University of Sheffield

# William Briggs Report


William Briggs

*Supervisor:* Dr Mark Stevenson

*Panel:* Dr Andreas Vlachos, Professor Richard Clayton

Department of Computer Science

March 23, 2018

## Declaration

All sentences or passages quoted in this report from other people's work have been specifically acknowledged by clear cross-referencing to author, work and page(s). Any illustrations that are not the work of the author of this report have been used with the explicit permission of the originator and are specifically acknowledged. I understand that failure to do this amounts to plagiarism and will be considered grounds for failure in this project and the degree examination as a whole.

Name: _____

Signature: _____

Date: _____

## Abstract

Something

# Contents

# List of Figures

# List of Tables

# Chapter 1

# Introduction

Medical literature poses interesting challenges for Natural Language Processing (NLP) researchers. The sheer volume of medical data makes it difficult for humans to process efficiently. One key task is the creation of systematic reviews. Systematic reviews are transparent reviews that aim to pull together and critically analyse relevant literature to a topical question. The process of creating a systematic review is rigorous and time consuming with varying degrees of complexity in-between steps. This report will look at the existing work done so far on using NLP as part of the systematic review process as well as the novel work by myself.

## 1.1 Steps of a Systematic Review

It is useful for us to break down the steps involved in creating a systematic review into subtasks. This way we can observe what techniques can be applied during the relevant sub tasks to improve the efficiency of the process. The following definitions are derived task simplifactions from the cochrane tutorial on systematic reviews: [8].

1. Question definition.
2. Relevant literature search.
3. Data Filtering.
4. Data Extraction.
5. Analysis and Data combination.

### 1.1.1 Question Definition

One of the best known techniques for formulating a systematic review question is known as the PICO strategy [9]. This technique focuses on exposing 4 pieces of information in

the systematic review question: patient population, intervention or exposure, comparison or control and outcome.

Example: (credit goes to [9])

"Is animal-assisted therapy more effective than music therapy in managing aggressive behaviour in elderly people with dementia?"

| | |
|---|---|
| P | elderly patients with dementia |
| I | animal-assisted therapy |
| C | music therapy |
| O | aggressive behaviour |

A potential point of interest would be attempting to generate these questions automatically given some literature context.

### 1.1.2 Relevant literature search

After formulating a question, systematic reviews need to search for the relevant literature that surrounds this question.

Large medical database-such as pubmed contain relevant studies that can be used to create the review. These databases are typically very large and require concise queries to efficiently retrieve data.

Naturally this can be modelled as an information retrieval problem. We have a large number of documents and we wish to retrieve the most relevant ones. One task for the 2017 CLEF conference was to produce a ranking of the most releavent documents for topics [6]. Many techniques have been proposed for ranking of relevant documents, with varying degrees of performance [2] [4] [7].

An important aspect of the relevant literature search step is the construction of the query. Query creators often apply filters (also known as hedges) to increase the effectiveness or/and the efficiency of the searching. Two key attributes for the query are the precision and the sensitivity (aka recall). By including synonymous phrases e.g: quality adjusted life or quality of well-being or disability adjusted life the sensitivity can be increased, but as expense of the precision. The creation of this query is a task that could potentially have some aspects of NLP applied to it.

### 1.1.3 Data Filtering

The data filter stage involves reducing the amount of documents returned by the initial query down to a smaller subset of relevant document. This is can also be referred to as the abstract screening phrase [6].

The length of this stage is highly dependant on how many documents were returned by the initial query, often in the excess of 5000 studies for a single query. In response to this, stopping criteria methods have been proposed that aim to optimize two key parameters; the effort and the recall. That is to say we want to get as many relevant documents as possible, whilst looking at the fewest. Examples of approaches include the knee method [10] and the target method [3]. Other techniques could be applied and evaluated such as curve fitting.

### 1.1.4 Data Extraction

The data extraction phase involves pulling the relevant information from the filtered subset of studies. Examples of important information includes how many people took part in the study and what the results were.

Being able to extract the relevant information from studies presents itself as an information extraction problem. The task to automate the process of extracting relevant information would reduce time and complexities of manually reviewing studies [5].

## 1.2 Overview of the Report

We will start by looking at relevant literature **??** in the field of affect analysis and sentiment analysis. We will look at specific techniques **??** of which can be applied to assign an emotion to some text. We will then introduce the complexities **??** of dealing with twitter data such as the large range of vocabulary and incorrect spelling.

As this project has the potential to be flexible in its aims and objectives, we will define some clear end goals **??**. We will then evaluate the existing #happysheffield **??** code using formal natural language processing evaluation metrics. At this point will have a clear baseline system performance, which will aim to beat **??**.

In the next section we will begin to plan **??** the project and decide what algorithms/techniques to apply. We will also look at the available APIs and frameworks we can use to reduce the workload. A risk assessment will also be produced **??**, where we will look at the risks of handling large volumes of twitter data, as well as other concerns. We will then finish our planning by producing a Gantt chart **??**.

# Chapter 2

# Literature Survey

Systematic reviews have many different stages that propose themselves as a candidate for automation. This section is going to look at the techniques that have been applied for some of these stages in previous literature.

## 2.1 Indexing and Querying Medline

Medline is a large collection of medical literature and data from around the world. Typically each entity will contain a title and an abstract containing some information on the study. Whilest Medline as a whole is very easy to access [1], the large size and complexity of the data makes it difficult to retrieve the relevant information.

Being able to create a reliant index of Medline would help with the effectiveness of the queries. As such existing medline indexes and IR systems have been created [**?**]

# Chapter 3

# Novel Work

In this section I am going to present the work I have completed so far. Two main areas of the systematic review process has been focused on. Stopping criteria and indexing/querying pubmed.

## 3.1 Sample Method to Stopping

As approach to determining when to stop looking at document abstracts returned by the query we are proposing a new sampling method. The first step of this method is to randomly sample a returned set of documents into a subsets.

$$U = \frac{|D|}{S} \tag{3.1}$$

Where $U$ is the computed randomised subset, $D$ is the document collection and $S$ is the sample size.

# Bibliography


[1]

[2]

[3] Alharbi, A., and Stevenson, M. Ranking abstracts to identify relevant evidence for systematic reviews: The university of sheffield's approach to clef ehealth 2017 task 2. *CLEF 2017* (2017).

[4] Cormack, G. V., and Grossman, M. R. Engineering quality and reliability in technology-assisted review.

[5] Cormack, G. V., and Grossman, M. R. Technology-assisted review in empirical medicine: Waterloo participation in clef ehealth 2017. *CLEF 2017* (2017).

[6] Jonnalagadd, S. R., Goyal, P., and Huffman, M. D. Automating data extraction in systematic reviews: a systematic review.

[7] Kanoulas, E., Li1, D., Azzopardi, L., and Spijker, R. Clef 2017 technologically assisted reviews in empirical medicine overview.

[8] Lee, G. E. A study of convolutional neural networks for clinical document classification in systematic reviews: Sysreview at clef ehealth 2017. *CLEF 2017* (2017).

[9] Nunn, J. cochranes. `http://cccrg.cochrane.org/animated-storyboard-what-are-systematic-reviews`.

[10] of Tasmania, U. pico. `https://utas.libguides.com/SystematicReviews/FormulateQuestion`.

[11] Satopa, V., Albrecht, J., Irwin, D., and Raghavan, B. Finding a "kneedle" in a haystack: Detecting knee points in system behavior.