Log 1

25 September 2016

**Objectives**:

* Seek articles/papers on open domain QA systems
* Find example applications
* Look for Open Source projects
* Gain an understanding of the report requirement

**Activities**:

The example applications and open source projects I have found have proven to be very useful.

Looking at these has given me more of an understanding on what it is I will be creating. The open source projects are likely to be quite fruitful in the development stages, as it is likely I will be using some of their elements. Take DBPedia, for example, it looks to be a great knowledgebase to retrieve information from.

Reading the papers and articles on this subject have given me more of a technical understanding. An understanding on the underlying principles of open domain question and answers systems, and also in general areas of AI. They will be of great use in the initial report, as well as in other reports, and also in development of my own system.

**Reflections**:

Until the initial report is completed, I will be reading and researching more articles and papers. I will also be on the look out for similar systems, which I will be testing to see their faults and advantages. It's likely the commercial applications will yield better results than the research based applications.

Log 2

13 October 2016

**Objectives:**

* Continue reading articles and reports
* Continue testing similar QA systems
* Continue writing the report

**Activities:**

After reading the brief on what is needed for the report, I have concluded that I will have five main sections.

The introduction, a project goals section, a development section, a planning section, and a research section.

The introduction will be a short summary of what I intend to do during the project duration, but more specifically what an Open Domain Question and Answer system really is.

The project goals section will outline my intended goals for the project. For example, which features I plan to implement will be core, and which features will be developed if time and development allow.

The development section will give an overview on how I intend to develop the system. This section will also feature some of the open source projects mentioned in the previous log book. These open source projects will be considered in depth in the initial report. Also in this section will be details on how I intend to develop certain aspects of the system, for example the natural language parser.

Within the planning section will be details on the methodology type I will be using, and why I decided to use this methodology. The section will also show details on timeframes, in the form of a Gantt chart and an explanation of the chart.

Finally, the research section is where I will talk about what I have learned and read during my research. Within this section I will also show sample open domain question and answer applications to show what it is that I want to achieve.

**Reflections:**

Writing this initial report has been important for two main reason. Firstly it has given me a greater understanding on the topic at hand. Secondly, through my research I have come across some open source projects, which could potentially be used within my own system.

Having only written what I would consider little throughout my university studies this report has proven to be more of a challenge than first anticipated.

Log 3

22 October 2016

**Goals:**

* Start work on the project itself
* Continue testing and researching similar systems
* Focus on other modules too

**Activities:**

This week I have not made much significant progress as my time was focused more on coursework for other modules. However, I was able to draw up rough diagrams on how the system will be developed. These are likely to be adjusted over the near future.

**Reflections:**

After having completed the report, my initial feelings have not changed. I am still in the mind-set of it being more difficult than first anticipated. However, finishing this report will allow for smoother project development, and an slightly easier interim report.

What I found crucial in writing this report was time management. It would have been easier for me to complete if I had managed my time better. This has shown me that I need to be able to functionally manage my time for not only the next report, but also for these logbooks, and especially for the project development. Time management for the project as a whole is crucial, as I will not only be working on this but also coursework for other modules.

One of my module choices for this year is Natural Language Engineering. This module will help in developing the natural language part of my system. In this module we will be using an open source piece of software called NLTK (Natural Language Toolkit). NLTK was mentioned in my initial report as a potential piece of software I will be using. Using it in another module will allow me to gain more in-depth knowledge of it, as I will be able to ask the lecturer for assistance. The module as a whole will also aid in my knowledge of natural language engineering, which I will be able to transfer into the project work.

Log 4

Thursday, 10 November 2016

**Objectives**:

* Figure out how to use DBPedia
* Figure out how to use SPARQL syntax to query DBPedia
* Learn about DBPedia's ontology
* Figure out how to parse natural language into a SPARQL query

**Activities**:

Currently I am unable to download a knowledgebase and run it locally, as my machine does not have sufficient storage. Instead of running it through an external hard drive I found a more elegant solution would be to pull directly from the DBPedia servers. This can be done using the RDF query language SPARQL.

I needed to learn a bit about DBPedia and SPARQL before I start the natural language processing.

There would be no point in starting the natural language processing without knowing what the natural language needed to be parsed into.

So far I am able to pull certain information from a DBPedia page. However, being able to manipulate this data successfully is something I need to work on.

**Reflections:**

This week I was unable to parse natural language into a SPARQL query as I am still getting to grips with SPARQL and DBPedia. I will also need to continue learning the ontology used in DBPedia, as this is crucial to retrieving the correct information.

For next week I will need to again focus on pulling information from DBPedia, and setup some sort of format to inject variables into a query. It might be better to have a selection of preformatted queries to inject variables into, as opposed to generating a new query each time. This is likely to save development time and make development slightly easier. However, this could also hinder the final product, and restrict the systems information retrieval.

Log 5

Wednesday, 23 November 2016

15:36

**Objectives**:

* Continue learning how to use DBPedia
* Continue learning SPARQL
* Continue develop the system
* Start to implement simple natural language processing
* Start outlining interim report

**Activities**:

Monday

I read documentation on DBPedia and tested the system with new prefixes and data types.

Tuesday

Today I took a break from development and focused on the interim report. I made a brief outline of what I want to discuss in the report, and also read up on the requirements for the report.

Thursday

I continued working on enhancing my knowledge on SPARQL and the structure of DBPedia.

Friday

Had a meeting with John. We discussed current progress, future plans, masters applications, and the interim report.

**Reflections:**

This week ended up being more research focused than development focused. With the interim report being due quite soon I feel like the next few weeks will be of a similar structure. While development is crucial, it's more important for me to research the tools I am using in-depth before implementation as they are quite complex.

The meeting with John proved to be very productive as we were able to outline key areas of focus for the report. Next week I will be touching up my diagrams and requirements to send to John for feedback, to make sure I am on the right track.

Log 6

Monday, 28 November 2016

13:09

**Objectives:**

* Work on interim report
* Find and Read articles and papers on QA systems, and tools being used
* Continue work on accessing specific elements on DBPedia

**Activities:**

Monday

I researched sources for the literature review section of interim report. Alongside this I created an outline structure of the interim report, and researched methods on writing a literature review.

Tuesday

Gathered more sources for the literature review. Created a draft requirements list and system architecture diagram.

Emailed John with a list of functional requirements, non-functional requirements, and a system architecture diagram for the interim report's SRS document.

Wednesday

Read Johns reply, and replied with more questions to fine tune this part of the document. After taking what John said into account I began to work on a better set of requirements and system architecture diagram.

Friday

Researched more articles and did some light work on the DBPedia information retrieval aspect.

**Reflections:**

This weeks work was a good start towards the interim report. A lot more work needs to be done and with a solid foundation I can expect to get a 2:1 grade for the report. I need to focus on laying the foundations for the SRS document, a crucial part of the waterfall methodology.

Log 7

5/12/2016

**Objectives**:

* Work on the SRS document
* Gather and read more sources
* Create relevant diagrams
* Create an actual structure for the report
* Start writing the literature review
* Create a mock GUI

**Activities**:

Monday

Using material on the CE301 Moodle page, and my initial report as a basis, I began to outline the initial draft of the interim report. I found that it would be best to have four main sections, and subsections under each of those. The main sections will be:

|  |  |
| --- | --- |
| 1.0 | Introduction |
| 2.0 | Project Management |
| 3.0 | System Requirement Specification |
| 4.0 | Literature Review |

Tuesday

Today I started writing the introduction in which I gave a brief outline of what to expect in the report. The introduction also includes details on the topic, and reminder of what the project is.

Wednesday

Designed a mock GUI for the system using inspiration from similar systems. Started the SRS document by working on the outline and writing the foundations of some on the subsections.

Thursday

Carried on with outlining the report, however in more detail than on Monday. I also continued to find and read upon more sources, specifically on the semantic web and DBPedia documentation.

Friday

Read more sources, specifically on natural language and information retrieval. Added more structure to the report and started writing the project management section, using the initial report, logs, and my own notes as sources of information.

**Reflections:**

With the report due on the 16th I feel as though I need to find more sources and read them in depth. The current layout of the report is subject to change as more sections are completed. This week and next will be entirely report and SRS focused. This is best as once the report and SRS document is completed the development will have a greater sense of direction.

Log 8

12/12/16

**Objectives**:

* Find and Read more sources
* Finalise the structure of the report
* Complete a draft
* Create more diagrams
* Finalise requirements
* Finalise and submit the report

**Activities**:

Monday

Continued to find and read more sources, and read more on current sources. Focus was on SPARQL and similar QA systems. Also worked on the SRS by expanding on last weeks work. The purpose, scope, performance requirements, and system requirements sections are completed. The project management section was also finalised, and the outline for the literature review.

Tuesday

Today I finished the SRS document adding the diagrams and the design sections. The UI mock-up was also completed today. I also read through and finalised the sources for the literature review, which I have started to write. The literature review will likely be completed by Wednesday evening, or Thursday at the latest.

Wednesday

The main focus for today was to complete the literature review. Initially I had a one page introduction, which I found to be unnecessary. Part of the introduction would be better suited for the literature review, so after some adjustments I used this to start the literature review.

Thursday

With the bulk of the report finished I spent Thursday finishing sections across the entire report, tweaking certain aspects, and editing the report for final submission. Submitted the report.

**Reflections:**

I am quite content with the quality of my Interim Report and I am expecting a 2:1 grade. The SRS document and updated project management model in the report will provide guidance and allow for greater control over the rest of the project.