Progress Evaluation: Milestone 4

Project: *The Phisherman - AI-Powered Phishing Email Detector*

Team Members:
Curtis Jones - *cjones2022@my.fit.edu*
Elton Batista - *ebatista2022@my.fit.edu*
Jordan Chesley - *jchesley2022@my.fit.edu*

Faculty Advisor: Khaled Slhoub - *kslhoub@fit.edu*

Client: Khaled Slhoub - *College of Engineering and Science: Department of Electrical Engineering and Computer Science*

Milestone 4 Task Matrix

Task	Completion	Curtis	Elton	Jordan	To Do
Implement remaining OAuth functions	*100%*	*N/A*	*N/A*	*Implemented the remainder of OAuth functionality (report / delete)*	*Completed*
Implement supporting functions from Worker/Queue and Database modules	*80%*	*Implemented basic worker/queue functionality for batch processing*	*Integrated Phish Tank and designed the process for refreshing the hashmap*	*Created the ER diagram structurizing the system's database and began database integration*	*Finish integration of database with the system to streamline process for returning users, test batch processing through worker queue and integrate with database functionality, implement hard-coded functionalities to assist accuracy and performance*
Refine methods on training AI models and train Artificial Intelligence model on new data	*70%*	*Improved implementation of datasets and researched new sources (LLM responses, private databases)*	*Strengthened detection layer by incorporating real-world threat-intelligence data.*	*N/A*	*Continue implementation through newly researched methods and databases*

Perform Testing Artificial Intelligence model	*.70%*	*Conducted testing on various methods to improve the AI model on real-world data*	*N/A*	*N/A*	*Continue testing through newly researched methods and databases*
Implement remainder of UI elements as outlined in the design documents	*90%*	*Assisted in merging codebases and implemented the remainder of required UI elements.*	*N/A*	*Made major refinements to the webapp's UI, including code structure and aesthetics*	*Implement remainder of screen images and finish integration of user statistics by utilizing our database.*

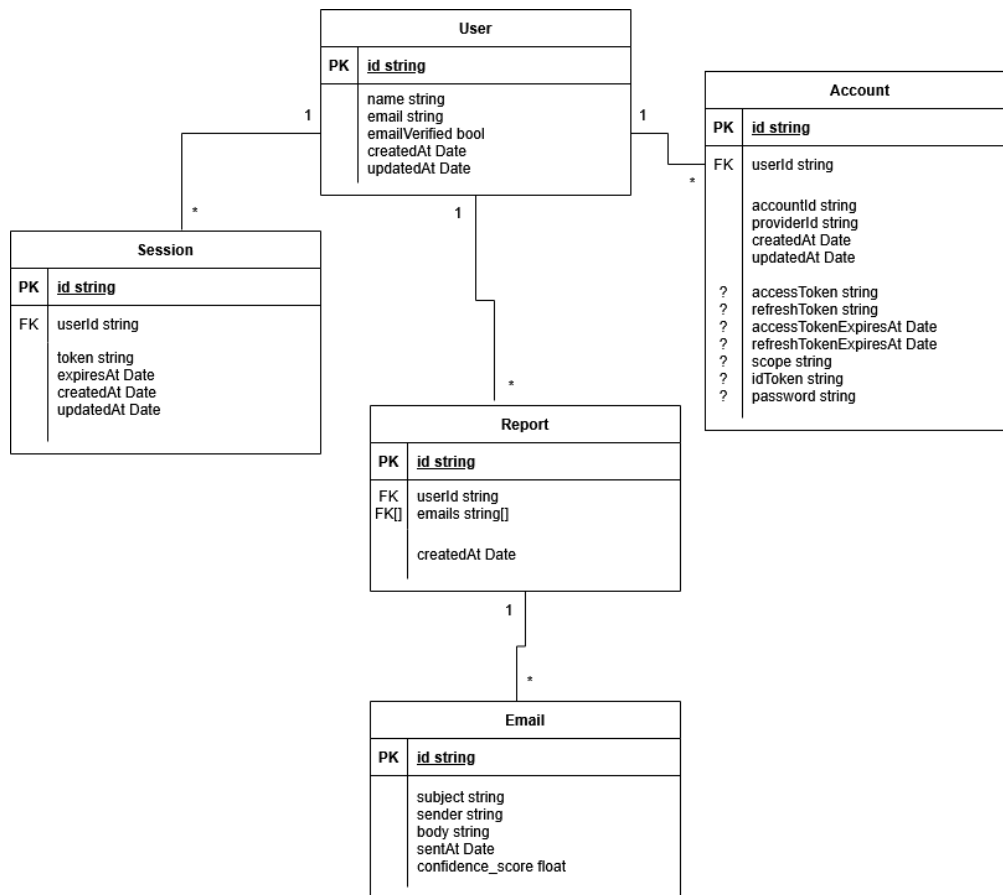# Milestone 4 Task Discussion:

Implement remaining OAuth functions

Implemented the remaining OAuth functionality to delete emails. Gmail does not have an API feature to report emails, but emails can be moved to the spam folder, so that feature was implemented, and these emails will be deleted in 30 days. All UI elements in this regard are fully functional.

Implement supporting functions from Worker/Queue and Database modules

Implemented basic worker/queue functionality to improve system performance by processing emails in batches. We also began the design and implementation of the Database into our microservice infrastructure.

Database ER diagram

User

PK	id string
	name string
	email string
	emailVerified bool
	createdAt Date
	updatedAt Date

Account

PK	id string
FK	userId string
	accountId string
	providerId string
	createdAt Date
	updatedAt Date
?	accessToken string
?	refreshToken string
?	accessTokenExpiresAt Date
?	refreshTokenExpiresAt Date
?	scope string
?	idToken string
?	password string

Session

PK	id string
FK	userId string
	token string
	expiresAt Date
	createdAt Date
	updatedAt Date

Report

PK	id string
FK	userId string
FK[]	emails string[]
	createdAt Date

Email

PK	id string
	subject string
	sender string
	body string
	sentAt Date
	confidence_score float

Refine methods on training AI models and train AI model on new data

Upgraded the model by training it on our new datasets and training methods. We researched new data sources, including LLM-generated emails and APWG databases. We are currently working on implementing hard-coded checks to improve accuracy in areas the AI model would be lacking, including checking the incoming IP address of the sender and attached URLs in the email.

Perform Testing AI model

Conducted testing using our refined methods and currently available data.

Implement remainder of UI elements as outlined in the design documents

Merged code bases and implemented the remainder of UI elements. All required UI elements are currently added while we move on to add the required UI images and aesthetics/animations.

Discussion of the contribution of each team member to the current Milestone
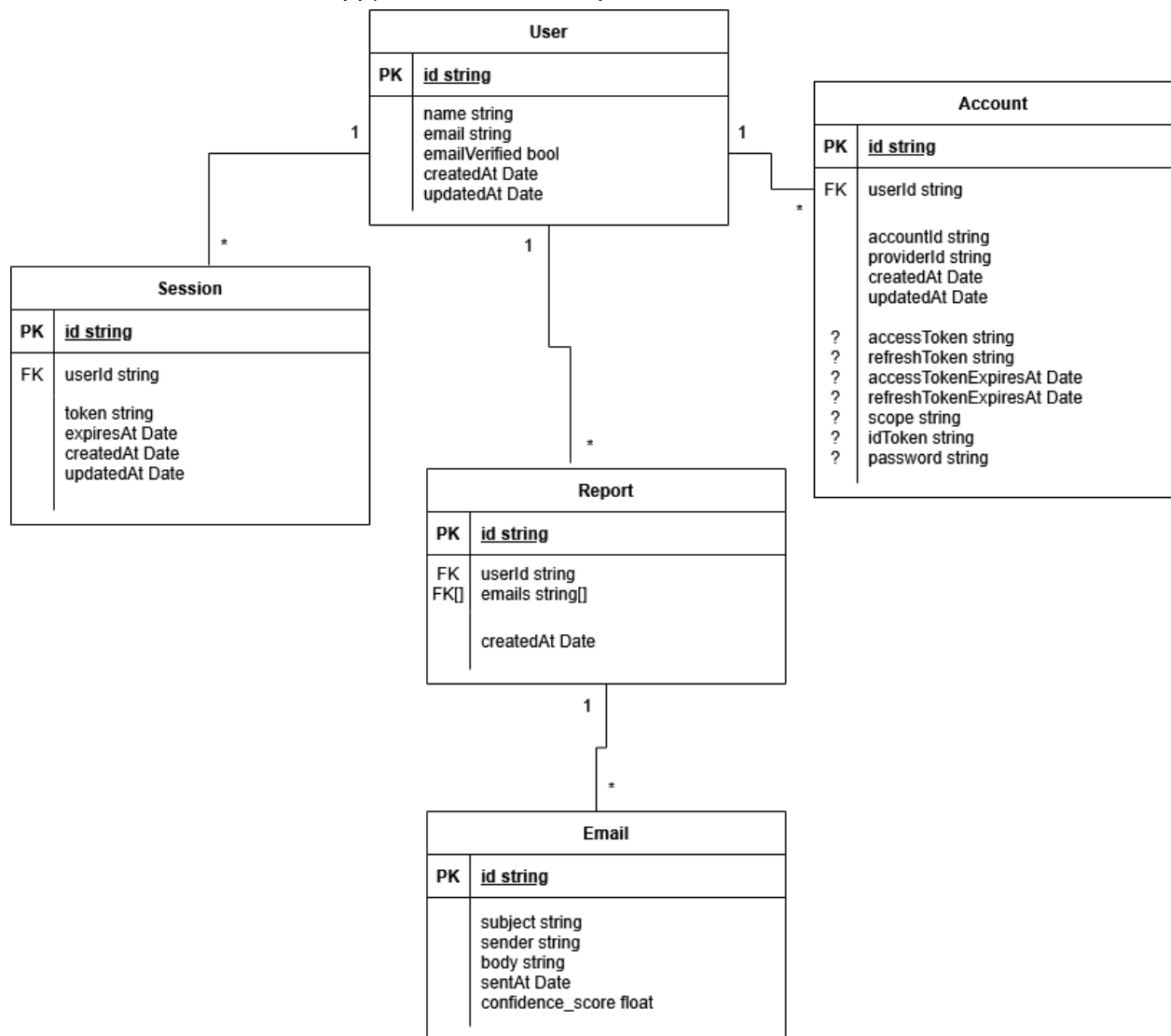
Curtis Jones

- Improved implementation of available datasets to reinforce the AI model.

- Conducted testing on various methods to improve the AI model on real-world data

- Researched different sources of datasets that would be beneficial to improving our model:
  - LLM generated email responses - Used by more recent phishing attacks due to the increased availability of LLMs
  - Collecting real world data - Generating datasets from data provided by peers
  - Private Databases - Using accredited databases through private access, such as the Anti-Phishing Working Group, which works with over 2200 well known institutions to collect and maintain phishing data. This group provides access to its data for free for academic purposes with strict criteria.
  - Alternative methods - Training the model based on data we have yet to explore, such as URLs and attachments. This could include sub-models or hard-coded checks using outside APIs to increase the accuracy of our system.

- Implemented basic worker/queue functionality to improve system performance by processing emails in batches.

- Assisted in merging our codebases and implemented the remainder of the required UI elements outlined in the design document.

Elton Batista

- Integrated PhishTank into our backend workflow by using its community-reported phishing URLs to maintain an updated hashmap of known malicious links.

- Designed the process for periodically refreshing this hashmap and checking any URLs extracted from incoming emails against the verified phishing list.

- Strengthened the system's early-stage detection layer by incorporating real-world threat-intelligence data into the email-scanning pipeline.

- Contributed to discussions on extending this approach to sender-IP reputation checks using services like Spamhaus for emails where OAuth exposes the originating IP.

- Helped outline how IP-based reputation scoring could complement URL-based detection to improve overall system robustness.

Jordan Chesley

- Implemented the remaining OAuth functionality to delete emails.

- Made major refinements to the webapp's UI, including code structure, optimizations, and aesthetics.

- Merged our multiple codebases into a monorepo.

- Created the ER diagram structurizing the system's database.

- Began designing the containerization process for our microservice structure (AI server, database, and webapp) with Docker Compose.

User

PK	id string
	name string
	email string
	emailVerified bool
	createdAt Date
	updatedAt Date

Account

PK	id string
FK	userId string
	accountId string
	providerId string
	createdAt Date
	updatedAt Date
?	accessToken string
?	refreshToken string
?	accessTokenExpiresAt Date
?	refreshTokenExpiresAt Date
?	scope string
?	idToken string
?	password string

Session

PK	id string
FK	userId string
	token string
	expiresAt Date
	createdAt Date
	updatedAt Date

Report

PK	id string
FK	userId string
FK[]	emails string[]
	createdAt Date

Email

PK	id string
	subject string
	sender string
	body string
	sentAt Date
	confidence_score float

Milestone 5 Task Matrix

Task	Jordan	Elton	Curtis
Implement remaining functionality from Worker/Queue and Database modules	*40%*	*40%*	*20%*
Finish testing and implementation of AI model functionality	*60%*	*30%*	*10%*
Finish UI implementation and styling according to the design document	*80%*	*10%*	*10%*
Implement security procedures	*10%*	*80%*	*10%*
Conduct evaluation of system and analyze results	*33%*	*33%*	*33%*
Create poster for Senior Design Showcase	*33%*	*33%*	*33%*

Milestone 5 Task Discussion:

Implement remaining functionality from Worker/Queue and Database modules

Complete the integration of the database into the system to streamline the process for returning users. Batch processing must be finished to implement supporting functionalities discussed throughout the progress evaluation, including hard-coded functionalities.

Finish testing and implementation of AI model functionality

Testing of the AI model must include newly researched methods. Additionally, we must utilize the database to provide administrative functionality to continuously improve the model. This requires a python script to retrain the model to include curated emails sent to the admin side of the system that users delete or mark as spam. Additionally, we must train on different models. Currently we use a RoBERTa based machine learning model for our emails; it may prove beneficial to include testing based on LLM models such as Ollama and ChatGPT to compare performance/accuracy. We must also include a separate LLM implementation that discusses with users how their email was found to be phishing for each email.

Finish UI implementation and styling according to the design document

Implement the remainder of the screen images, aesthetics, and animations, including the integration of backend functionality not yet completed with existing UI elements.

Implement security procedures

Design, implement, and test security procedures to protect the privacy of users. This includes storing tokens using AES-256 encryption, salting email hashes with a user-specific key, and enforcing HTTPS on all system endpoints. Legal acknowledgements have to be implemented into the UI as well.

Conduct evaluation of system and analyze results

Conduct a full evaluation of the system to analyze the results of our system implementations.

Create poster for Senior Design Showcase

Create the poster to showcase the project at the Senior Design Showcase

1. Date(s) of meeting(s) with Client during the current milestone:

   2/3
   2/19

2. Client feedback on the current milestone
   - ... (if Client and Faculty Advisor are the same, write "see Faculty Advisor Feedback below")
   - ...
   - ...
3. Date(s) of meeting(s) with Faculty Advisor during the current milestone: …

   2/3

   2/19

4. Faculty Advisor feedback on each task for the current Milestone
   - Task 1:
   - Task 2:
   - Task 3:
   - Task 4:
5. Faculty Advisor Signature: ____Dr Slhoub_____ Date: ___2/23/2026_____