

## Supervisor meeting #2

### Agenda:

Overview of progress to date, including write-up

- Literature review completed
- EDA, statistical analysis and first ontology (POC) completed
- Main ontology almost completed; currently loading dummy data for testing

Discuss report structure:

- Introduction
- Background (literature review)
- Ethical and Professional Considerations (light because of no primary research)
- Methodology
  - EDA in Python on NIJ data to form hypotheses
  - Statistical analysis in R on NIJ data to test hypotheses
  - Build NIJ PoC ontology with real (public, anonymised) data
  - Build MOJ ontology using MoJ Metadata and dummy data
- Evaluation
  - Evaluation of EDS, statistical analysis and both ontologies
  - Questions:
    - I found a performance issue with the ontology build so I changed the design which addressed it, but created a limitation. So I have two designs (very small difference), but I think it's worth calling out why it was changed. Should that be in the implementation section or evaluation?
    - Do I need to include written tests in an appendix or is it sufficient to just describe how I tested it?
  - Learning
  - Conclusions

Confirm report presentation. The guide says Times New Roman 12 point with 1.5 spacing. That's what I'm using and fine to continue, but Times New Roman is very dated and would prefer to use Arial if permitted.

Further to my initial agenda, I've written a few specific questions that I'd like to ask once we've reviewed the overall status of the dissertation (one is actually a suggested answer to a question that I already raised in the agenda):

1. I found a performance issue which meant a small but critical design change. I wasn't sure whether to put that in the implementation or evaluation section so I've put it into implementation and left evaluation to evaluate the final implementation.
2. Do I need to write test scripts and include them in an appendix, or is it enough to describe how it was tested?

3. Do I need to include the Python EDA and R statistical analysis in appendices (they don't paste well into a document) or just provide a link to them on Github?
4. I'd like to include a recommendation for further research to enhance Protégé and other ontology tools to be able to store metadata alongside data to overcome a limitation with my ontology design. Is that a step too far and unrealistic for further research?
5. The learning section is described as "Personal reflection that involves an in-depth and critical analysis of what you have learned, as well as skills gained/improved during the project". Should it therefore be written in the first person, unlike the rest of the dissertation?

### **Summary of the meeting:**

The dissertation to date was discussed, including the reasons for having two ontologies; one to test the concept with real data and the second to build a useable ontology for England and Wales with only a small sample of dummy data.

The need to present test data was discussed. I have already generated test data so I need to ensure that is included in the dissertation (appendix), including how and why it was created.

The basis for setting the correlation threshold at 0.1 was discussed. At the moment there is no reference to why 0.1 was selected, or alpha 0.05, so these need to be justified with literature.

The specific question that I had were addressed as follows:

Q1. Do I need to write test scripts and include them in an appendix, or is it enough to describe how it was tested?

A1. Screen shots of outputs is fine to show that testing was done.

Q2. Do I need to include the Python EDA and R statistical analysis in appendices or just provide a link to them on GitHub?

A2. Yes, it's best to include them if possible with the GitHub link in addition, but it's not worth spending hours on. Perhaps take screen shots of the first few pages to evidence that they exist.

Q3. I'd like to include a recommendation for further research to enhance Protégé and other ontology tools to be able to store metadata alongside data to overcome a limitation with my ontology design. Is that a step too far and unrealistic for further research?

A3. Yes, this is OK

Q4. The learning section is described as "Personal reflection that involves an in-depth and critical analysis of what you have learned, as well as skills

gained/improved during the project". Should it therefore be written in the first person, unlike the rest of the dissertation?

A4. Everything should be 3rd person, including the personal reflection.

Q5. The guide says the report should be Times New Roman 12 but that font is very dated. Can Arial be used instead?

A5. It's not a hard rule so Arial is fine

Q6. What should go on the cover sheet? Is there a template or example I can use?

A6. It's up to me what to put on. It should include things like: Title, author, supervisors, word count.

We agreed to meet again in approximately one month, but were unable to secure a date whilst on the call. We agreed that I would email a draft of the dissertation in around two weeks, and then arrange to meet in approximately a further two weeks.