**Interim Report**

**Candidate Number** – 164547

**Supervisor** – Dr Julie Weeds

**Working Title** – Fact-Checking Journalism

**Introduction**

The purpose of my project is to provide a tool to journalists that can be used to fact check a particular aspect of their article or document against an external database.

Rather than scanning for all facts within a document - which may prove difficult and inaccurate – I’ll focus my tool on fact-checking a specific area of interest about named entities. In particular, I aim for journalists to be able to fact check if their claims about named entities working for a specified organisation are true. Even though I’ve narrowed down my aim to fact checking organisational roles of named entities, the hope is that beyond this final year project, it could be extended to fact checking other aspects of named entities as well.

It should be noted that my focus will be on providing a tool that journalists can manually use to check their facts as opposed to creating an automated service. I intend to do this by providing some sort of link to the external database concerning the fact in question. However, automated fact checking is something I wish to explore as well and is therefore an extension of my project.

Since ‘fake news’ is such a prominent topic currently, I’m motivated to create this tool which can help journalists mitigate false information in their article before being published. As an extension, I would also like to provide this software to readers, so they can fact check articles themselves. However, time constraints will have to be taken into consideration.

**Relevance**

My project is relevant to my degree course as it will display a variety of skills that I’ve learnt so far and will hope to improve on. The focus is to incorporate Natural Language Processing (NLP) and Machine Learning (ML) techniques to perform the fact-checking scan.

Additionally, I’ll be making use of the skills I’ve learnt in Python to code the back end and Java to code the front end. Since I’ve never used PyCharm before, that’ll be a useful skill to add to my repertoire along the way as it’s a commonly used IDE in industry. Finally, some of the project management skills I learnt during the Software Engineering module in my second year will underscore the project.

**Related Work**

The structure of a fact checking task can be best summarised by four steps that are laid out in a paper by Andreas Vlachos [1]:

1. Extract statements from an article/document to be fact checked.
2. Construct appropriate questions/queries about the statements for the external database to answer.
3. Obtain answers to those queries from the external database.
4. Extension; if constructing an automated fact checking service, reach a verdict using that answer.

Statement extraction (the first step) is a well-explored area of NLP and is really where a lot of my attention will be focused.

Extensive research has already been carried out in the field of automated fact checking (AFC). I’ve found that even in AFC, however, human supervision is currently still a necessity. Lucas Graves [2] points out that “Both researchers and practitioners agree that the real promise of AFC technologies for now lies in tools to assist fact-checkers to identify and investigate claims, and to deliver their conclusions as effectively as possible.”

Naeemul Hassan [3] lays out four factors that contribute towards the “Holy Grail” of computer-based fact checking systems:

**“Fully automated**: It checks facts without human intervention. It takes as input the video/audio signals and texts of a political discourse and returns factual claims and a truthness rating for each claim (e.g., the Truth-O-Meter by PolitiFact).

**Instant**: It immediately reaches conclusions and returns results after claims are made, without noticeable delays.

**Accurate**: It is equally or more accurate than any human fact-checker.

**Accountable**: It self-documents its data sources and analysis, and makes the process of each fact-check transparent. This process can then be independently veriﬁed, critiqued, improved, and even extended to other situations.”

Clearly, these characteristics are unattainable given my time frame. It’s therefore clear to me that my primary aim should be to provide a tool that instead assists human fact checkers as quickly, accurately and effectively as possible, instead of focusing on the automated aspect.

A lot of fact checking software currently available focuses on so called ‘fake news’ in politics. One such piece of software is called ClaimBuster, a tool that can identify claims within a text that are most likely to be false and therefore need checking. It does this by providing the user with a quantitative measure about how check worthy the text is. Hassan [4], the co-founder of ClaimBuster, explains “Given a sentence, ClaimBuster gives it a score between 0.0 and 1.0. The higher the score, the more likely the sentence contains check-worthy factual claims. The lower the score, the more non-factual, subjective and opinionated the sentence is.”

An example of what ClaimBuster does is illustrated below. I’ve created a dummy piece of text, emulating something that may be said by a politician and submitted it to the software to observe the kind of sentences it thinks are check-worthy.

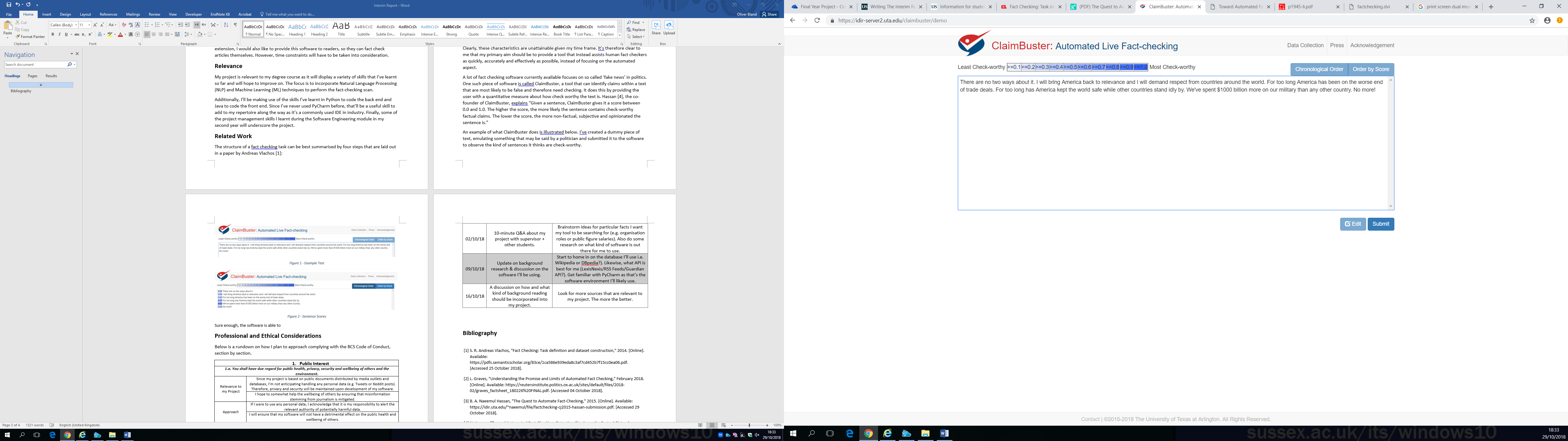


Figure 1 - Example Text

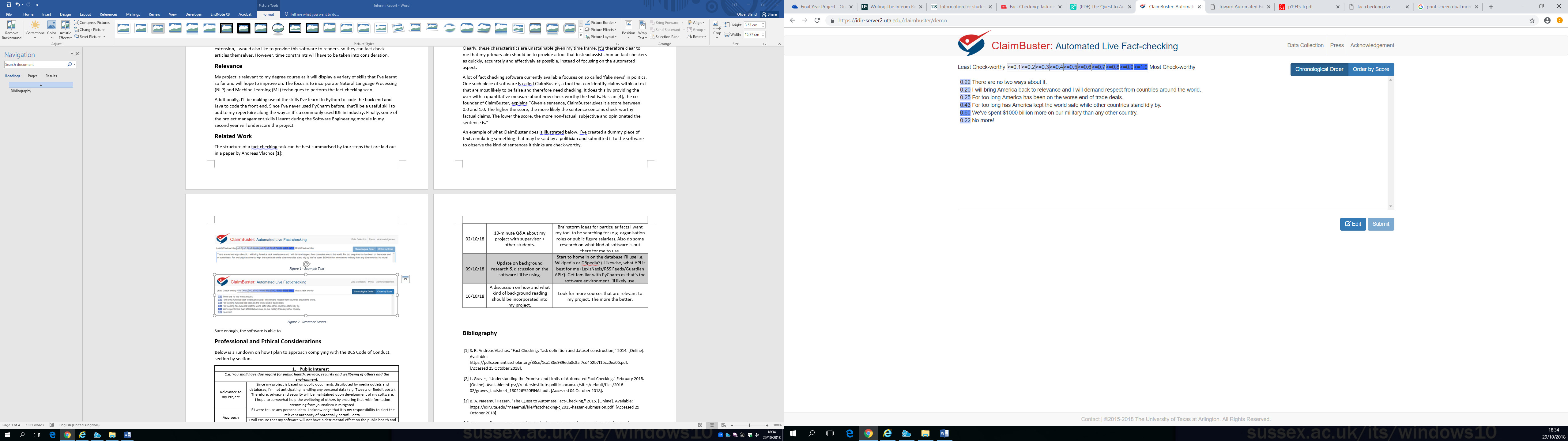


Figure 2 - Sentence Scores

Sure enough, the software assigns a higher score to the sentence that is worth fact checking than any of the other sentences. We can therefore see that at least when it comes to political facts, the software is very helpful in fact checking. However, I hope to provide a product that’s different in the following ways:

1. I’d like to provide a product that specialises in one, different domain – organisational roles. Submitting the sentence “We need only look at the success that Bill Gates has had as CEO of Facebook to see that the field is extremely prosperous” yields a check-worthy score of 0.28. Yet this sentence clearly contains a false fact when it comes to the organisational role of Bill Gates. I want to provide a tool that allows the journalist/user to notice a fact like this and avoid misinformation.
2. Instead of scoring all sentences and providing that information to the user, I’d like to build a tool which scans a document and only returns sentences which it thinks needs fact checking.

**Professional and Ethical Considerations**

Below is a rundown on how I plan to approach complying with the BCS Code of Conduct, section by section.

|  |  |
| --- | --- |
| 1. **Public Interest** | |
| ***1.a. You shall have due regard for public health, privacy, security and wellbeing of others and the environment.*** | |
| Relevance to my Project | Since my project is based on public documents distributed by media outlets and databases, I’m not anticipating handling any personal data (e.g. Tweets or Reddit posts). Therefore, privacy and security will be maintained upon development of my software. |
| I hope to somewhat help the wellbeing of others by ensuring that misinformation stemming from journalism is mitigated. |
| Approach | If I were to use any personal data, I acknowledge that it is my responsibility to alert the relevant authority of potentially harmful data. |
| I will ensure that my software will not have a detrimental effect on the public health and wellbeing of others. |
| ***1.b. You shall have due regard for the legitimate rights of Third Parties.*** | |
| Relevance to my Project | Potential third parties of my software include any person who does not directly benefit off my fact-checking tool (i.e. Non-journalists). |
| Approach | A point of emphasis of mine is to always keep in mind privacy regulations of third parties. |

**Interim Log of Supervisor Meetings**

|  |  |  |
| --- | --- | --- |
| Date | Discussion Topic | Future Actions |
| 26/09/18 | Initial discussion about what my project would entail & first steps. | Background research; particularly research articles in similar areas as the domain I’ll be working in. |
| 02/10/18 | 10-minute Q&A about my project with supervisor + other students. | Brainstorm ideas for particular facts I want my tool to be searching for (e.g. organisation roles or public figure salaries). Also do some research on what kind of software is out there for me to use. |
| 09/10/18 | Update on background research & discussion on the software I’ll be using. | Start to home in on the database I’ll use i.e. Wikipedia or DBpedia?). Likewise, what API is best for me (LexisNexis/RSS Feeds/Guardian API?). Get familiar with PyCharm as that’s the software environment I’ll likely use. |
| 16/10/18 | A discussion on how and what kind of background reading should be incorporated into my project. | Look for more sources that are relevant to my project. The more the better. |

# **Bibliography**

|  |  |
| --- | --- |
| [1] | S. R. Andreas Vlachos, "Fact Checking: Task definition and dataset construction," 2014. [Online]. Available: https://pdfs.semanticscholar.org/83ce/1ca586e939eda8c3af7cd452b7f15cc0ea06.pdf. [Accessed 25 October 2018]. |
| [2] | L. Graves, "Understanding the Promise and Limits of Automated Fact Checking," February 2018. [Online]. Available: https://reutersinstitute.politics.ox.ac.uk/sites/default/files/2018-02/graves\_factsheet\_180226%20FINAL.pdf. [Accessed 04 October 2018]. |
| [3] | B. A. Naeemul Hassan, "The Quest to Automate Fact-Checking," 2015. [Online]. Available: https://idir.uta.edu/~naeemul/file/factchecking-cj2015-hassan-submission.pdf. [Accessed 29 October 2018]. |
| [4] | N. Hassan, "Toward Automated Fact-Checking: Detecting Check-worthy Factual Claims by ClaimBuster," 2017. [Online]. Available: http://ranger.uta.edu/~cli/pubs/2017/claimbuster-kdd17-hassan.pdf. [Accessed 07 October 2018]. |