Second report of Open Banking Data Analysis Project

Team 45 06/12/2019

Updates from the second meeting with clients:

The meeting was about the issues we had since the first report, and set tasks for the next sprint.

Restatement of the objectives of this project:

- Understand WebApp architecture and infrastructure
- Understand and implement MVC into our WebApp
- Understand noSQL database
- Learn HyperCube data structuring

The structure of a WebApp is the most important:

- Learn about MVC (Model - View - Controller) structure, model = database, View = render HTML template, Controller takes up HTML template as basic view and inserts database, performs application logic and sends it back out
- Understand the entire architecture of WebApp -> clients try to connect to the server, hit firewall, go through web server(Apache that sort). Behind the web server, the internal HTTP server links to the backend end database.
- Use python frame, probably Django
- Learn the interaction between a WebApp framework and a noSQL database
- Learn about web servers (e.g. Apache, lighttpd and so on)

About MarkLogic database:

- MarkLogic is mainly for production, it is usually really fast. The current poor performance might be due to the use of the cheapest plan on Azure and the developer version of MarkLogic. Current conditions of our system should also be checked to see if minimum requirements are met.
- Since there are not many resources about using Django with MarkLogic, MongoDB could be considered. We need to study the differences between these two databases and choose the better one.

About data and Hypercube:

- Clients will check with the Italian team who have built an opening banking platform so they'll possibly provide us with sample data
- Regarding HyperCube data examples, Japanese team may be reluctant to send as there could be some confidentiality involved in the files

- https://pythonhosted.org/cubes/ to learn about implementing hypercube using python
- Focus on putting together a WebApp framework (priority for learning objective), do not focus on data at the moment.

About forward looking insight (Open BI or Power BI):

- If such libraries exist, it is better to understand what happens inside these libraries and use them instead of doing from scratch
- Power BI and Open BI, we should use it but not make the backend frame abstract

DevOps methods used at NTTData:

- Agile methods, do in sprints
- We need to set milestones for our project

Progress so far:

NoSQL database(taken care by Yuheng):

- Connected
- Created the database on Azure virtual machine
- Set up the inner http server for MarkLogic database

- Next steps:
    - Take a look at the MongoDB, which is another NoSQL database that has more documentations ad works with python. Try to decide between using MongoDB and MarkLogic.
    - Try to get the NoSQL database work with Django

Sample Code and Django Framework (taken care by Lib Kai):

- Tried to use tink.com to retrieve sample data and realized it does not provide sample data
- Learnt how to embed PowerBI visualizations into web app
- Started learning how to use Django

Back-end and coursework logistics (Taken care of by Raghib)

- Gained a basic understanding of integrating data from database into PowerBI (either directly using MarkLogic or through Python script using library “pandas”)- client probably does not want to spend money on license of PowerBI so looking into OpenBI as alternative.
- Created basic structure for deliverable website for project, hosted on website: http://students.cs.ucl.ac.uk/2019/group45/index.html

- Next Steps:
  1. Need a more thorough understanding of the structure of web apps.
  2. Start looking at python framework and working out the design process- we need to use MVC model so since I am working on the back-end I need to understand my role in this.

Milestones for the project:

12/12/2019 - 19/12/2019: the whole team works on Django and understands the structure for the webapp Goal: to get a working framework for the webapp.

19/12/2019 - 26/12/2019: Apache with Django/ Start working on frontend/ Deploy figures using PowerBI/ OpenBI

26/12/2019 - 2/1/2020: get every part together and have the first prototype.

2/1/2020 - 16/1/2020: work on deliverables, including a video, a poster, a presentation and an individual report.