

Minutes of the Fifth Client Meeting

Terrific Group 19

Saturday 9th March 2022

Chair	Mingen Xiao
Secretary	Jing Han
Members	Zhaotong Cui Yufeng Jiang
Apologies	None.

1 Time and Place

The fifth client meeting for the Master of Computing & Innovation Group Project was held in **Zoom** remotely, at **6pm on Friday 08th April 2022**.

2 Quorum Announcement

The Chairman announced that a quorum of the group was present and that the meeting, having been duly convened, was ready to proceed with its business.

3 Summary of Previous Meeting

Mingen Xiao briefly recapped Thursday's meeting which was largely focused on the showing technical solutions, project requirements discussion and the upcoming group milestone.

4 Group Milestone

4.1 Overview

All the team members contributed on making business case and draft plan for the upcoming group milestone 1. The team also focus on designing a form as the object of the database by applying mysql.

4.2 Detailed Presentation

Mingen Xiao and Zhaotong Cui contributed on writing the business case of the report. In particular, Zhaotong Cui mainly focus on completing the executive summary and project motivation, where Mingén Xiao focus on completing the project activities, the benefits and the goals of the tool and the further directions of the project after finishing the semester. Yufeng Jiang and Jing Han contributed on writing the draft plan for the group milestone 1. In particular, Yufeng Jiang mainly focus on completing the schedule by applying the milestone templete form, and Jing Han mainly focus on completing the roles allocation and the communication plan.

- The first week was mostly lost to the late scheduling of selecting project topics, so bythe end of the first week, only a brief meeting with self-introduction between

group members had occurred.

- In the second week the team had a brief role allocation and started focusing on technique research. The first client meeting was held which mainly introduce the details of the project and discuss about the expectations of every step of the project.
- In the third week, upon to what the team had known for the project details so far, the client showed the key point 'data lineage' for the preferred project, which is a enterprise system with lots of data, objects and tables storing in the database. The challenge is to achieve visualisation and generate links between objects. The aim is to build a web application which can load the script of the sequence of database, that could be from 100-200 script into 2. The function is to analyse the links between the tables in the database. SQL lineage could be the crucial skills for the software.
- Throughout the course of the entire project so far, all team members have a general frame of how to start the projects and make the final decision of the project topic so that the team can further going with the technique research and start learning corresponding skills. The team also had well prepared agenda in the general group meeting before the first client meeting and collect important questions that needed to ask for the client to make a clear project plan.
- In the fourth week, the team found the project has more potential details to be discussed. Each member had new questions. The team had well prepared agenda in the general group meeting before the second client meeting and collect important questions that needed to ask for the client to make a clear project plan.
- The client meeting in this week mainly talked about the details of the project and discuss about the expectations of every step of the project.
- The client demonstrated the basic functionality of the project. This project is a visual website. Users input data and use the underlying tools of the website to obtain the relationship between related data, and display it in visual graphics, allowing customers to intuitively see the relationship between stock data. The challenges faced by this project are as follows:

1. Medium data scale, hundreds of orders of magnitude.

2. Using a given analysis tool, how to create interfaces.
 3. How the tool recognizes multiple logical relationships between the input data and correctly outputs the visualization.
 4. Create a classification query function on the website, and each type of data can be output and compared through the same name.
- In order to overcome the difficulties in the project, the team should learn SQL basic syntax, T-SQL, subqueries SQL and SQL joins. Besides, learn how to use the given tool or use another ready-made mature tool.
 - Throughout the course of the entire project so far, all team members are going with the technique research and start learning corresponding skills.
 - In this week's presentation, the problem that the team firstly discussed was the complex data flow and too much data, so searching data is time-consuming. Database user needs a tool to help them to search meaningful information, track data flow, create a map of data flow. Need to get insight of the database or data information. Generate data map, dictionary of the data.
 - Then comes to solution:
 1. Parsing-transform data into useful structure.
 2. Storing useful data/information in a database and visualize it through data map and data dictionary.
 - During week 6, the whole team has detailed scheduling of lectures and briefly start planning on the milestone report.
 - The team in this week had more general group meetings than before because we all mainly focus on the business case and the draft plan, which introducing our project, team allocation and planning for the milestone 1.
 - The team also had a client meeting in this week which mainly talking about the stage of the technical solution

5 Individual Milestone Reports

5.1 Mingen Xiao

- Attended all meetings, including one client and three team meetings helping with requirements gathering, and all additional ad hoc meetings.
- Participated in everything needed in the fifth week activities.

- Contribute on taking notes of the meetings and collecting important information to write the minute.
- Contribute on the business case & draft plan.
- Play the role of organizing the work allocation for the business case & draft plan
- Mainly focus on the business case part, especially on the part of project activities, benefit of our tool, the project goal and its future directions after the semester.
- Play the role of editing the format and checking the details of the business case & draft plan to make sure the team submitted the best version to the assignment page.
- Get ready for the design of the database by following the requirement.
- Frequently communicate with the client if there are any questions in the team about the project.

5.2 Zhaotong Cui

- Attend all activities during the week
- Attending client meetings, helping to take notes and gather information
- Learning how to use the third-party data lineage tool and visualising the initial plan
- Participate in the first chapter of the full pitch presentation
- Researching the implementation of data catalogues and finding ways to parse source documents

5.3 Yufeng Jiang

- Attended all meetings, including one client and two team meetings plus all additional ad hoc meetings
- Made slides for the pitch presentation
- Asked the clients some suggestions about the presentation and the project.
- Learned knowledge about data lineage, data mapping and data dictionary.
- Took an active role in to coordinate the division of labor among group members.

5.4 Jing Han

- Attended all meetings, including one client and three team meetings helping with requirements gathering, and all additional ad hoc meetings.
- Participated in all the activities in the fourth week.
- Contribute on project pitch presentation.
- Doing further research for data lineage & python.
- Help to record the client meeting content and requirements.
- Contribute on project presentation slides, modify slides logic and content.

6 Project Administration

Jing Han did all the recordings and drafts in the meeting and post it to the group chat. Yufeng Jiang analysed what the key points recorded during the client meeting and suggested that we should communicate in a more efficient way. Zhaotong Cui suggested all the members join in the discord and post all the documents on it to help sharing ideas. Mingen Xiao declared all the members should show what he did in each week in order to follow the time schedule, and also:

- Fix github structure
- Communicate with the client

Front-end:

Activity	Team member	Planned due date
Convert csv file into json format so that the system can read the data	Zhaotong Cui	15 th April 2022

Back-end:

Activity	Team member	Planned due date
Get and display the output after extraction in the requiring format.	Yufeng Jiang	15 th April 2022

- Fix GitHub structure
- Commit code to GitHub
- Communicate with the client in Slack

7 Requirements Elicitation

Asked the client for a number of input datasets as the standard to design the form in the database, and the details of the meaning of the project requirement. Discuss with the client about the expected output and case statement.

7.1 User Requirements

- As a user, my files are usually including 100-200 scripts with the csv format.
- As a user, I want to move the nodes on the dataflow map so that I can control the shape it looks like.

7.2 System Requirements

High Priority:

- Create a database with mid-size payload to store the data.
- Create an input button that the users could upload their csv files.
- Generate code to convert csv files into json format.
- Create an external function in the front-end to make the nodes in the dataflow map moveable.

Low Priority:

- Optimize the layout of the dataflow map.
- Create a filter button in the webpage with additional functions to sort and store the data in different arrays with different properties for filtering.
- Upgrade the payload size of the software.

7.3 Browser Support

- Google Chrome

- Mozilla Firefox 3.0+
- Microsoft Internet Explorer 6.0+

7.4 Communication

- Face-to-face meeting
- Zoom meeting onshore
- Text in Slack

7.5 Tasks

- Mingen Xiao will focus on converting csv file into json format so that the system can read the data.
- Zhaotong Cui is going to extract case and join statement.
- Yufeng Jiang will modify the output of the system until matching what the client expected.
- Doing more research of how to write a wonderful milestone report
- Discuss about the roadmap such as deliverables and timelines

7.6 Lists of Tasks to do

Done:

- Briefly understand the skills that the team need to get familiar with to complete the project.
- Doing research of data lineage.
- Learn fundamental skills of python.
- Have a brief understanding of SQL and start using it.
- Understand the details in the requirement document provided by the client.
- Doing research of csv format.
- Design a form as the object of the database by applying mySQL.
- Front-end developers create webpage with basic framework.
- Successful extract the subqueries and the name of the subqueries.
- Successful extract the comments in every query.

In Progress:

- Learn how to input files from convert csv format into json format.
- Extract case and join statement.
- Standardize the output format of comment and subquery statement.
- Printing output correctly as client expected.

7.7 Glossary

Data-Lineage:

- Data Lineage
- Dataflow maps
- Meta-data
- Data extraction
- Data mining
- Data parsing
- Data catalog
- The extraction of case and join

Python:

- Tuple

SQL:

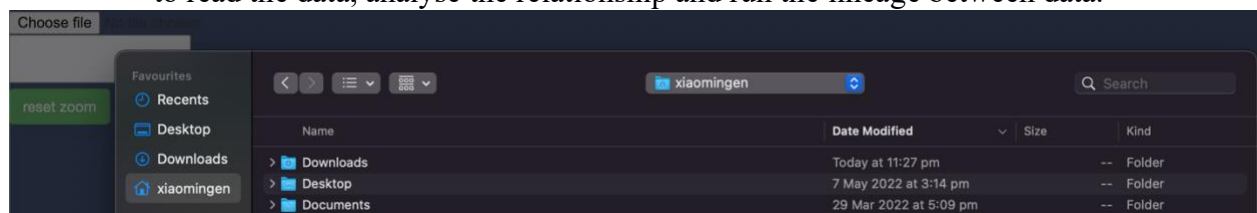
- Subquery
- Join clause

7.8 Non-functional

- Simply provide csv files as input to get the output.

7.9 Interface

- Graphical user interface to achieve visualisation of the dataflow map.
- The software needs to be highly compatibility because data lineage can be widely used for the users in different fields and generate various functions.
- The input files will be the csv format files, and convert to json format, for the system to read the data, analyse the relationship and run the lineage between data.



8 Adjournment

The next meeting is a group meeting and will be decided by the client.
The meeting closing time will be decided by the client.