**CS673 Software Engineering**

**Team 1 - Alpha**

**Meeting Minutes**

# Meeting 2

**Date and Time:** 15:21 PM Thursday, September 18, 2025

**Place**: Room 101, 1010 Commonwealth Ave, Boston, MA 02215

**Participants:** Ming Leong "Lucas" Tsui, Sujan, Xinyu "Iris" Wang

**Minutes taker:** Xinyu Wang

**Timekeeper:** Xinyu Wang

**Goals:**

* Understand the current situation with handling PDF tables and how to convert it into structured tables
* Clarify the next steps for database connection, data processing and how to present through UI
* Define responsibilities between backend(database management) and frontend(UI, visualization, security)
* Align expectations on output, deliverables and testing approach

**Discussions:**

1. Data Management

* Already extract some pdf
* Extracting data from PDF is doable(using OCR).But It’s better if we can figure out how to use ML(increase depth)
* Current challenges: database connection, large data volume(1000+ pdf documents), and security concerns(data cannot be exposed)

1. Database Management

* Data includes property contacts, rental agreement, transactions,etc.
* Main task now: connect database to python (via Flask) / [ import mysql.connector]

1. Backend Work

* Python will be used to handle calculations
* Schema is required
* Need to implement basic analysis first, then consider ML models later for predictive judgements

1. Frontend & UI

* UI must presenting results in a way that supports decision-making (such as flagging poor-performing properties)
* UI can build basic framework first, once backend is stable, it can run really fast

1. Team collaboration:

* Everyone should clarify what they promise to deliver to avoid delays
* UI and backend work must stay in sync

1. Testing & Deployment

* For now, focus on SQL and Python connecting test
* Later stages: CI/CD pipelines for automated testing
* Ensure files and database steps are clear

**Key Decisions:**

* OCR is ok, but machine learning is better
* Need to figure out the data analysis scope
* UI part: prepare basic framework
* Python-SQL connection using Flask

**Action Items:**

**For the iteration 0 release (ddl:next thursday 6:00 a.m.)**

* ~~Xinyu need to handle software test document~~
* ~~Lucas need to handle software design document and SPPP(Xinyu helps writing 5&6)~~
* ~~Boya need to handle the risk management document~~
* ~~Xinyu create the presentation (discussing with Lucas)~~

**For Software development lifecycle**

* ~~Zihan needs to clarify the difficulties of ML compared to OCR. If ML doesn’t work, she needs to change her working methods and plans.~~
* ~~Xinwen try to build the UI interface (what to display)~~
* ~~Xinyu figure out what she can test at the moment~~
* ~~Boya should deal with the security problems and JIRA management~~
* ~~SJ convert data in .bak format to flask and brief Lucas how to manipulate using Python.~~
* ~~Lucas provide Schema to SJ~~

# Meeting 1

**Date and Time:** 4:22 PM Thursday, September 11, 2025 (EDT)

**Place**: Room 101, 1010 Commonwealth Ave, Boston, MA 02215

**Participants:** Boya Zhao, Ming Leong "Lucas" Tsui, Sujan, Xiwen "Sean" Fang, Xinyu "Iris" Wang, Zihan Wang

**Minutes taker:** Leong Ming Tsui

**Timekeeper:** Ming Leong Tsui

**Goals:**

* Articulate product goals and adjust until everybody confirms it is doable within this semester given their current skillset.
* Onboarding: share contacts, class schedule, git, JIRA. Why: so everybody could know when and when not to communicate and work, and how. Discord DMs are very distracting.
* Have everyone work on something to kickstart the project to measure members capabilities and probe points of friction.
* Test AI-assisted workflow: use AI to recall meetings and draft minutes; connect AI to repo for suggestions to save time to learn how and integrate AI into workflow

**Discussions:**

* Contextualization: Lucas explained that the client oriented value of the project is to make data analysis easier given that documents recording transactions come in a multitude of formats.
* Coordination: Xinyu needs to know what language Zihan uses to start her work. Zihan confirms she will use python.
* Obstacle: Zihan expressed concern over the complexity involved in training a machine learning algorithm from scratch given our limited time and infrastructure. Training a machine learning algorithm to read raw pdfs is beyond team capabilities. Alternatives like OCR take more time-cost than project scope allows.
* Obstacle: there is a real security risk of data leaks given that source code and dataset is publicly accessible. Boya can implement security measures like account log in, 2FA, but risk of data breaches remains unsolved.
* Exploration: we choose Flask & Django for backend. We choose given Sujan’s proficiency.
* Exploration: We found out that JIRA is actually easy to use based on friends' comments and it’s good to learn some new tools for everybody.
* Exploration: we don’t have a professional way to present a UI proposal so Lucas asked Xiwen to improvise.
* Obstacle: Lucas is speaking too fast.
* Obstacle: the requirement and reporting documentations are scattered around different platforms.
* Obstacle: Lucas foresees a possible need to incorporate Docker which will create a heavy integration burden but the team is pressured into starting implementing.

**Key Decisions:**

* Reduce project scope. Instead of using machine learning algorithms to read non-readable pdfs, we use a machine learning algorithm to analyze an existing sample dataset to analyze high-value anomalies, e.g, properties with exceptionally high maintenance cost, or letting agents with higher maintenance quotes.
* Use Flask for backend as this is what Sujan is good at.
* Use Jira for project management. Boya will onboard us on Jira.

**Action Items:**

* ~~Lucas: upload sample data to GitHub for~~
  + ~~Sujan: to start working on the backend.~~
  + ~~Zihan: to start experimenting with data analysis in the way as mentioned in Key Decisions.~~
* ~~Lucas: send project docs to Sujan and Boya~~
* ~~Lucas: fill in reporting documents.~~
* ~~Xiwen: given the ratios to report as indicated in Key Decisions, draft a UI layout in any chosen format (not done).~~
* ~~Xinyu: nothing (nothing to test or integrate at the moment).~~
* ~~Boya: Make us do all our coordination work on JIRA.~~