

Enabling Text Analytics

Jason Cairns

November 20, 2019

University of Auckland

Introduction

- The intention of this project is to develop a user friendly program that that allows for the performance of common text analyses.
- It is intended to be capable of integrating with iNZight, with a similar user base.
- The application is to be flexible enough to deal with a range of text formats, and produce attractive output quickly.

- Text Analytics is comprised of a variety of processes and techniques to extract information from text and provide a high-level overview.
- The text almost always requires some initial processing.

Lemmatisation

Stopwords

Sentiment

Associated Words

Summarisation

Program Description

Development Areas

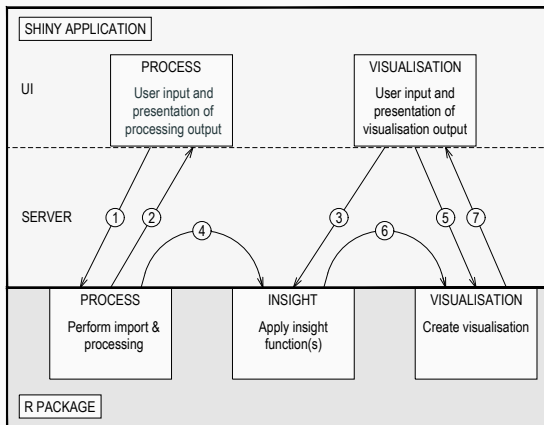
Processing

Insight

Visualisation

Program Architecture

PROGRAM ARCHITECTURE



LEGEND

- ① File name and processing options are selected in the user interface, and are passed as arguments to the processing functions from the package.
- ② The processing functions yield a processed dataframe, which is rendered in the interface.