



JISC Project Plan Template

Overview of Project

1. Background

NaCTeM is the first publicly funded text mining centre in the world, supporting (with services and access to appropriate tools) the analysis of large collections of documents to discover previously unknown information. The establishment of NaCTeM has been as a result of an existing project currently funded by JISC (which has contributed £979,000) with additional funding from the BBSRC (Biotechnology and Biological Sciences Research Council) and the EPSRC (Engineering and Physical Sciences Research Council). The Centre is operated by the Universities of Manchester (lead partner) and Liverpool, with unfunded associate partners at the University of Geneva, the University of Tokyo, the University of California, Berkeley and the San Diego SuperComputer Centre and is currently focused on biosciences.

NaCTeM is one of the users of the National Grid that can demonstrate a direct relationship between Grid computing and the feasibility of its work, as many of the processes it undertakes would not be possible without Grid computing. NaCTeM's services are free of charge for usage by members of higher and further education institutions. Substantial institutional commitment to the establishment and on-going success of the text mining centre has been demonstrated by the partner universities, each of which have funded the secondment of at least the equivalent of one full-time member of staff to the project. Finally, the Centre is at a critical point as it has a global advantage in the text mining field and it is vital that it pushes into new areas to retain this lead for the UK.

The current project, **Automatic Summarisation for Systematic Reviews using Text Mining (ASSERT)** centres around providing for broader institutional involvement in text mining through a community call, while at the same time developing an exemplar service for the social sciences domain. An additional text mining expert for NaCTeM will participate in projects in the arts and humanities or other domains resulting from a community call. In parallel, the text mining expert will work with domain experts from within the social sciences. This background work developing an exemplar systematic review service will provide a foundation for the project work.

As a whole, ASSERT strongly contributes to the outcomes required by the e-Infrastructure programme:

- Greater participation by the social sciences in e-Research.
- Greater usage of the Grid for social science and arts and humanities based e-Research.

Currently there is very little use of text mining within the social science or arts and humanities areas. Despite this, there are potentially huge advantages to using text mining methods to save time for researchers, open up new areas of research and encourage new ways of doing research. There are also tangential benefits of demonstrating how the Grid can be used to help research with the social sciences and arts and humanities.

ASSERT thus interfaces with the social sciences community in two ways:

1. Development of a social sciences summarization exemplar service based on NaCTeM tools.
2. Support for the social science community projects involving text mining funded by the JISC as the result of a community call.

2. Aims and Objectives

The overall aim of ASSERT is to encourage greater participation by the social sciences community in e-Research by developing text mining technology (summarisation service) to facilitate the production of systematic reviews and to support a number of community projects related with text mining applications.

Before undertaking any new policy, practice, research or before making any other decisions it can be useful to find out *what is already known* about an issue in a fair, unbiased manner, in order to be of any scientific value to the community. This knowledge may include the findings of individual research studies that might, alone, be limited in their applicability and vulnerable to bias. In order to minimise this bias, therefore, a large number of people and organizations, such as the Cochrane Collaboration (<http://www.cochrane.org>), the Centre for Reviews and Dissemination's guidelines (<http://www.york.ac.uk/inst/crd/report4.htm>) have developed methods for locating research evidence and synthesising it in order to inform decision-making. They have developed ways of conducting literature reviews of research in a systematic way, which provide users with a 'short-cut' to relevant evidence.

Systematic reviews usually proceed along the following stages:

(i) First, extensive *searches* are carried out in order to *locate* as much relevant research as possible according to a query. These searches often include electronic databases, scanning reference lists and searching for unpublished literature. This stage includes the definition of a set of inclusion and exclusion criteria on which the researchers base their searching.

(ii) Then the mass of data retrieved by this process is *screened* until only the most relevant and reliable literature remains to form the focus of the review.

(iii) Finally, the literature is *synthesised* and summary reports are written to inform policy and practice. The summaries of research that are produced in this *systematic* way are then used to help users of research to make evidence-informed decisions.

The overall objectives of the proposed exemplar service are:

1. to develop cost-effective and rapid methods for locating relevant studies for input to a systematic review using a combination of text mining techniques;
2. to apply a suite of text mining tools that will support novel methods of information management in the domain of social science systematic reviews (document clustering, information extraction and text summarization);
3. to demonstrate the applicability of the text mining technology in social sciences, in cooperation with the Evidence for Policy and Practice Information and Co-ordinating Centre (EPPI) currently heavily involved in producing systematic reviews;
4. to disseminate the benefits of text mining to the wider social sciences community via NaCTeM services.

We have identified a key partner to participate as service users and domain experts in the proposed service development. The EPPI is part of the Social Science Research Unit at the Institute of Education in London, and has been involved in research synthesis since 1993. It is part of the ESRC's National Centre for Research Methods (NCRM), focusing specifically on building upon existing approaches to research synthesis to build an integrating framework that accommodates diverse types of information and research.

The existing NaCTeM tools need to be minimally tailored in order to meet the requirements of the new project. This will involve the addition of some minor features and some tuning to make the tools more effective when applied to problems of systematic reviews and in the domain of social sciences. However, the bulk of the work will be to produce a tool for a **summarisation service**. Additionally, the summarisation tool must interoperate effectively with the existing NaCTeM tools. For this reason, we intend to make the new tool compliant to the NaCTeM software infrastructure, Unstructured Information Management Architecture (UIMA). UIMA is an open, industrial-strength, scalable and extensible platform that has been developed by IBM¹ since 2001. UIMA provides an interoperability layer which allows for the composition of multiple analysis tools into a single application. A recent special edition of the IBM Systems Journal describes the framework in detail (<http://www.research.ibm.com/journal/sj43-3.html>)

The overall objectives of the proposed community call support are:

1. to liaise with domain experts from other institutions who will use text mining as an integral part of their projects supported under the community call
2. to configure appropriate text mining workflows for the specific purposes of the supported projects.

Currently, this task is performed mostly manually and encounters many problems. The proliferation of information in textual form means that the quantity of potentially relevant literature retrieved in the early stages of a review can become unmanageable. Reviewers have been accustomed to sacrificing specificity in their searches in order to ensure they have not missed any relevant studies. They conduct searches that yield large numbers of 'hits'. They then download the titles and abstracts, usually into bibliographic software, and look through them manually. This process is often called 'screening' and is the most time-consuming part of the review. Reviewers are finding that they sometimes need to scan through tens of thousands of titles and abstracts that are retrieved from large databases such as ERIC (<http://www.eric.ed.gov/>) and Medline (<http://www.nlm.nih.gov/>) to decide whether or not they meet the inclusion criteria for a review.

3. Overall Approach

3.1 Strategy

¹ Main contributions are from teams based at IBM Thomas J. Watson Research Center (New York), IBM Haifa Research Laboratory (Israel), IBM Development Laboratory Boeblingen (Germany) and IBM Almaden Research Center (California)

A number of phases can be identified in ASSERT required to make the summarisation services in social sciences a reality: (1) requirements analysis, (2) tool customisation (3) tool development (summarisation module) and (4) service exemplar development.

(1) The requirements analysis phase will involve close collaboration with the EPPI and will involve the following:

- definition by the user of the criteria for the classification of systematic reviews. Systematic reviews are currently classified according to inclusion and exclusion criteria, which are elicited from systematic reviewers.
- gathering of types of information required to produce a summary from social science related documents. In order to achieve this we will incorporate *viewpoints* (as defined by EPPI-Centre staff) to generate a summary, e.g. what is the background of a paper, methodology, outcomes, conclusions etc.

(2) The tool customisation phase will focus on the existing tools:

- T-MEMM part-of-speech tagger (bi-directional maximum entropy markov model)
- T-CFG parser (context free chunker)
- Enju deep syntactic parser
- TerMine terminology extraction system based on the C-value measure (hybrid system based on statistical knowledge and linguistic information)

For **parsing**, we will use our own deep linguistic HPSG parser, Enju (<http://www-tsujii.is.s.u-tokyo.ac.jp/enju/>). The efficient parsing algorithm of Enju and the wide-coverage probabilistic grammar it uses can effectively analyse the syntactico-semantic structure of complex sentences and provide us with the types of evidence needed in ASSERT. The annotated corpora will be kept as local copies and converted into XML to allow social science research groups to have indirect access via NaCTeM.

In order to proceed with the subsequent tasks, we assume that we have as a starting point a set of documents supplied by the user e.g. EPPI, which will then be clustered and summarised.

For **document clustering**, we will use the existing open source software CLUTO (<http://glaros.dtc.umn.edu/gkhome/cluto/cluto/overview>). CLUTO is a software package for clustering low- and high-dimensional datasets and for analyzing the characteristics of the various clusters. CLUTO is well-suited for clustering data sets arising in many diverse application areas including information retrieval, web, GIS, etc. CLUTO's distribution consists of both stand-alone programs and a library via which an application program can access directly the various clustering and analysis algorithms implemented in CLUTO. This software contains versions that run under Linux, UNIX and Windows environments. The goal of document clustering is to assign documents based on the topic they discuss. The produced clusters, also called **topic-clusters**, should ideally correspond to a topic that is shared by all the documents they contain and by no other document in the collection. Identifying the topic of a document is not a straightforward procedure. Current research in document clustering tends to use actual words and their frequencies that are contained within a document in order to identify the topic of a document. However the same word may be used to denote the same

topic (polysemy, ambiguity) or different forms of the same word appear in text (variation). These occurrences may potentially divert the document clustering algorithm, leading it to incorrect decisions. Thus, any background knowledge (ontology) may enhance the clustering results attained by a document clustering algorithm. If we can use ontologies to build links that correlate different terms appearing within the documents of our collection, then we can safely expect an increment in the quality of our clustering solution. However, in such domains the classification may be of more use especially if we have a good understanding of the different settings in which the objects participate. Documentation of the software and a user manual can be found at <http://glaros.dtc.umn.edu/gkhome/cluto/cluto/download>.

We note that we will **leverage** from existing tools for document clustering since these are adequate for our purposes, allowing us to focus on the summarisation task.

(3) The core component of the project is the development of the summarisation module.

Summarisation takes as input the sentences extracted from our customised tools and is based on the identification of topics in the documents and the selection of salient sentences for each topic.

Next we will construct a classifier which will categorize the input information according to the viewpoints defined in step (1). For that purpose, machine learning techniques will be used on an annotated test corpus annotated (according to the **viewpoints** defined by the user). We will use a classifier based on Support Vector Machines (SVMs), which uses a number of features such as n-gram frequency, dependency relation, sentence position, etc. We will evaluate our classifier independently to determine its performance with respect to the specific task. The quality of the classifier is very important because it may affect the overall quality of the summarization system. The output of this step is a set of automatically extracted sentences that include the viewpoints as discovered by the classifier.

The last step produces the summary. It takes as input the sentences extracted in step 3. These sentences now contain an annotation of the viewpoints which make up the summary. We adopt the following strategy to summarize information:

- we divide the document to be summarized into sub-sections according to the viewpoints given by the user (sectioning);
- we extract the most salient description from each viewpoint, and
- we exclude redundant information scattered over the input documents.

An important aspect of this sectioning step is the statistical analysis (e.g. term frequency, sentence location, clue phrases, etc.) of the input documents and the extracted sentences. Examples of sectioning are: **background, conclusions, methodology** etc. Sectioning will be based on the viewpoints provided by the EPPI reviewers.

We adopt a practical solution to summarization as it is still very difficult to generate comprehensible summaries from an internal linguistic representation. In addition, domain specific documents use a number of technical terms (and variants) for describing the same concept. Hence, it is crucial to carefully perform the statistical analysis to improve the quality of a summary, incorporating terminological variations such as synonyms, acronyms, etc. Our summarization system is based on a systematic terminological analysis which is important for domain specific areas.

All terms in the documents are mapped into concepts using thesauri such as the HASSET Thesaurus <http://www.data-archive.ac.uk/search/hassetSearch.asp>, the National Public Health Thesaurus and the British Education Thesaurus. We will use as features concept- concept pairs by examining co-occurrences within sentences. The weights of concept-concept relations are calculated by using the frequency of the co-occurrences.

To detect a set of topics in the source documents, we shall apply EM clustering to the source sentences represented by the features (i.e., concept-concept pairs). Unlike the k-means clustering, EM clustering does not assume each cluster to have the same number of instances (sentences), which is a good characteristic for the topic detection.

(4) This phase will focus on the development of a service exemplar. This exemplar will serve a number of purposes. Firstly, it will address real problems encountered by the social sciences community. Secondly, it will demonstrate the effectiveness of the tools developed by the Centre as a means of solving problems. The initial focus of this activity will be to link the exemplar with the identified requirements of the social sciences community (EPPI). This will have the added benefit of ensuring a close interaction and cooperation with EPPI throughout the project.

The distribution of work between NaCTeM and EPPI is shown below; B, D, E, G (NaCTeM), A,C,F,H (EPPI)

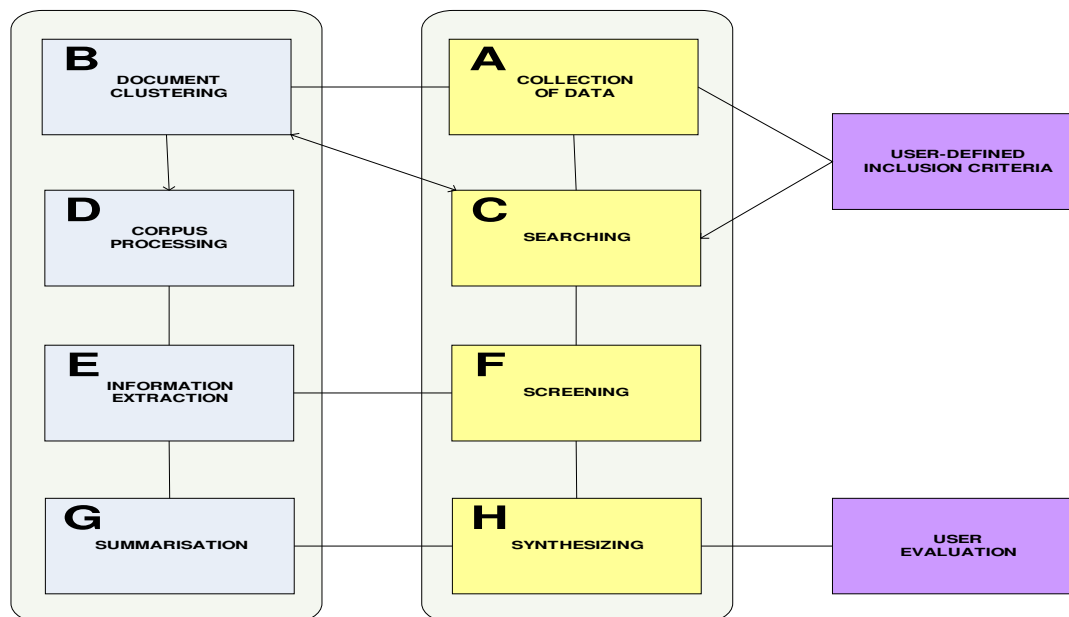


Figure 1: Architecture of the project and distribution of the work

3.2 Important Issues

External users will be able to access to the Centre’s text mining tools integrated with the software infrastructure at the end of the second year of ASSERT.

At the earliest feasible stage, we will expose our summarisation system to EPPI community, expanding to the wider social sciences community once our tool has been demonstrated to exhibit the following characteristics:

- robustness
- efficiency
- scalability

Scalable, robust, efficient and rapidly responsive services for very large collections, simultaneous requests, the need to consult large-scale resources (corpora and thesauri) are critical aspects of the service provision.

- Focus on *user-need* related development whilst using experts to feed-in research results
Early involvement of the EPPI users will ensure this characteristic.
- Documentation and awareness are also crucial aspects of the user experience and will be core activities during the development work.

3.3 Scope and boundaries of the work

Text mining technologies have the potential to revolutionize the way we approach research synthesis. Using automated techniques rather than people to perform this time consuming task, reviews would be completed much more quickly and would therefore be much cheaper. They would also be timelier – and become a more useful mechanism for informing decisions which need to be taken relatively quickly. Moreover, if text mining technologies carry out this task, then the development of complex electronic search strategies which combine sensitivity with the retrieval of a manageable number of references would be less critical. Search strategies could simply be as sensitive as possible – with computers managing the screening of hundreds of thousands of potentially relevant references.

Our work in ASSERT will be application-oriented and will concentrate on the following:

- customising existing tools and
- developing a new text mining component (summarisation module).
- providing scalable text mining software

Wherever possible we will take advantage of existing work by the Grid and cluster computing communities. The close involvement of Manchester Computing, one of the partners in the JISC-supported National Grid Service, and the wider e-science community will be invaluable in this respect.

The aims of ASSERT are to produce a summarisation and sectioning tool for systematic reviews. We will not enhance or customise the document clustering algorithms for our purpose.

3.4 Critical success factors.

ASSERT will have a strong **user focus**; we will be developing a text mining service for the social sciences community. We have identified a specific problem within systematic reviews and we will address this problem.

Our evaluation will be performed from two aspects:

- evaluation of the text mining components; performance of the summarisation component
- user oriented evaluation: feedback from EPPI whether text mining technology actually facilitates the task of searching and screening in systematic reviews.

Critical success factors in the success of the project will include:

- **Positive user responses** to our services. It is important that we conduct an ongoing dialog with our user communities.
- **Scalability** of the tools that we provide. The capabilities of text mining have not been yet explored in social sciences, and ASSERT will be the first to do so. Existing tools like qualitative based analysis (QALDAS) content analysis do not deploy text mining techniques. ASSERT will enhance the capability of existing tools used in social sciences.
- The ease with which the software infrastructure allows other tools to **interoperate** will be another important factor in our success. We believe that adoption of the tools and infrastructure provided by the Centre will be greatly increased if other developers are able to straightforwardly integrate their own tools with ours. This integration must therefore be based on the cooperation of loosely-coupled systems, avoiding any requirement for major rewriting of existing software. **Interoperability** is the ability to combine modules and exchange data, meta-data and other resources to maximise their re-use <http://www.ukoln.ac.uk/interop-focus/about/leaflet.html>. There are several ways to achieve interoperability:
 - Modules can communicate through remote procedure calls
 - Web Services
 - common APIs
 - common data exchange formats. In this case, every module has to comply with the output format of the previous components,
- **Open standards** are a vital mechanism in achieving interoperability and acceptance amongst the wider community. We will adopt existing standards where they are available.

4. Project Outputs

The output would be a service that would automatically collect information about specific topics (as specified by the user), consult with knowledge bases using a combination of text mining techniques to augment the search topics, and automatically provide relevant summaries. This service would assist systematic reviewers in the social sciences in classifying and summarizing the thousands of abstracts and full texts of primary research studies according to the reviewers' viewpoints. It would be backed up by an exemplar based on the practices and documents used by the EPPI-Centre, since this is a good application of the usefulness and applicability of text mining tools and techniques in systematic reviewing. The exemplar would summarize an article or a set of articles by matching the relevance of documents to the user defined criteria by:

- Clustering documents relevant to the user's information need by improving searching using a combination of text mining techniques and existing knowledge sources (thesauri);

- Identifying and classifying the relevant types of information using information extraction;
- Generating summaries by “condensing” the classified information.

The service will provide also

- Documentation about how to use the service
- Training support for the software

5. Project Outcomes

We envisage the major intangible outcome of the project will be a general raising of awareness of text mining in the social sciences community, and of the tools that are available. NaCTeM has recently organised a dedicated workshop on “Bridging qualitative and quantitative research methods for social sciences using text mining” funded by the ESRC National Centre for e-Social Science <http://www.ncess.ac.uk/events/agenda/textmining/>

Through direct involvement in the community projects (the e-Research framework) we will identify requirements for the following text mining applications such as:

- sentiment analysis
- content based analysis
- qualitative analysis tools
- forensic linguistics
- authorship identification
- etc.

NaCTeM is already acting, via the organisation of workshops, tutorials etc as an **educator** in text mining technology to biosciences.

The new project will allow us to enhance the take-up of text mining in the social sciences.

Potential beneficiaries and users of the project outputs will include researchers in:

- qualitative inductive research
- social sciences
- e-social science

A dedicated workshop related with the issues, problems and solutions for systematic research synthesis and a demonstrator of the text mining tools deployed for that purpose, will be organised in conjunction with the users (EPPI) and the National Centre for e-Social Science.

6. Stakeholder Analysis

Stakeholder	Interest / stake	Importance
NaCTeM	<p>NaCTeM will customise text mining tools for the UK social sciences community and in particular EPPI.</p> <p>NaCTeM will consolidate text mining as a scientific activity in its own right, by demonstrating its applicability in a variety of domains and text types.</p> <p>NaCTeM will ensure service provision in social sciences</p> <p>NaCTeM will revolutionise the way systematic reviews are conducted.</p>	High
University of Manchester	<p>Enhancing the image and visibility of the University as leading text mining for the social sciences community</p> <p>Making the University a focus of excellence in text mining</p> <p>Investment in NaCTeM and text mining research</p>	High
University of Liverpool	<p>Linking work on digital libraries with text mining</p> <p>Demonstrating the applicability of text mining in social sciences</p> <p>Preparing for the community call and porting text mining into humanities and arts.</p>	Medium
HEFCE (JISC)	Allocation of significant resources to the Centre with	High

	<p>expectation of development of service that is relevant to the needs of the social sciences community.</p> <p>Strategic alignment through capital programs</p> <p>Expanding into e-Research, e-Social Science programmes</p>	
EPPI	<p>Improving the way they conduct systematic reviews</p> <p>Allocation of resources to annotate a gold standard</p> <p>Provision of requirements analysis</p> <p>Evaluation of text mining tools for their needs</p>	High
National Centre for e-Social Science, e-Research, ESRC	<p>NaCTeM will engage in complementary activities, leading to synergistic and cross-group activities to develop new insights and further both individual aims and UK-wide collective success.</p>	High
E-Infrastructure	<p>Alignment with other programs</p> <p>Community engagement</p> <p>Alignment with National Grid services</p>	High
Academic users in social sciences	<p>NaCTeM will provide services and support to the wider social sciences community</p>	Medium
Respondents to the community call	<p>Support from text mining services</p> <p>Engagement with NaCTeM</p> <p>Seeking catalysis of a text mining community in social</p>	High

	sciences in the UK	
--	--------------------	--

INSERT RISK ANALYSIS HERE (LANDSCAPE) FILE ASSERT_RISK_ANALYSIS

8. Standards

Name of standard or specification	Version	Notes
UIMA		Infrastructure for text mining pipeline
W3C web service standards		Maintaining web services
XML		Mark-up language
DocBook		XML-based structural-level mark-up language.

9. Technical Development

Technical development will adhere to the general provisions of the NaCTeM service development process documented in the NaCTeM project plan. This includes the setting up of proposed solutions and use cases, system architecture and service roadmap. We will exploit the NaCTeM Subversion source code control repository and policy to ensure that service development follows best practice. This covers source control, source code standards, daily builds, automated tests and versioning.

10. Intellectual Property Rights

The summarisation software will be released under a licence agreement in accordance with the JISC policy and the existing tools of the National centre. The released software will be available in perpetuity. Copyright University of Manchester.

- Enju: copyright held by University of Tokyo
- T-MEMM part-of-speech tagger (bi-directional maximum entropy markov model) : copyright held by University of Tokyo
- T-CFG parser (context free chunker) copyright held by University of Tokyo (all available through NaCTeM and covered by NaCTeM consortium agreement and licencing)
- TerMine: copyright held by University of Manchester
- CLUTO open source software copyright University of Minnesota

Project Resources

11. Project Partners

University of Manchester

Main contact: Dr Sophia Ananiadou

School of Informatics
 University of Manchester
 Manchester
 M60 1QD

e-mail: Sophia.Ananiadou@manchester.ac.uk

Phone: (0161) 306 3092

Role: Leading the project. 1 full-time research associate, fully-funded by JISC.

EPPI

Main contact: Dr James Thomas

Social Science Research Unit
 Institute of Education, University of London
 18 Woburn Square
 London
 WC1H 0NR

telephone: +44 (0)20 7612 6844

fax: +44 (0)20 7612 6400

email: j.thomas@ioe.ac.uk

Role: problem provider, requirements, evaluation, no funding from JISC

12. Project Management

The project manager for this project will be Dr Sophia Ananiadou, School of Informatics, University of Manchester. The work will be managed in line with the JISC project management guidelines.

13. Programme Support

Consultation regarding the related upcoming community call.

14. Budget

Detailed budget for the proposal, profiled over Academic Year

	2005/06 AY	2006/07 AY	2007/08 AY	Total
Staff salary ²	£3,107	£38,773	£36,964	£78,844
Overhead ³	£4,472	£53,894	£49,618	£107,984
Contribution from University of	-£3,014	-£36,394	-£33,576	-£72,984

² Assumes pay award of 4% per annum and includes overheads.

³ Overhead includes contribution from JISC for estate costs, indirect costs and PI as calculated using the TRAC method, as supplied by John Keane, University of Manchester.

Manchester ⁴				
Dissemination	£1,800	£6,800	£6,800	£15,400
Equipment	£8,856			£8,856
Community call		£161,000	£200,900	£361,900
Total	£15,221	£224,073	£260,706	£500,000

Detailed Project Planning

15. Work-packages

Work-package	Summary
1 Project Management	Managing the day-to-day activities of the project. Ensuring that deliverables are delivered on schedule. Writing plans and reports.
2 Requirements Gathering and Evaluation	Test data gathering, requirements analysis by users, set up of evaluation methodology, creation of gold standard. Ongoing third party evaluation throughout project and testing.
3 Information Extraction	Customisation of existing text mining tools for social science applications
4 Document Clustering	Use of document clustering software and integration into the NaCTeM pipeline
5 Summarisation	Development of a scalable summarisation engine; integration into existing NaCTeM software infrastructure
6 Service Exemplar	Development of service exemplar which demonstrates the full capabilities of the summarisation tool.
7 Dissemination	Development of a Roadmap for availability of summarisation service, presenting work to conferences, organisation of a workshop
8 Support for the Community Call	Promote text mining through dissemination activities and engagement with user groups

⁴ There is a strong level of commitment from Manchester University to recruit an additional text mining expert as they feel that this resource is vital to the continued growth of NaCTeM. They are therefore prepared to make a significant contribution to the overheads for this person.

16. Evaluation Plan

Evaluation is an essential part of a practical discipline like automatic summarization. In general, evaluation of automatic summarization is categorized into two types: intrinsic evaluation which tests the system in itself; extrinsic evaluation which tests the system in relation to some other task.

Two kinds of intrinsic evaluation are typically carried out using summarization.

- The first is a quality evaluation which measures how a summary reads. This can be assessed by on-line evaluation (e.g., having subjects grade summaries for readability) or off-line evaluation (e.g., metric to measure readability).
- The second type of intrinsic evaluation is the degree of informativeness which measures how much information from the source a summary preserves at different levels of compression. This can also be assessed by on-line or off-line evaluation similarly to the quality evaluation.

The idea of an extrinsic evaluation is to determine the effect of summarization on some other task. A variety of different tasks have been proposed, for example, finding documents relevant to one’s need by reading summaries, effort required to post-edit a summary to bring it to some acceptable state, etc.

In order to conduct proper evaluation, the creation of a gold standard would be required. This task can only be conducted by the EPPI. In case, due to lack of resources, we are not able to have access to a gold standard other evaluation strategies will be investigated and adopted. These may include comparison of a generated summary with an existing systematic review to determine readability, lack of redundancy, informativeness, production time, etc

Timing	Factor to Evaluate	Questions to Address	Method(s)	Measure of Success
Tool Evaluation (Summarisation System)				
Month 14	Readability, content of system output according to requirements	Is the system useful to support systematic reviewers?	Evaluation metrics for summarisation and feedback from users	Performance rates according to existing metrics and response from users (70% rate system as useful)
Month 20	Scalability, speed	Is the speed of the system adequate for large scale use?	Metrics	75% of users rate system as useful for daily use
Dissemination				
Month 22	Awareness of usefulness of text mining for social sciences	Is the social scientist aware of the usefulness of text mining?	Evaluation questionnaire after workshop	70% of respondents award satisfactory score or better

17. Quality Plan

Timing	Compliance With	QA Method(s)	Evidence of Compliance
Lifetime of project	W3C Web standards	Compliance with NaCTeM Website QA Policy, which requires compliance with XHTML 1.0 Transitional and CSS 2	Automated tests are run daily to ensure compliance
Lifetime of project	Fitness for purpose	NaCTeM Software Development Process requires development of unit tests, integration tests, load tests. Integration of tools with the Software Infrastructure will be tested with the Software Infrastructure Compatibility Test Suite.	Software passes automated tests
Lifetime of project	Best practice for software development	Compliance with NaCTeM Software Development Process	Production and validation of documentation, test plans (unit, module, system, regression), profiling, mini-milestones. Technical training as required.
Lifetime of project	Adherence to specifications	NaCTeM Software Development Process requires development of automated acceptance tests to validate that implementations conform to user requirements	Implementations pass automated acceptance tests.
Lifetime of project	Web service standards	Compliance with appropriate W3C Web Services standards	Automated testing
Lifetime of project	Accessibility standards	Compliance with NaCTeM Website QA Policy, which requires compliance with W3C Web Content Accessibility Guidelines 1.0 with a conformance level of	Automated testing

		Triple-A. Compliance with NaCTeM Software Accessibility QA Policy.	
Lifetime of project	Documentation	Compliance with NaCTeM Software Development Process, which requires full documentation of software APIs	Automated checking of documentation during daily build. Regular code reviews.

18. Dissemination Plan

Timing	Dissemination Activity	Audience	Purpose	Key Message
M3 onwards	Incorporation of project information into NaCTeM web site	Social science community	Awareness, Information, Involvement	Roadmap for availability of summarisation service Support for community call
Lifetime of project	Conference presentations and posters	Peer community User community Funding community	Inform Promote Engage	R & D promulgation Education Technical engagement
End of project	Workshop on text mining in social sciences	User community Peer community	Inform Disseminate Promote	Awareness raising

19. Exit and Sustainability Plans

Project Outputs	Action for Take-up & Embedding	Action for Exit
Software Service	Ensure tool address users' needs. Ensure tool is stable and easy to use Ensure availability from web site. Promote use through dissemination activities.	Been involved with community from initial stage of project Ensure source code is well documented. Investigate commercial exploitation and/or open source

		licensing (via OSS Watch) Long-term service provision.
Software Support	Ensure availability from web site. Provide documentation Integration with NaCTeM help desk	Long term service provision
Community Call Support	Promote through dissemination activities and engagement with user groups Help to build a community of social science text miners via community call project support	Engage with further projects with community as appropriate Continuation within long-term service provision. Pursue integration of community call project outputs in overall NaCTeM service offering (involve OSS watch)

Project Outputs	Why Sustainable	Scenarios for Taking Forward	Issues to Address
Software	Based on proven and state-of-the-art research Address real needs of users Generic tools which can be applied to different problems.	Partnering with commercial suppliers Provision of commercial hosted services in different domains	Focus on users' requirements IPR and licensing arrangements. Academic versus commercial usage.

Appendixes

Appendix A. Project Budget

	2005/06 AY	2006/07 AY	2007/08 AY	Total
Staff salary ⁵	£3,107	£38,773	£36,964	£78,844
Overhead ⁶	£4,472	£53,894	£49,618	£107,984
Contribution from University of Manchester ⁷	-£3,014	-£36,394	-£33,576	-£72,984
Dissemination	£1,800	£6,800	£6,800	£15,400
Equipment	£8,856			£8,856
Community call		£161,000	£200,900	£361,900
Total	£15,221	£224,073	£260,706	£500,000

INSERT APPENDIX OF WPS HERE IN LANDSCAPE (FILE ASSERT_APPENDIX)

⁵ Assumes pay award of 4% per annum and includes overheads.

⁶ Overhead includes contribution from JISC for estate costs, indirect costs and PI as calculated using the TRAC method, as supplied by John Keane, University of Manchester.

⁷ There is a strong level of commitment from Manchester University to recruit an additional text mining expert as they feel that this resource is vital to the continued growth of NaCTeM. They are therefore prepared to make a significant contribution to the overheads for this person.