Minutes of the Sixth Client Meeting

Terrific Group 19

Tuesday 3rd May 2022

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

1 Time and Place

The sixth client meeting for the Master of Computing & Innovation Group Project was held in **Zoom** remotely, at **6pm on Tuesday 3rd May 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 last week's meeting which was largely focused on the decision of the topic and technical details of the project.

4 Group Milestone

4.1 Overview

All the team members contributed on double check for the project management, showing features of what the software currently have and discuss for more possible skills to perfect the functions.

4.2 Detailed Presentation

All the team members contributed on writing the Milestone1 report. In particular, Yufeng Jiang and Zhaotong Cui mainly focus on the parts of activities, planned output and achieved output of the milestone1 report. Jing Han and Mingen Xiao contributed on writing team reflection on progress and team reflection on managing problems.

- The first week was mostly lost to the late scheduling of selecting project topics, so by the 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 an 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 from100-200 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.
- During the meetings in week 5, 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.
- In 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.
- During this week, The whole team focuses on the milestone report this week, completing the individual part of their own. And we find that what we recognized and did is a little bit different from what we planned, that means what we had known is a little bit different from the Business Case and draft.
- The team also had a client meeting in this week which mainly talking about the stage of the technical solution.
- As for the technical stage, we completed the basic framework of the front-end page for data map this week. And for the data catalog, we can now extract subqueries and comments, but have difficulties in extracting case and joins statements.

5 Individual Milestone Reports

5.1 Mingen Xiao

- Attended all meetings, including one client and two team meetings helping with requirements gathering, and all additional ad hoc meetings.
- Participated in everything needed in the sixth week activities.
- Contribute on collecting the problems of the project that the team currently facing and organizing the requirements listed by the client to complete the agenda for the week 7.
- Contribute on the first Milestone Report. Mainly focusing on the team reflection part.
- Frequently communicate with the client if there are any questions in the team about the project.
- Doing research for converting csv file into json file.

5.2 Zhaotong Cui

- Attended all meetings.
- Asked some questions about the milestone report draft
- Learn and use open source tools to visualised data flow mapping
- Completed part of the tasks of join extraction
- Learn and use the python package to convert csv files into json format files
- Took an active role in to coordinate the division of labor among group members.

5.3 Yufeng Jiang

- Attended all meetings.
- Complete the extraction of case statement, waiting for testing results from client
- Improve agenda for 4.08
- Asked the clients some questions about the milestone report of the project
- Built the basic front-end page for data flow map
- Took an active role in to coordinate the division of labor among group members.

5.4 Jing Han

- Attended all meetings.
- Participated in everything needed in the sixth week activities.
- Contribute on collecting the problems of the project before the client meeting.
- Wrote the Minutes for week6.
- Contribute on the first Milestone Report. Mainly focusing on the team reflection part.
- Doing further research for open source.

6 Project Administration

Jing Han did all the recordings and drafts in the meeting, post it to Discord, and reminded the team member to focus on the coding part and never stick on one little step for a long time.
Zhaotong Cui tried a new way to deal with the technique problem, and told some tips to others.
Mingen Xiao remind the team the deadline for finishing the Milestone report is coming, we should make some progress on it.
Yufeng Jiang gives several suggestions on how to get a better understanding about the open source.

Front-end

Activity	Team member	Planned due date
Create an information bar to show the further details related to the node such as the column and type while clicking on the node.	Mingen Xiao	11th May 2022

Back-end

Activity	Team member	Planned due date
Extract case statement	Yufeng Jiang	7th May 2022
Extract join statement	Zhaotong Cui	9th May 2022
Standardize the output format of comment and subquery statement	Jing Han	9th May 2022

7 Requirements Elicitation

Asked the client to help find a way to extract case statement. And show our new progress.

7.1 User Requirements

- As a user, I want to check the details of every node as quicker as I can.

7.2 System Requirements

High Priority:

- Create an information bar in the webpage to show the further details related to the node such as the column and type while clicking on the node. The data needs to be collected by calling the functions to get the variables from the input document.
- Standardize the output format of comment and subquery statement.

Low Priority:

- Optimize the default display structure of the nodes and edges for the dataflow map.

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

- Establish the activity for every team member.
- Check the progress of the project and manage the assignments for the following week.
- Mingen Xiao needs to create an information bar to show the further details related to the node such as the column and type while clicking on the node.
- Yufeng Jiang needs to solve the problem that cannot be passed all the tests when extract case statement.
- Zhaotong Cui needs to extract join statement.
- Jing Han needs to standardize the output format of comment and subquery statement.

7.6 Lists of Tasks to do

Done:

- 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.
- Successful extract the subqueries and the name of the subqueries.
- Successful extract the comments in every query.
- Learn how to input files from convert csv format into json format.
- Front-end developers create webpage with basic framework.

In Progress:

- Create an information bar to show the further details related to the node such as the column and type while clicking on the node.
- Fix some edge conditions for case statement extraction.
- 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
- The extraction of case and join
- Meta-data
- Data extraction
- Data mining
- Data parsing
- Dataflow maps
- Data catalog

Python:

- Tuple

SQL:

- Subquery
- Join clause

7.8 Non-functional

- Widely use in different fields and appropriate for various purposes of the users.
- Simply provide csv files as input to get the output.
- Effective portability of the system with fast performance in different environment.

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 held in **Zoom**, at **5pm on Thursday 06 May 2022**.
The meeting closed at 7pm.