

Thematic analysis of TAIC's body of Knowledge Proposal

James Thompson - 22nd November 2023

Introduction

As with many projects this one really has no limit. An endless list of improvements, upgrades and developments could be thought up. But with a deadline just over a month away things will have to be cut to get a proof of concept. This proposal has been prepared in response to [PRJ_RES_Thematic_Analysis_RFP_2023.pdf](#) and discussion with Chris and Ingrid.

As verification and trust in the model engine and its output is paramount that is where the focus of time has been spent. There are quality of life improvements to both the web app and surrounding infrastructure that would make it closer to being able to be rolled out. Depending on the speed and success of the first couple of weeks will determine how much of that work can be completed.

Things to be done

To help make it clear what I plan / think should be done here is my “translated” things to do. These are more technical and specific than the requirements listed in the output section. I have also added rough estimates for how long these things will take to be completed.

Item list

1. Add quote extraction and references to act as evidence with output. (3 days)

There will be two sections where evidence needs to be provided. Firstly in the safety theme summary for each report and secondly in the evidence for each predefined theme weighting.

This will involve updating the prompts and model interactions to get it to actually provide references from the reports. I will then have to create a reference checker to make sure that the stated references are correct.

2. Add in multimodal support (Optional 1 day)

Currently the engine only works for marine accidents. Adding in support for rail and aviation requires additional support in the engine for multiple modes. However it should not require a fundamental change to how the engine functions.

3. Extract metadata from the reports (1 day)

It is mentioned that hubstream provides metadata about the newer reports. However currently legacy reports do not have any of this data. It would be best to add this into the engine output for each report. It would allow better searching and even importing into hubstream.

4. Upgrade the validation set (1 day of work for me but may be a couple of days for the expert to actually create the set)

Normal statistical and machine learning techniques would use a training set and a test set so that we know how good the model is working. In our case there is no training to be done therefore we can only use a validation set. Currently there is a validation set created by me which shows the example outputs and then compares the model output to the examples. This gives us a validation metric to know how close it performs to what we would expect.

The validation set needs to update to allow for the metadata and the quote extraction. Furthermore it would be good to upgrade the validation set so that it is an expert's opinion as opposed to the average human (me).

5. Upgrade the searcher (2 days)

Currently the web app has a simple keyword search. This could be upgraded to add more search options. It would be good if it had an experience that is somewhat meeting what is described in paragraph 21.

- Safety themes
- Keywords
- Meta data (Boatsize, fatalities etc)

6. Data exportation (1 day for simple data exportation)

I can see two cases where data exportation would be needed.

Firstly exporting the engine database into a common data format as described in paragraph 24. This can simply be a dump of all the data in the database converted into a tabular format.

Secondly a user wants to have a “take home” summary of their search results. This would be an investigator wanting to make some search then get a summary report which shows the output of the engine. This would be a more permanent and archival search result than the interactive experience that would be found on the webapp.

7. Add administration dashboard to the webapp (2 days)

The webapp is currently just a searcher. It would also be good if the project and engine can be managed from the webapp directly. This would help the web app solve point two in “be user-friendly” paragraph 24.

This dashboard could have the configuration for the model parameters. It would also allow the user to run the engine with fresh data. It would add a GUI to the current command line application.

8. Upgrade the web app to use a database (2 days)

Currently the webapp uses a simple file structure system uploaded with the website code to the server. This was done due to having no time at its creation to worry about databases. However this is restricting and will prevent what can be done moving forwards.

It would be better to use a database which allows cleaner data management. It won't make any obvious changes to the user but will remove the technical debt that has been taken onboard.

Excluded

Due to the time rush there does not seem to be much time available for more "research" styled efforts.

This means that a few things are not currently included in the proposal. But depending on how things go could be included.

- Thorough analysis of differences between pre-built model types. This would include making some tests that would let me compare the output of gpt 3.5 and gpt 4.0 as well as LLMs offered by other companies.
- Testing out finetuning of a model. Currently the model is being used "off the shelf". However openAI has opened up access to fine-tune the model to what you want it to do. This would mean that we could feed all of the reports as training data to the LLM and then talk with the model directly. However the effectiveness of this will take time to develop and test.

Timeline

Below are rough estimates on roughly how the project will develop. As I intend on spending 20 hours a week this translates to about 3 days of work a week. There are 5 weeks from this week till the deadline. If there is flexibility in work hours I would be able to work more hours to keep this timeline on track.

Week 1

- Organise project specifics, contract organisation
- Organise the creation of the validation set

Week 2

- Add in the quote extractions and reference
- Integrate the new validation set

Week 3

- Add in extraction of metadata
- Upgrade the searcher so that it can use all of this new metadata

Week 4

- Upgrade the engine and webapp to use a database
- Add in data exportation from the database through the webapp

Week 5

- Add in admin dashboard that allows user to run the analysis from the webapp
- Update documentation
- Prepare for presentation

Request for clarification

I hope that I have correctly interpreted the project guidelines. However there were some areas where I wanted to know a bit more information that would help with executing the project.

Presentation to TAIC's Senior Leadership Team or Knowledge Transfer System project board.

I just wanted to know a bit more about this as depending scope of the presentation I

- Was the presentation expected to take place prior to christmas?
- What was my role in this presentation and would it be some form of powerpoint presentation?

Multi-modal support

Currently the engine only supports Marine accidents. This is just by chance and there are not any large technical problems that prevent other reports being included.

In the document it is implied that all three transport modes are supported. Was this something that you wanted to do?

Predefined themes

In paragraph 24 it mentions "identify predefined themes in the TAIC's published inquiry reports". By predefined do you mean generated by a subject expert? If so, when would these be able to be provided by?

Conclusion

I believe that a presentable engine output and viewer is possible by the deadline of the 20th. However with the various quality of life improvements required it may be pushing it to have it all done in the 90 hour time frame. To help facilitate all the items being completed I would be able to work more than 20 hours a week.