

# Microsoft Capita Team 2 / Bi-Weekly Report 2

**Date:** 28/10/2016

**Team:** Lambros Zannettos, Nathan Liu, Junwen He

## Overview

In the two weeks since the last report, we managed to gather more information on the technologies we can use for the project, and more importantly, on the nature and requirements of the project itself. We met with more people from Capita, and got a copy of the SIMS software along with installation instructions and a set of sample data.

## Meeting summary

*Wed, 19 October 2016*

Meeting with Scott Bell, from Capita. This was a meeting with both our team and team 1. We were given installation files for the SIMS software and sample databases to go with it. Scott also gave us more insight into what Capita does, and what they aspire to do with the data extracted from the SIMS databases currently held at the individual schools.

*Wed, 19 October 2016*

This was a meeting with the academic supervisor for our team project, Dr Harry Strange. We gave him updates on our progress and research so far, and we were given some valuable advice on the way forward. Since Dr Strange's field is Machine Learning, he also gave us some advice on how to tackle any Machine Learning aspects of the project which may arise.

*Thu, 20 October 2016*

Following the meetings on the 19<sup>th</sup> of October, our team had a brief meeting to discuss the new information we got and to assign roles for the current work we have to do. We started to do research on the Azure Machine Learning platform since it will mostly likely be the machine learning platform of choice.

*Mon, 24 October 2016*

This was a meeting for us to bring each other up to date with our progress with the tasks we set out to complete. We also helped each other finalize the installation of SIMS and get the hang of how to query the database to start seeing some data and playing around with some ideas.

*Mon, 28 October 2016*

Over the past few days, we have all been researching different ideas regarding the functionality we would like to develop, and this meeting was about sharing what we have found with each other.

## Tasks Completed

1. Had meeting with project mentor at Capita, Scott Bell.
2. Met with our academic supervisor.
3. Assigned roles for the next couple of weeks.
4. Got files for the SIMS software and the sample database.
5. Installed SIMS on VMs.
6. Started learning to use the tools provided by Microsoft.

## Problems to be resolved

1. Get SIMS and it's accompanying database working on our Windows/Mac machines (natively or in a VM) perfectly. The installation is not that straight forward but Scott Bell at Capita has been very helpful and we are hoping to be up and running soon.

2. We have installed Windows VMs onto team mates using Linux and Mac OS operating systems however the VMs are running very slow. We are hoping to obtain Windows machines from Microsoft in the upcoming week.

### **Plan for next two weeks**

- Get in touch with the IoE and put a team together.

This week, our client will have a meeting with professor Richard Noss from the IoE to start talks on how the IoE will collaborate with Capita and our team on this project. Following this initial meeting, and after the formation of a team at the IoE, we are hoping to start establishing a clearer set of goals for the project to aim towards.

- Put the project website up and start populating the blog.

We have already started putting together some material for the website offline, but within the next few days we need to decide on what tools we are going to use to build the website and start putting it online.

- Acquire Windows machines from Microsoft

We are not yet 100% convinced that we will need to borrow Windows machines, as we have been able to get SIMS up and running (almost perfectly) using VMs. Depending on how stable the system is in the next few days, we will decide whether we need to get in touch with Microsoft about borrowing machines.

- Further explore MS Azure Cloud Services/ML Studio

Azure Cloud Services offers a huge variety of tools, including the Machine Learning Studio which will potentially be useful for our project. We have put together a playlist of relevant educational courses on Lynda, and we are planning to finish watching these within the next few days. This will give us a good idea and understanding of the available tools and possibly help inspire the next steps of the project.

- Further research NLP conversion

One of our ideas is to create an interface whereby natural language questions (e.g. *How many pupils at school X are under 15 years old?*) will be converted to SQL queries.

### **Individual reports**

*Lambros Zannettos:*

My role for the next step of the project was to chase up the collaboration with the IoC, and take any necessary steps to establish the connection. It turned out that our client will make the first contact, and the team at the IoC will hopefully start being put together this week. I also helped setup the VMs we used to install SIMS, and put together a playlist of relevant educational material on Lynda which we are now going through to better understand the tools available to us (Azure/ML Studio) and other Data Science basics. I have also been chasing up on some ideas for the functionality, like Natural Language Interfaces to Databases, OpenNLP and Kueri.

*Nathan Liu:*

My role for the next step of the project is to get SIMS up and running and connected to the sample database, and help the rest of the team to set it up on their computers as well. I ran into problems with the set up with difficulties in authentication. I have had a quick look at tutorials on the Azure Cloud Machine Learning Platform. I also watched some introductory online lectures on FTP and Internet of Things and uploaded a few sample files onto our

group's directory on the UCL CS server using FileZilla (needed for the website later on in the project).

*Junwen He:*

My role for the next step of the project is to start building the website for the project and upload it to the hosting space provided by UCL. I followed the guide to connect my PC with the group directory in UCL CS server using FileZilla, then I tested it by uploading a test website and it did work. And on the following weeks, I'm going to build our website with some basic tools and record what we have done on the website, and keep it updated.