

Media and Social Science Projects

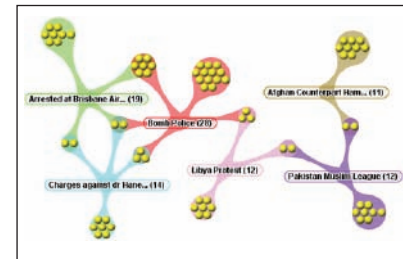
BBC News Browser

The aim of this pilot project is to facilitate improved browsing and presentation of online news stories through a combination of term extraction, concept clustering and visualization techniques.



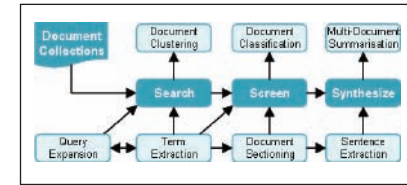
Key features include

- Automatic clustering of results into important topics with readable summary labels
- Dynamic document maps showing relationships between clusters in real-time
- Highlighting of significant terms for effective skimming and focused browsing
- Detection of similar documents to inform you of the wider context of the content



Assert

The Automatic Summarisation for Systematic Reviews using Text Mining (ASSERT) project centres around providing for broader institutional involvement in text mining and related activities. As part of this we will be developing an exemplar service for the social sciences domain, looking at how text mining techniques can support the systematic reviewing process. For this we are working closely with the EPPI-Centre, an organisation concerned with providing unbiased reviews of literature to support policy making.



Searching: extensive searches are carried out in order to locate as much relevant research as possible according to a query. This process is improved by using query expansion techniques based on the most important concepts (terms), similarities among terms but also ontologies and thesauri.

Screening: narrows the scope of search by reducing the collection to only the relevant documents to a specific review. The aim is to highlight key evidence and results that may impact on the policy.

Synthesis: correlates evidence from a plethora of resources and summarises the results. This process is improved by using multi-document summarisation.

The ASSERT project will develop and test text mining tools to support efficient access to evidence through a web portal. Currently, access to evidence is either available through general web search engines or specialist bibliographic databases. The scope of bibliographic databases is often narrow, necessitating multiple searches, whereas the scope of a search engine such as Google is very broad, resulting in huge numbers of irrelevant hits. We therefore propose to evaluate the extent to which text mining is able to overcome these problems.

Key features include

- Interactive query expansion to maximize coverage of an unbiased search
- Document clustering and visualization to aid exploration and overview of search results
- Document classification and sectioning to find relevant information within specific types of text.
- Automatic and custom built filters to narrow the focus of the collection into a manageable size
- Extractive summarization of a document to aid in fast indicative evaluation
- Chronological informative summary from multiple documents to help develop synthesis

Assist

Two new projects building upon the features developed in the ASSERT project and resulting from a community call for social science applications of text mining.

Case Study 1:

UK Educational Evidence Portal

This project is with the Evidence for Policy and Practice Information and Co-ordinating Centre (specifically its user involvement team) to work with NaCTeM to develop and evaluate an innovative search engine – using text mining – for a portal of education evidence, relevant to education practitioners and policy-makers. If successful, this project would be a high profile exemplar of the utility of text mining in the social sciences, with application beyond the single case described here.

Case Study 2: Frame Analysis of Media

This project is with the National Centre for e-Social Science to work with NaCTeM to develop and evaluate an innovative search engine – using text mining – for a portal of education evidence, relevant to education practitioners and policy-makers. If successful, this project would be a high profile exemplar of the utility of text mining in the social sciences, with application beyond the single case described

MANCHESTER
1824

The University
of Manchester



The National Centre
for Text Mining

Providing Text Mining Services to the UK academic community
www.nactem.ac.uk

The University of Manchester
Oxford Road
Manchester
M13 9PL

tel +44 (0)161 306 6000
www.manchester.ac.uk

Royal Charter Number RC000797
J1837 02.08