Minutes of the Ninth Client Meeting

Terrific Group 19

Tuesday 31st May 2022

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

1 Time and Place

The eleventh general client meeting for the Master of Computing & Innovation Group Project was held in **Marjoribanks/126/SANTOS Lecture Theatre**, at **5pm on Tuesday 31th 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 the previous client meeting which was largely focused on the presenting a list of activities that the team achieved and asking for more functions that the product may need to generate to get better performance.

4 Group Milestone

4.1 Overview

Most of the team members contributed on the list of activities that established just after the previous client meeting. Most of the activities for the project have been completed before the client meeting in this week so that the team did a great presentation to show the performance of the product that currently have. The team also discussed with the client for more potential enhancements of the product, including the final deployment before the end of the semester.

4.2 Detailed presentation

- 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 showedthe 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 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.
- 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.
- There are more general group meetings than before in week 6 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 week 7, 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.
- In the general client meeting in week 8, Yufeng Jiang firstly displayed the output that the software currently generated in our own website to analyse with the client that whether our results are expected.
- After that, the client Roberto made a conclusion that our software structure is in the right direction but still need to improve the output because the printout information still not completely matched with what it should look like.
- Zhaotong Cui was then explaining why our output printout as the current view and made an assumption about how to achieve the correct structure of the output.
- Mingen Xiao also asked and showed a plan of how the front-end should be modified in order to printout the correct output.
- After discussing the technical solutions, Mingen Xiao showed the draft of the milestone1 report and explained the individual stages of progress that the group has completed.
- The client Roberto provided further suggestions of how to make a more appropriate milestone1 report for the progress of the project.
- Before the meetings ends, Jing Han asked a couple of more questions that we prepared about the codes and the milestone1 report.
- In week 9, Yufeng Jiang displayed the new achievements of the team, shows some new functions of the data flow map page. i.e Shows the dependence of the nodes, use Size function to change the node size according to its level.
- Yufeng Jiang display how we solved the problem that cannot be displayed all nodes when there are too many nodes.
- The client checked some cases in the search function and give some advice for improvement.
- There is a new requirement, the client want the team to implement to show the dependency of all nodes when searching for a specific node, not just show the first level relationships.

- Jing Han asked whether there is new requirement for the back end, since the whole team was blocked there. Maybe save time to initialize other new functions than just focus on one thing is a better strategy.
- The client said, in this project, the detail things should be designed and discussed by the team itself and how to implement the project aim is an open question, there is no need to wait until the client ask.
- In week 10 the meeting with the client was long and important. The meeting started on Thursday briefly displayed what activities the team did during the period before the client meeting.
- However, since the meeting time is too short, the client decided to make another meeting on Friday. During the meeting, the team displayed more details of functions of the system including introducing the step to use the functions and explaining the codes of the functions.
- The client was then suggesting for more activities that the team should complete in order to make better performance of the project.
- The team also had a brief discussion with the client about the plan to process testing for the project. In particular, the client represented that the accuracy, functionality and stability is the crucial properties of the product, which helped the team to find the appropriate directions to plan the tests.
- During the meeting, Yufeng Jiang played the role of showing how the system works and explaining the code.
- Mingen Xiao played the role of a user to show how the functions fit the requirements by using the sample input provided by the client.
- Zhaotong Cui discussed the questions that the team met with the client and asked the tools used appropriate for the new activities.
- Jing Han played the role of recording all the important information during the client meeting.
- During week 11, Yufeng Jiang present the team shows what the team achieved in last week. And mentioned a difficulty the team faced. When use the search function to show a specific node and the related nodes, the team use the recursion method, therefore the searching time is more than 30minutes.
- The client mentioned when use the search function, the accessible time complexity is 15 seconds, if the team cannot show all the related node within 15seconds, then showing the first level relationship is accessible.
- Mingen Xiao asked a question about how to make the navigation bar stable when turning into another page.
- The client gave the suggestion, exploring build the navigation bar on the existing one.
- Jing Han asked whether there are new requirements and discussed how to develop the efficiency of the whole team.
- In this meeting, the client post new requirements, the team should implement following functions in this week.

  1. Have two colors of the edges, one for one type of relationship (objects that depend on and object on which depends)

  2. Set colors of the objects according to the type of object (view, user table and store procedure)

3. Change the information that you display when you are hitting the node

4. Add a navigation bar with the following sections data flow map, cases statements, subqueries and data catalog (I′ll can explain it later)

- 5. Add a section to load the csv file.
- In the final week of the semester (Week 12), the team planned to have not only one client meeting because there will be a lot of details about the project that need to communicate with the client before completing delivery. For example, performing the functionality and discuss about the deployment to the public.
- Yufeng Jiang displayed the new functions that he achieved in between the previous client meeting till the current one. And also explained how the code works and the logic to achieve the functions.
- Mingen Xiao also shown the new functions that he achieved with the explanation of how does the code work and the logic of the code. And also asked for several outcomes of the output for the web page navigation function and discuss with the client to get the preferred performance.
- Jing Han asked a number of questions about the final project deployment and discuss with the client and check whether there are any specific requirements for the deployment.
- Zhaotong Cui had an issue on his activities, which was about the print out information of the subquery statement, and made a discussion with the client in order to get suggestions from the client to find the appropriate solutions.

5 Individual Milestone Reports

5.1 Mingen Xiao

- Contribute on the activities.
- Attended all meetings, including two client meetings and a number of team meetings and all additional ad hoc meetings, where focus on perfecting the functionality, discuss the final deployment and preparing for writing the poster and the final presentation.
- Participated in everything needed in the week 12 activities.
- Elected the chair of the team, contribute on taking notes of the meetings and collecting important information to write the minute.
- Contribute on the activity that perfecting the layout of the navigation bar as well as the dataflow map to satisfy the client’s requirements that should make the website looks more clear, easy controlling and larger view of the dataflow map.
- Contribute on the activity that only show the first-level search instead of the combination of the first level and the second-level search while searching specific information in the dataflow map.
- Continue working on the activity that showing case statement and subquery statement in new web pages.
- Contribute on the new activity that processing multiple sql entries instead of just one.
- Contribute on modifying the code style, such as putting indentation and remove useless code parts.
- Contribute on editing the README of the project.

- Doing research for achieving final deployment to public for the product.
- Communicate with the client about the project activities.
- Start working on the poster and the final presentation.

5.2 Zhaotong Cui

- Attended all meetings, including client and team meetings.
- Participated in assigning team tasks.
- Created a new web page, completed, and optimised the case statement and sub-query pages.
- Encapsulation the back-end open-source tools and integrated the front and back-end code.
- Completed the framework of the data catalogue page by using. Dash_core_components library.
- Adjusted the style and layout, beautified the upload button component.
- Tried to deploy the project using Heroku.

5.3 Yufeng Jiang

- Attended all meetings, including client and team meetings, where focus on checking new and improved features.
- Participated in assigning team tasks.
- Complete the upload function for subquery and case statement page. Improve the current code structure to fix bugs.
- Change the layout of the subquery and case statement page according to client's feedback. Try to improve the search function to increase efficiency.
- Try to improve the search function to increase efficiency. Create a new page and upload button for data catalog page.
- Improve the page layout. Create the filter function for data catalog page.

5.4 Jing Han

- Attended all meetings, including client and team meetings.
- Participated in everything needed in weekly activities.
- Contribute on collecting the problems of the project before the client meeting.
- Make notes during the client meeting, post all the client meeting records to team members.
- Wrote the Minutes and agenda.
- Wrote the Poster.
- Prepare for the final presentation.
- Test all the code on Linux system, and optimize it, in order to run the code on it smoothly.
- Design Subquery statement extraction page test case.
- Design Unit Tests for Data catalog web page.
- Try to deploy the project website on the AWS Cloud.

6 Project Administration

Continue working on some of the existing activities and the client did suggest a number of enhancements for the current functions and additional features that could

achieve before making delivery of the project:

Activity	Team member	Planned due date
Display case statements in the case statement page	Yufeng Jiang, Zhaotong Cui	03rd June 2022
Display subquery and the name of the subquery in the subquery page	Zhaotong Cui, Yufeng Jiang	04th June 2022
Display objects detailed information in the data catalog web page.	Zhaotong Cui Mingen Xiao	04th June 2022
Set to accept multiple sql entries.	Yufeng Jiang	04th June 2022
Connect different pages with a navigation bar in the webpage	Mingen Xiao	04th June 2022
Complete project deployment to the public	All team members	05th June 2022
Create a filter function for the data catalog	Yufeng Jiang, Jing Han	09th June 2022
Search for only first-level dependencies for the searching function.	Mingen Xiao, Yufeng Jiang	10th June 2022
Complete the README of the project in GitHub	Yufeng Jiang Mingen Xiao	12th June 2022

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

7 Requirements Elicitation

Present the new functions to the client and explain the logic of the code. Ask for any improvements of the product from the client and discuss about any specific requirements of the final project deployment.

7.1 User Requirements

- As a user, I would like to clearly view the lists of case statement, subquery statement and data catalog.
- As a user, I may have the input file with multiple sql scripts.
- As a user, I would like to check the data clearly, for example, viewing different information in different pages.

7.2 System Requirements

High Priority:

- Create pages for showing case statement, subquery statement.
- Accept the input file that contains multiple sql scripts and read them one by one.
- Connect different pages with a navigation bar in the webpage.

Low Priority:

- Create a page for showing data catalog.
- Separate the sql statement by various types.
- Complete the extraction of join statement.

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

- Perfect the functionality of the project.
- Establish the new activities suggested by the client.
- Decide the priority to achieve the activities.
- Allocate the person to charge the activities.
- Do some research on making a project poster.
- Get ready for making slides for the final project presentation.
- Prepare for the final project presentation.
- Mingen Xiao needs to:
    - Perfect the connection between different pages with a navigation bar stabilised on the side of the webpage.
    - Help the team member to achieve the function to display the subquery statement, especially contribute on the approach to get input files from the users.
    - Complete project deployment.
    - Modify the grammar of the project code, such as indentation and useless

code parts.

    - Edit the README.

- Yufeng Jiang needs to:

    - Debug and perfect any functions of the product, including extending the function of loading input files to aceept the product to read multiple sql scripts.

    - Help the team member to achieve the function to display the data catalog, especially contribute on the approach to get input files from the users.

    - Complete project deployment.

    - Create a filter function for the data catalog.

    - Edit the README.

- Zhaotong Cui needs to:

    - Complete the functions that read the input file, get the case statement, subquery statement and data catalog to print them out in the web pages.

    - Complete project deployment.

- Jing Han

    - Complete project deployment.

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.
- Solve the problem that cannot be displayed all nodes when there are too many nodes.
- Extract case statement.

- Standardize the output format of comment and subquery statement.
- Printing output correctly as client expected.
- Stabilize the network display of dataflow map of the product.
- Set colours of the objects according to the type of the object.
- Detailed the information that displayed in a text box when hovering over the node.
- Add a section to load the csv file.

In Progress:

- Create a webpage for showing case statement.
- Create a webpage for showing subquery statement.
- Create a webpage for showing data catalog.
- Complete the connection between different pages with a navigation bar in the webpage.
- Extend the function of loading input files to let the product to read multiple sql scripts.
- Achieve project deployment.
- Create a filter function for the data catalog page.
- Correct the searching function that for only first-level dependencies.
- Edit project README

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
- Annotation

- A synchronous generator iterator
- Callback
- Extension module
- __future__
- hash-based pyc

SQL:

- Subquery
- Join clause
- Schema

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.
- Accurate and visible data lineage analysis.
- Maintainable system.
- Easily use in different operating systems since the code structure is simple.

7.9 Interface

- Navigation in different web pages with different addresses separated by button controls in order to achieve graphical user interface to get visualisation of the dataflow map and the lists of the data information such as case statement and subquery statement.
- The software needs to be highly compatibility because data lineage can be widely used for the users in different fields and generate various functions.
- Interface for users which allow them to upload the input csv file directly in the webpage, and convert to json format so that the system can read the data.

8 Adjournment

Since it is the final week before the end of the semester, if there are any further client meeting, it will be decided by the client.
The meeting closing time will also be decided by the client.