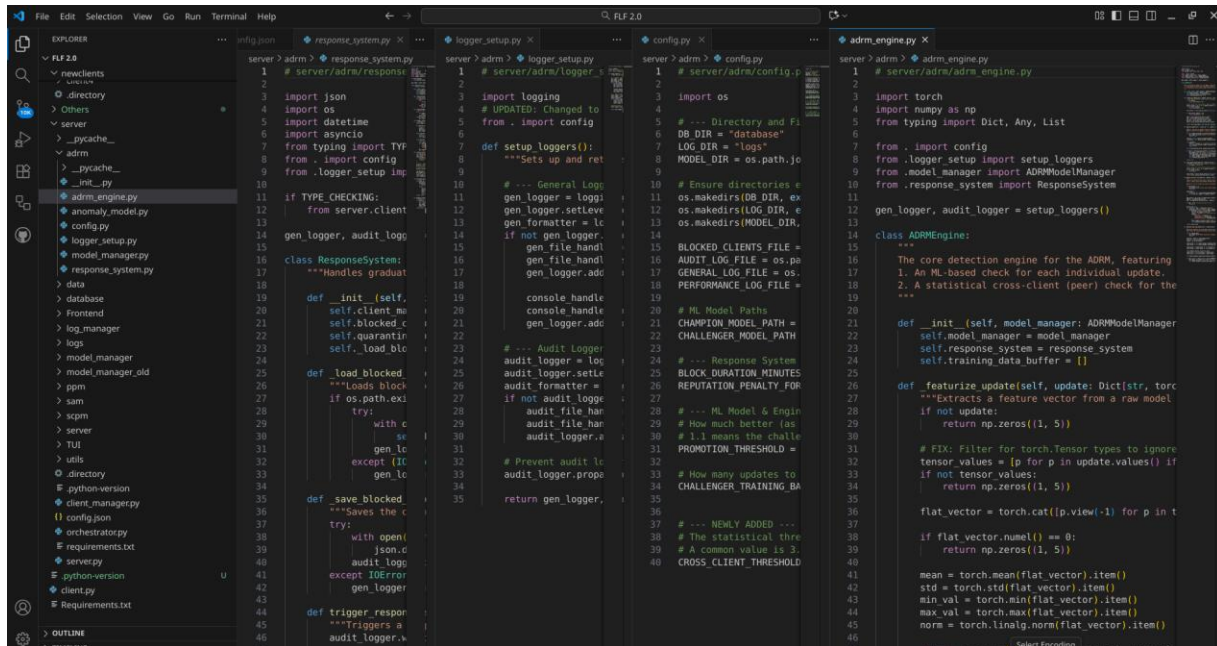


6.3 System Testing



6.4 System Codes

ADRM – Server



```

server > adm > response_system.py
1 # server/adm/response_
2
3 import json
4 import os
5 import datetime
6 import asyncio
7 from typing import TypedDict
8 from . import config
9 from .logger_setup import
10
11 if TYPE_CHECKING:
12     from server.client
13
14 gen_logger, audit_logg
15
16 class ResponseSystem:
17     """Handles graduat
18
19     def __init__(self,
20                 self.client_ma
21                 self.blocked_c
22                 self.quarantin
23                 self.load_blo
24
25     def load_blocked:
26         """Loads block
27         if os.path.exi
28             try:
29                 with c
30                     se
31                     gen_lc
32             except (IO
33                 gen_lc
34
35     def save_blocked:
36         """Saves the c
37         try:
38             with open(
39                 json.c
40             ) as f:
41                 audit_logg
42                 gen_logger
43
44     def trigger_respor
45         """Triggers a
46         audit_logger.k

```

```

server > adm > logger_setup.py
1 # server/adm/logger_
2
3 import logging
4
5 # UPDATED: Changed to
6 from . import config
7
8 def setup_loggers():
9     """Sets up and ret
10
11     # --- General Logg
12     gen_logger = loggi
13     gen_logger.setLevel
14     gen_formatter = lc
15     if not gen_logger:
16         gen_file handl
17         gen_logger.add
18         console handl
19         console handl
20         gen_logger.add
21
22     # --- Audit Logg
23     audit_logger = log
24     audit_logger.setLe
25     audit_formatter =
26     if not audit_logge
27         audit_file har
28         audit_file har
29         audit_logger.s
30
31     # Prevent audit lc
32     audit_logger.prope
33
34     return gen_logger,

```

```

server > adm > config.py
1 # server/adm/config.p
2
3 import os
4
5 # --- Directory and Fi
6 DB_DIR = "database"
7 LOG_DIR = "logs"
8 MODEL_DIR = os.path.jo
9
10 # Ensure directories e
11 os.makedirs(DB_DIR, ex
12 os.makedirs(LOG_DIR, e
13 os.makedirs(MODEL_DIR,
14
15 BLOCKED_CLIENTS_FILE =
16 AUDIT_LOG_FILE = os.pa
17 GENERAL_LOG_FILE = os.
18 PERFORMANCE_LOG_FILE =
19
20 # ML Model Paths
21 CHAMPION_MODEL_PATH =
22 CHALLENGER_MODEL_PATH
23
24 # --- Response System
25 BLOCK_DURATION_MINUTES
26 REPUTATION_PENALTY_FOR
27
28 # --- ML Model & Engin
29 # How much better (as
30 # 1.1 means the challe
31 PROMOTION_THRESHOLD =
32
33 # How many updates to
34 CHALLENGER_TRAINING_BA
35
36 # --- NEWLY ADDED ---
37 # The statistical thre
38 # A common value is 3.
39 CROSS_CLIENT_THRESHOLD
40

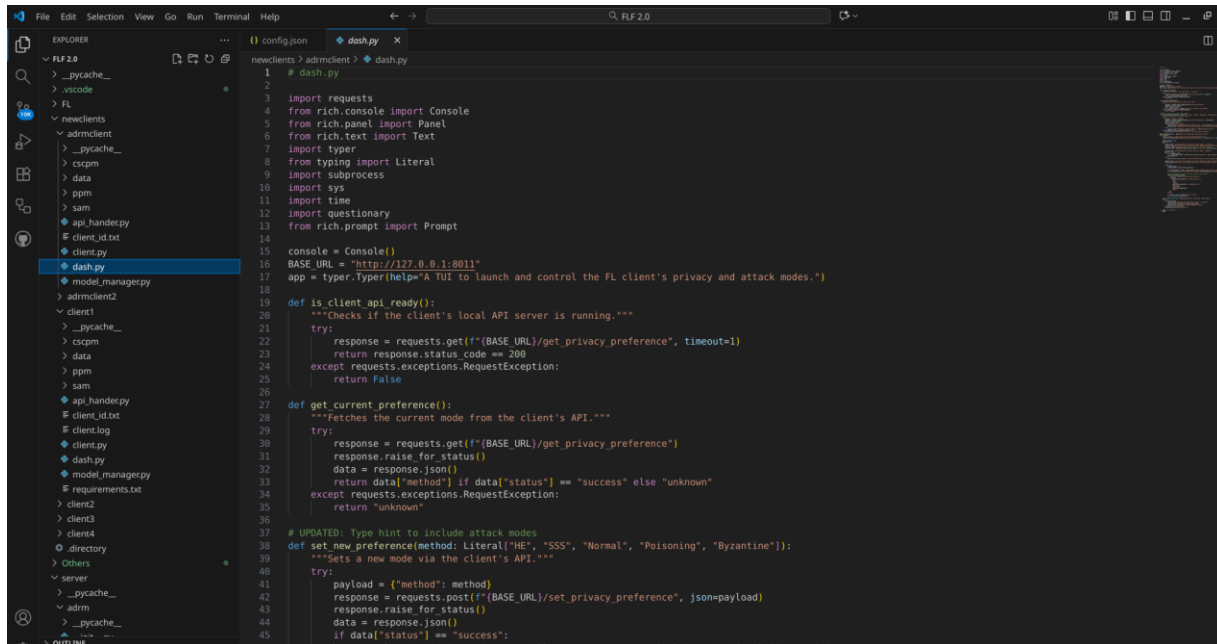
```

```

server > adm > adm_engine.py
1 # server/adm/adm_engin
2
3 import torch
4 import numpy as np
5 from typing import Dict, Any, List
6
7 from . import config
8 from .logger_setup import setup_loggers
9 from .model_manager import ADRMModelManager
10 from .response_system import ResponseSystem
11
12 gen_logger, audit_logger = setup_loggers()
13
14 class ADRMEngine:
15     """
16     The core detection engine for the ADRM, featuring
17     1. An ML-based check for each individual update.
18     2. A statistical cross-client (peer) check for the
19     """
20
21     def __init__(self, model_manager: ADRMModelManager):
22         self.model_manager = model_manager
23         self.response_system = response_system
24         self.training_data_buffer = []
25
26     def _featureize_update(self, update: Dict[str, torch.Tensor]) -> torch.Tensor:
27         """Extracts a feature vector from a raw model
28         if not update:
29             return np.zeros((1, 5))
30
31         # FIX: Filter for torch.Tensor types to ignore
32         tensor_values = [p for p in update.values() if
33                         if not tensor values:
34                             return np.zeros((1, 5))
35
36         flat_vector = torch.cat([p.view(-1) for p in t
37
38         if flat_vector.numel() == 0:
39             return np.zeros((1, 5))
40
41         mean = torch.mean(flat_vector).item()
42         std = torch.std(flat_vector).item()
43         min_val = torch.min(flat_vector).item()
44         max_val = torch.max(flat_vector).item()
45         norm = torch.linalg.norm(flat_vector).item()
46

```

ADRM – Client



```

server > adm > config.py
1 # server/adm/config.p
2
3 import os
4
5 # --- Directory and Fi
6 DB_DIR = "database"
7 LOG_DIR = "logs"
8 MODEL_DIR = os.path.jo
9
10 # Ensure directories e
11 os.makedirs(DB_DIR, ex
12 os.makedirs(LOG_DIR, e
13 os.makedirs(MODEL_DIR,
14
15 BLOCKED_CLIENTS_FILE =
16 AUDIT_LOG_FILE = os.pa
17 GENERAL_LOG_FILE = os.
18 PERFORMANCE_LOG_FILE =
19
20 # ML Model Paths
21 CHAMPION_MODEL_PATH =
22 CHALLENGER_MODEL_PATH
23
24 # --- Response System
25 BLOCK_DURATION_MINUTES
26 REPUTATION_PENALTY_FOR
27
28 # --- ML Model & Engin
29 # How much better (as
30 # 1.1 means the challe
31 PROMOTION_THRESHOLD =
32
33 # How many updates to
34 CHALLENGER_TRAINING_BA
35
36 # --- NEWLY ADDED ---
37 # The statistical thre
38 # A common value is 3.
39 CROSS_CLIENT_THRESHOLD
40

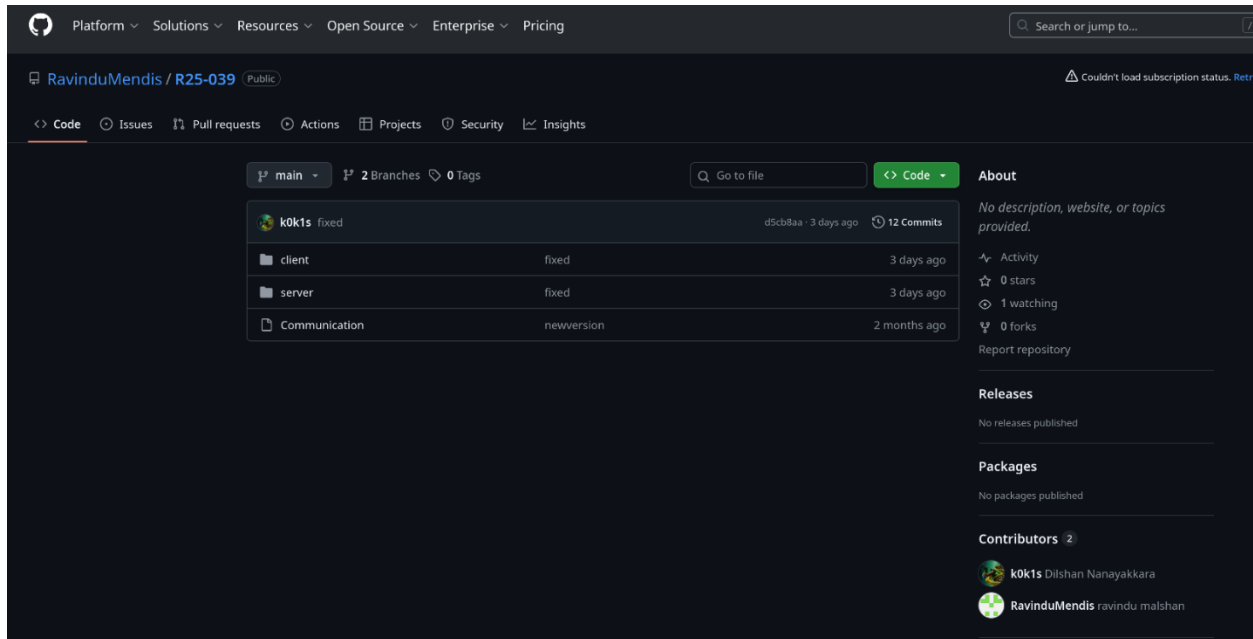
```

```

server > adm > dash.py
1 # dash.py
2
3 import requests
4 from rich.console import Console
5 from rich.panel import Panel
6 from rich.text import Text
7 import typer
8 from typing import Literal
9 import subprocess
10 import sys
11 import time
12 import questionary
13 from rich.prompt import Prompt
14
15 console = Console()
16
17 BASE_URL = "http://127.0.0.1:8011"
18 app = typer.Typer(help="A TUI to launch and control the FL client's privacy and attack modes.")
19
20 def is_client_api_ready():
21     """Checks if the client's local API server is running."""
22     try:
23         response = requests.get(f"{BASE_URL}/get_privacy_preference", timeout=1)
24         return response.status_code == 200
25     except requests.exceptions.RequestException:
26         return False
27
28 def get_current_preference():
29     """Fetches the current mode from the client's API."""
30     try:
31         response = requests.get(f"{BASE_URL}/get_privacy_preference")
32         response.raise_for_status()
33         data = response.json()
34         return data["method"] if data["status"] == "success" else "unknown"
35     except requests.exceptions.RequestException:
36         return "unknown"
37
38 # UPDATED: Type hint to include attack modes
39 def set_new_preference(method: Literal["HE", "SSS", "Normal", "Poisoning", "Byzantine"]):
40     """Sets a new mode via the client's API."""
41     try:
42         payload = {"method": method}
43         response = requests.post(f"{BASE_URL}/set_privacy_preference", json=payload)
44         response.raise_for_status()
45         data = response.json()
46         if data["status"] == "success":
47             console.print(f"Mode set to {method}.")
48         else:
49             console.print(f"Failed to set mode to {method}.")

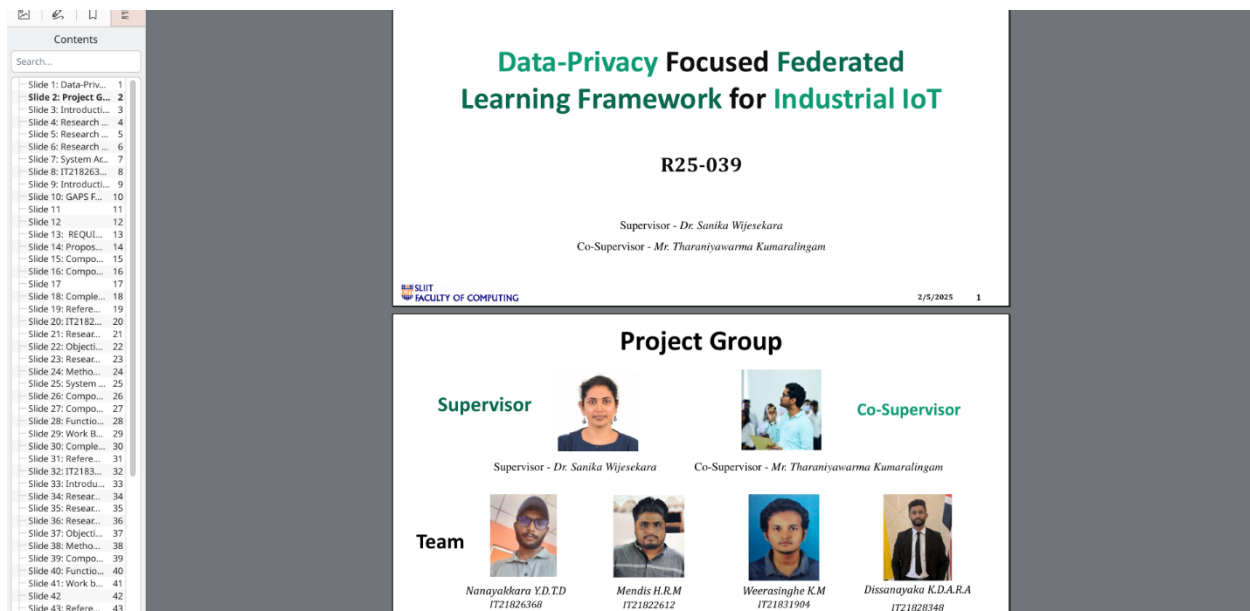
```

7. GitHub Upload



8. Documentation

8.1 Proposal



8.2 Presentation 1

Contents

Search...

- Slide 1: Data-Priv... 1
- Slide 2: Project Gr... 2
- Slide 3: Introduct... 3
- Slide 4: Research... 4
- Slide 5: Research... 5
- Slide 6: Research... 6
- Slide 7: System Ar... 7
- Slide 8: IT218263... 8
- Slide 9: Introduct... 9
- Slide 10: GAPS F... 10
- Slide 11... 11
- Slide 12... 12
- Slide 13: REQU... 13
- Slide 14: Propo... 14
- Slide 15: Compo... 15
- Slide 16: Compo... 16
- Slide 17... 17
- Slide 18: Comple... 18
- Slide 19: Refere... 19
- Slide 20: IT2182... 20
- Slide 21: Resear... 21
- Slide 22: Object... 22
- Slide 23: Resear... 23
- Slide 24: Metho... 24
- Slide 25: System... 25
- Slide 26: Compo... 26
- Slide 27: Compo... 27
- Slide 28: Functio... 28
- Slide 29: Work B... 29
- Slide 30: Comple... 30
- Slide 31: Refere... 31
- Slide 32: IT2183... 32
- Slide 33: Introdu... 33
- Slide 34: Resear... 34
- Slide 35: Resear... 35
- Slide 36: Resear... 36
- Slide 37: Object... 37
- Slide 38: Metho... 38
- Slide 39: Compo... 39
- Slide 40: Functio... 40
- Slide 41: Work b... 41

Real-world testing

There is a lack of large scale IoT datasets for testing privacy preserving techniques under real-world conditions. Existing research often relies on synthetic data, limiting the generalizability of results.

Regulatory compliance

Privacy preserving methods need to be compatible with existing data protection laws like GDPR, but there's a gap in evaluating how these solutions perform in compliance contexts.

Resistance to privacy attacks

While current privacy preserving methods aim to protect data, the robustness of these methods against evolving privacy attacks in IIoT remains insufficiently addressed.

Energy consumption and efficiency

Many privacy-preserving methods are computationally intensive, posing significant challenges to resource-constrained IIoT devices, affecting their energy efficiency.

Methodology


Approach:

- Analyze existing FL privacy vulnerabilities.
- Combine HE and DP for enhanced privacy.
- Optimize techniques for IIoT-specific constraints.
- Validate Using real-world Datasets

Key Techniques:

- Homomorphic Encryption (HE):** Encrypts gradients, allowing computations on encrypted data without decrypting it. Prevents data leakage even if adversaries intercept communications.
- Differential Privacy (DP):** Ensure that individual data points cannot be separated by adding controlled noise to gradients. Balances model accuracy with privacy.

System Architecture



8.3 Presentation 2

Contents

Search...


- > OVERALL 1
- > ADMIN COMPONENT... 15
- > COMPONENT 2 PRL... 23
- > COMPONENT 3 SEC... 33
- > COMPONENT 4 SCPM... 40
- > GENERAL END SLIDES 48

DATA-PRIVACY FOCUSED FEDERATED LEARNING FRAMEWORK FOR INDUSTRIAL IOT


R25 - 039

PROJECT GROUP

Team




Nanayakkara Y.D.T.D
IT21826368



Mendis H.R.M
IT21826212

Supervisors



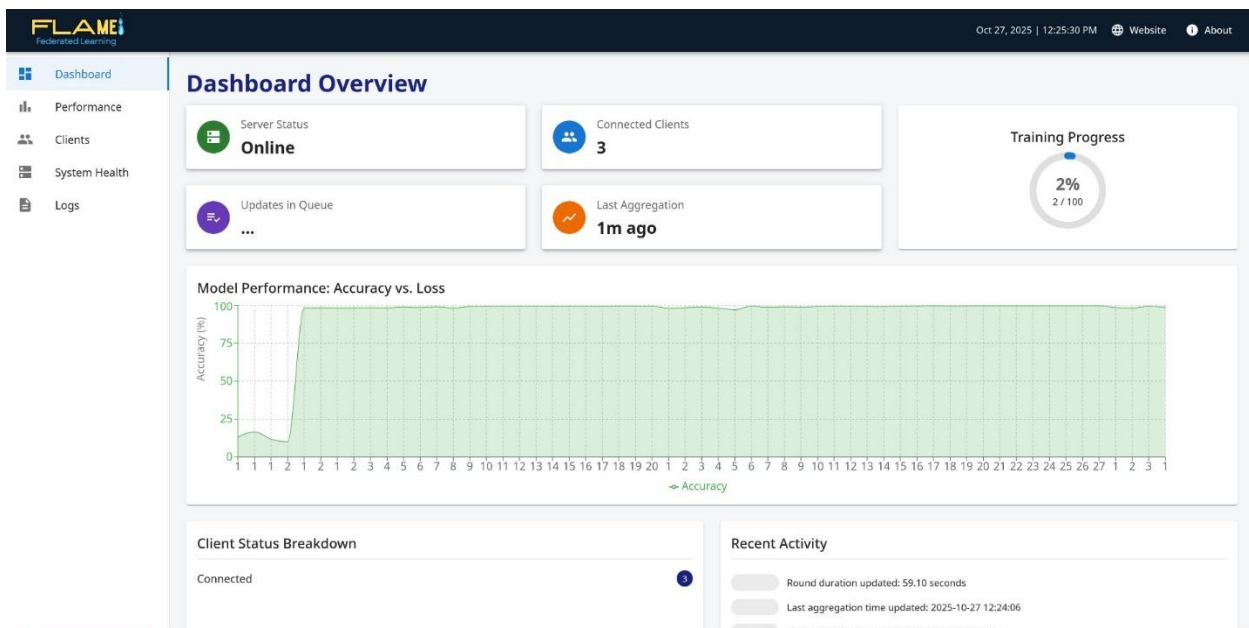
Mr. Amila Nuwan Senarathne
Supervisor

8.4 Final Presentation



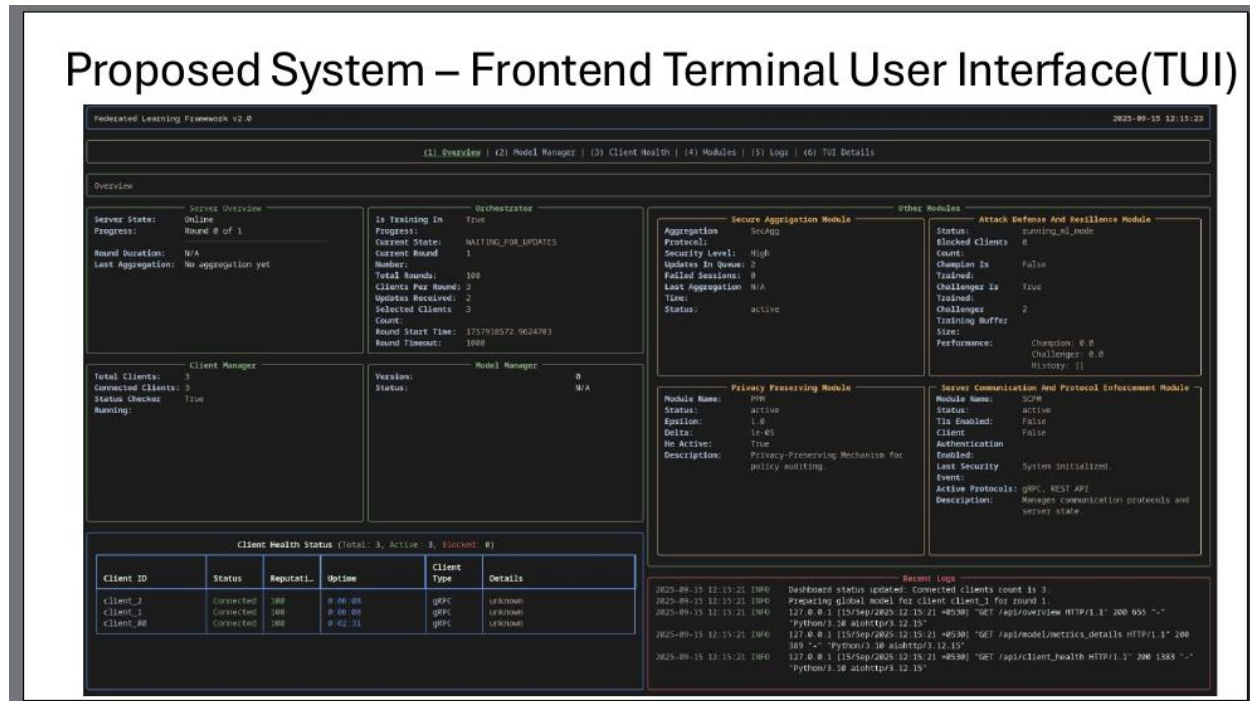
8.5 Final Product

Web Portal Frontend



Terminal User interface (Frontend)

Proposed System – Frontend Terminal User Interface(TUI)



8.6 Research Paper

III. Conference Appetence

To Tharindu D Nanayakkara <dilz.nanayakkara@gmail.com> 📧

10/29/25, 11:27 AM

Acceptance Notification

Dear Tharindu D Nanayakkara,

Congratulations! We are pleased to inform you that your paper has been accepted as a regular paper to be presented at the 7th International Conference on Advancements in Computing 2025.

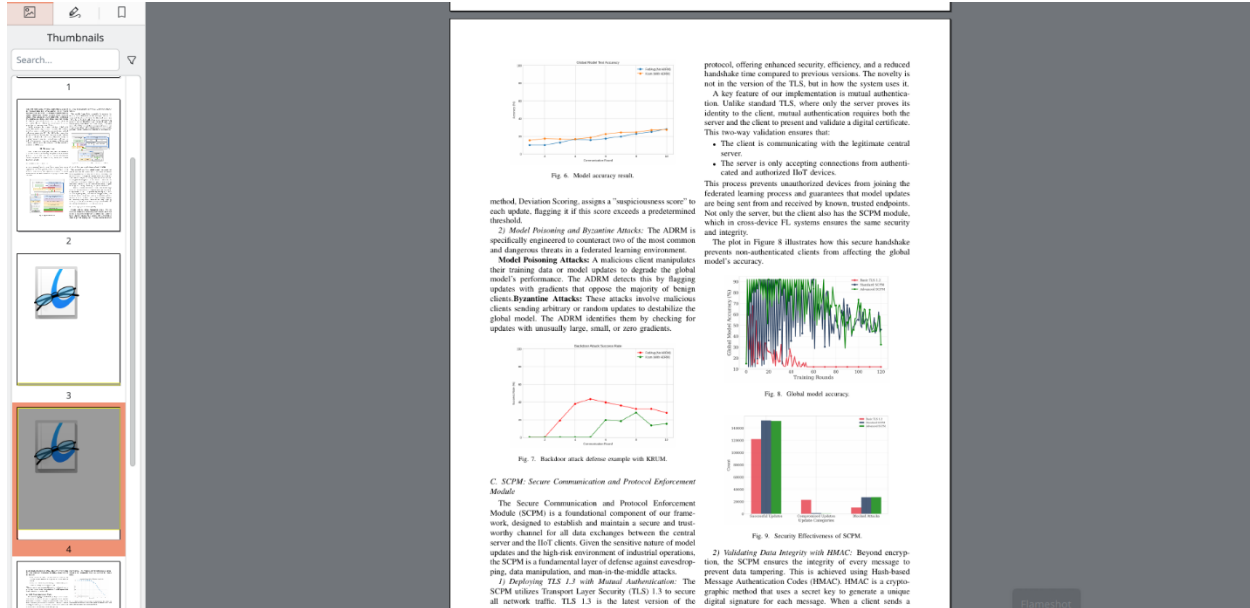
Paper ID: 469

Paper Title: Data-privacy based Federated Learning Framework for Industrial IOT

Please visit <https://cmt3.research.microsoft.com/7ICAC2025/Submission/Index> to view the reviews given during the double-blind review process.

When preparing the camera-ready version of your paper, please address all the review comments and follow the camera-ready guidelines given in the <https://icac.lk/for-authors>

Please note that the camera-ready deadline is 10th November 2025 and camera-ready submission portal on CMT will be available starting from 22nd October 2025.



9. CDAP upload



CDAPSubmissionCloud

Private group

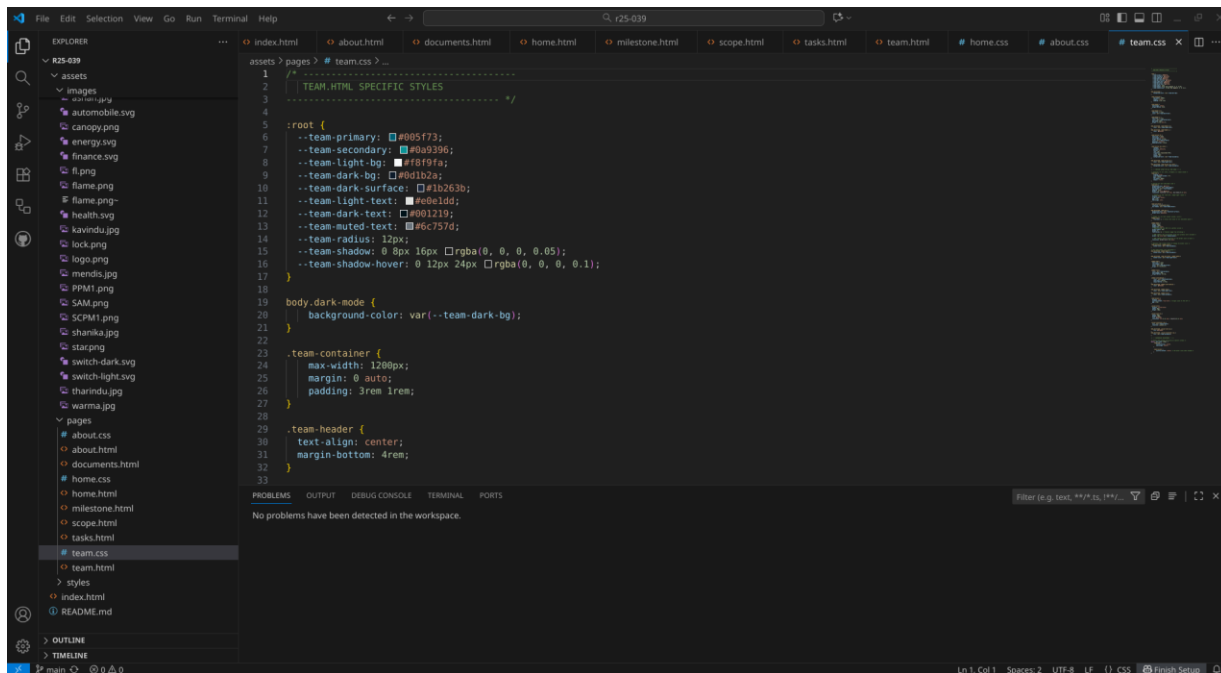
+ New Upload Edit in grid view Share Copy link Add shortcut to OneDrive Download Export to Excel Automate Integrate Sync

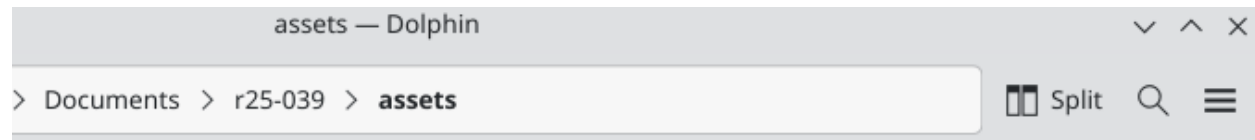
2025RegCloud > R25-039-Students

Name	Modified	Modified By
1. Project Proposal	January 27	Tharaniyawarma Kumaralingam
2. Progress Presentation - 1	January 27	Tharaniyawarma Kumaralingam
3. Progress Presentation - 2	January 27	Tharaniyawarma Kumaralingam
4. Research Paper	January 27	Tharaniyawarma Kumaralingam
5. Final Report & Presentation	January 27	Tharaniyawarma Kumaralingam
6. Check List Documents	April 29	CDAP SLIIT
7. Website	January 27	Tharaniyawarma Kumaralingam
8. Log Book	January 27	Tharaniyawarma Kumaralingam
Marking Schemes	January 27	Tharaniyawarma Kumaralingam
Project Registration Documents	January 27	Tharaniyawarma Kumaralingam
Panel Comments for the Students.xlsm	September 20	CDAP SLIIT

10. Website

10.1 Development





pages



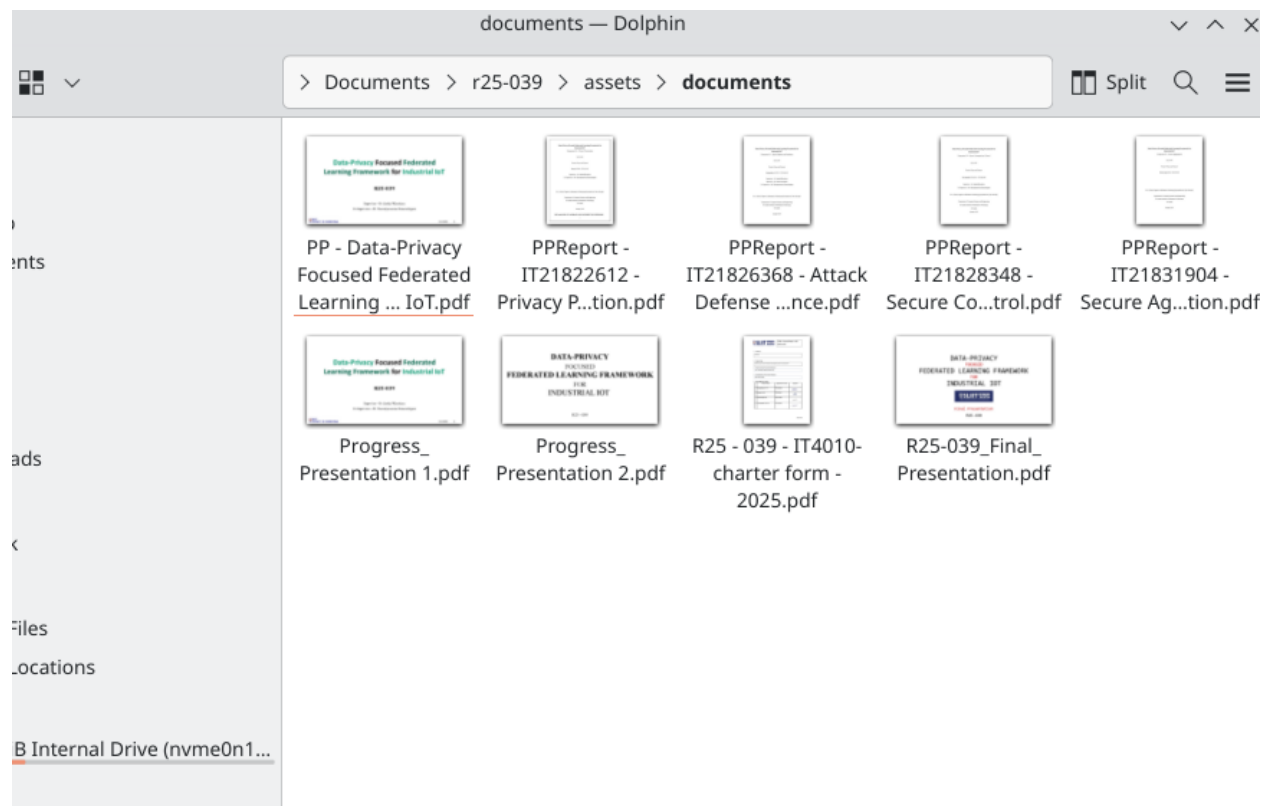
styles

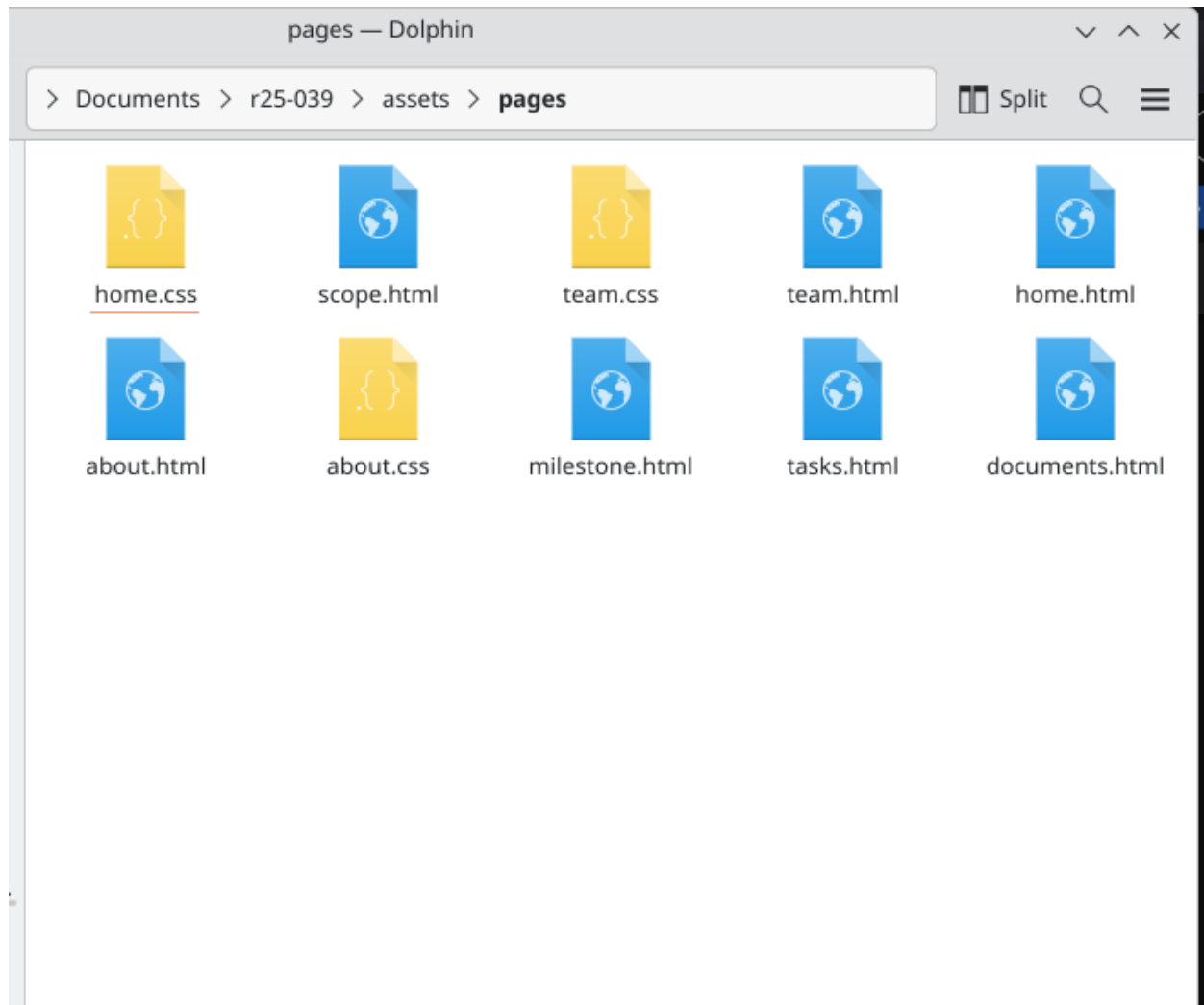


images




documents





10.2 Finalize



Home About Scope Milestones Tasks Docs Team

A Privacy-First Federated Learning Framework

Building secure, resilient, and efficient ML models without compromising data privacy. Designed for the demands of the Industrial IoT.

F . L . A . M . E

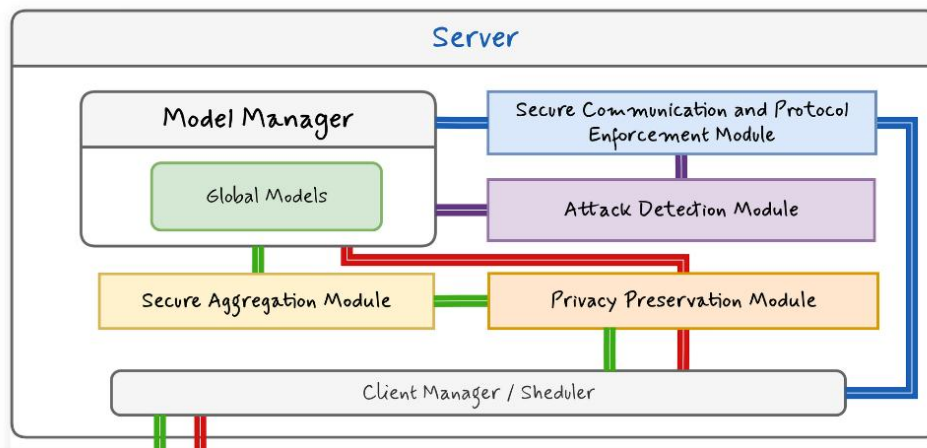
Federated Learning Aggrigated Modular Environment

[Explore Features](#) [View on GitHub](#)

Privacy-Enhanced Federated Learning Framework

Our comprehensive Federated Learning (FL) System Framework is engineered to significantly augment the privacy, security, and operational resilience of machine learning models deployed in decentralized and distributed environments. The framework is composed of four interconnected core modules, collectively guaranteeing data integrity, defense against adversarial attacks, and authenticated inter-component communication.

Framework Overview and Architecture





Mr. Amila Senerathne

Supervisor



Dr. Sanika Wijesekara

External Supervisor



Mr. T. Kumaralingam

Co-Supervisor

