Project ref. no.	*IST-1999-11748*
Project acronym	**LIMBER**
Project full title	**Language Independent Metadata Browsing of European Resources**
Security (distribution level)	Confidential
Contractual date of delivery	Feb 2002
Actual date of delivery	21st Feb 2002
Deliverable number	
Deliverable name	LIMBER Final Report
Type	Report
Status & version	Final 1.0
Number of pages	24
WP contributing to the deliverable	
WP / Task responsible	CLRC
Other contributors	Intrasoft International, NSD, UKDA
Authors	Anne Etheridge *UKDA*, Michael Wilson *CLRC,* Terje Sande *NSD,* Antonis Ramfos *Intrasoft International*
EC Project Officer	Mr. Kimmo Rossi
Keywords	multi-lingual thesaurus, thesaurus management system, metadata, standards
Abstract (for dissemination)	This report describes the final state of the LIMBER project and its deliverables.

Executive Summary

This is the final report of the LIMBER IST funded project that produced 5 major products:

ELST: a multilingual thesaurus (English, French, Spanish, German) for the social science domain;

A metadata representation and format for the metadata entries for data sets held in social science data archives;

An internationalised user interface to a data archive system to access metadata, and to the multilingual thesaurus management system localised to English, French, Spanish and German languages

A text categorisation tool to automatically index metadata entries with keywords to allow keyword search on the metadata

The Thesaurus Management System (TMS) to store and serve the thesaurus from a server to a client, and with an API to allow access from other clients.

The project was completed to budget, but with a three month overrun on time due to staff illness and departure from the partners.

1. Contents

EXECUTIVE SUMMARY 2

1. CONTENTS 3

2. INTRODUCTION 4

3. THE PROBLEM ADDRESSED 4

3.1 User requirements and resulting product profile 5

4. INTERNATIONALISING THE QUERY PROCESS 7

5. WORK DONE 9

5.1 Conformance to the contract 14

5.2 Deliverables 14

5.3 Evaluation and User Involvement 14

5.4 Standardisation Activities 15

5.5 Project management 15

6. USERS 16

7. OBJECTIVES 16

8. PROBLEMS ENCOUNTERED 17

9. PLANNING 18

10. EXPLOITATION 19

11. PROJECT PUBLICATIONS. 21

12. DELIVERABLES PRODUCED 23

13. OTHER PROJECT 'PRODUCTS' 24

2. Introduction

This is the final report of the LIMBER IST funded project. Limber produced 5 major products which are described in section 2 below:

ELST: a multilingual thesaurus (English, French, Spanish, German) for the social science domain;

A metadata representation and format for the metadata entries for data sets held in social science data archives;

An internationalised user interface to a data archive system to access metadata, and to the multilingual thesaurus management system localised to English, French, Spanish and German languages

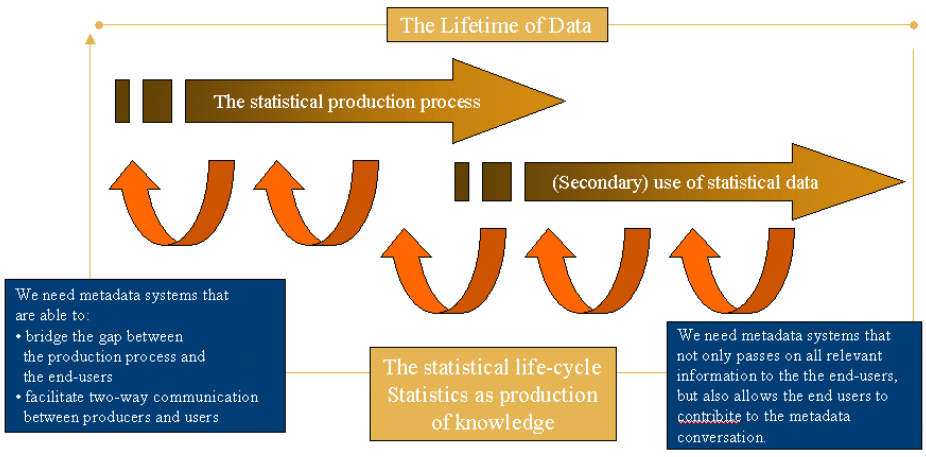
A text categorisation tool to automatically index metadata entries with keywords to allow keyword search on the metadata

The Thesaurus Management System (TMS) to store and serve the thesaurus from a server to a client, and with an API to allow access from other clients.

LIMBER ran from Jan 2000 until Dec 2001 for 24 months. Of these, as was planned, no work was done in the first 3 months which were used to recruit staff for those partners using the additional costs funding model. The project was originally planned to be completed by Sept 2001, but was extended for a further 3 months due to illness and departure of employees of two partners, who required the extra time to complete the contacted work.

3. The Problem Addressed

Most countries have a national archive for social science data containing coded socio-economic and environmental information. These data archives vary in size, but generally contain a few hundred gigabytes of data. They are accessible to some extent through the Internet, and are rapidly moving to provide web-based access.



In a more integrated Europe, it is becoming increasingly important to allow access to such information across national and linguistic boundaries, so that decision makers can be provided with a broader range of real data as evidence on which to base their planning decisions.

The LIMBER project aims to enable this by providing a uniform metadata description which can be used to allow these archives to be accessible in a range of languages. The results will also support a range of potential applications involving the integration and interoperability of datasets within and across archives based on the existing tool Nesstar.

3.1 User requirements and resulting product profile

An analysis was made of the business model behind the data archives, that identified the users at various stages. Use cases were drafted for each of these users. The role of automated tools to index metadata and access it were then presented at a workshop including archive users and LIMBER user group representatives.

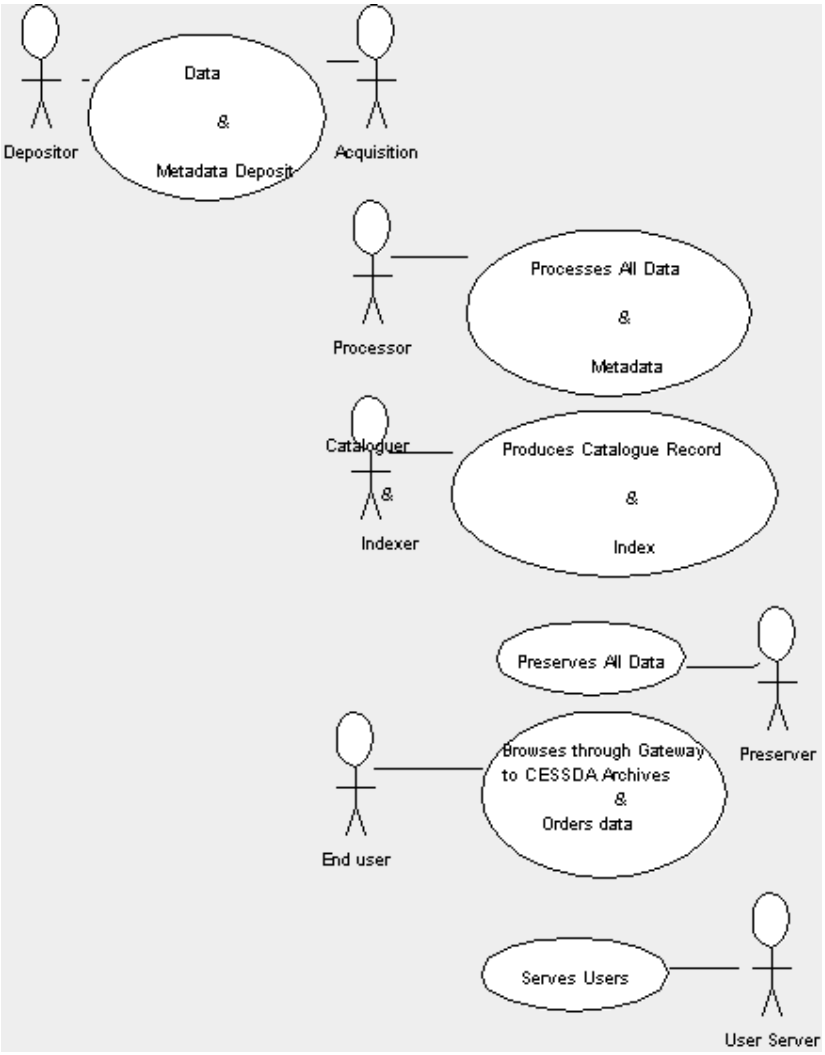
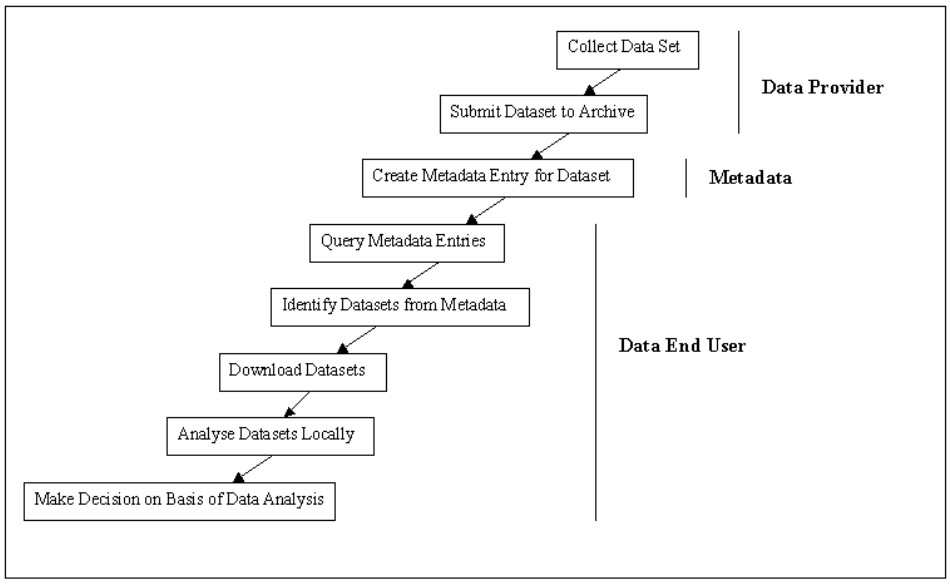


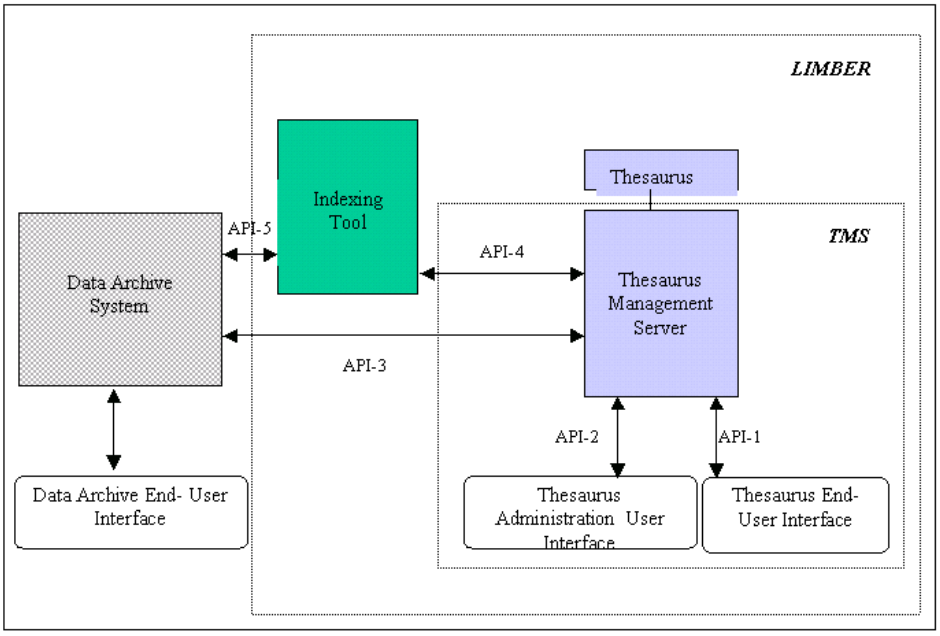
Figure: The business model of the Data Archives.

The workshop produced a detailed set of requirements for the functionality of the tools, and for the internationalisation of the interface. These requirements were used as the basis of the system architecture.

The flow of information through an archive identifies the actors and stakeholders in the process, and shows where support is required for archivists, contributors and end users.



This resulted in an architecture for the LIMBER system that addressed both the end user to provide queries using terms from a multilingual thesaurus hosted on a server and returned data which could have its labels etc.. translated through the same server, but also for the archivist to provide an indexing tool to help produce the metadata that was required to support the search.



4. Internationalising the Query process

As a result of the work achieved, a simple query will provide access to metadata, which in turn ca provide access to data, that can be viewed in different languages, where the different parts of a presentation of the data arise from clearly different sources as shown in the figures below:

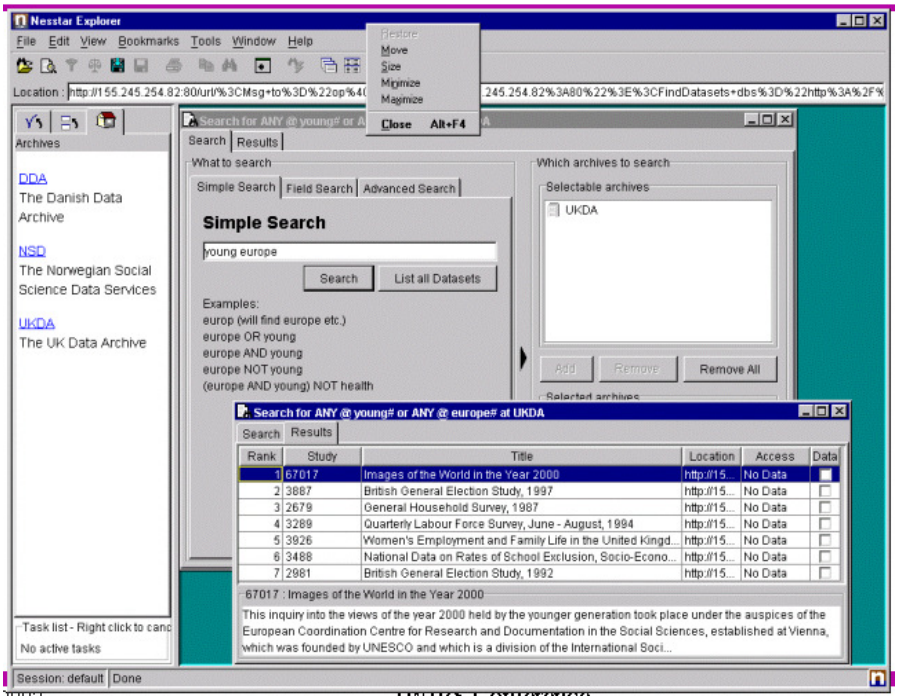


Figure: The NESSTAR query interface and returned metadata

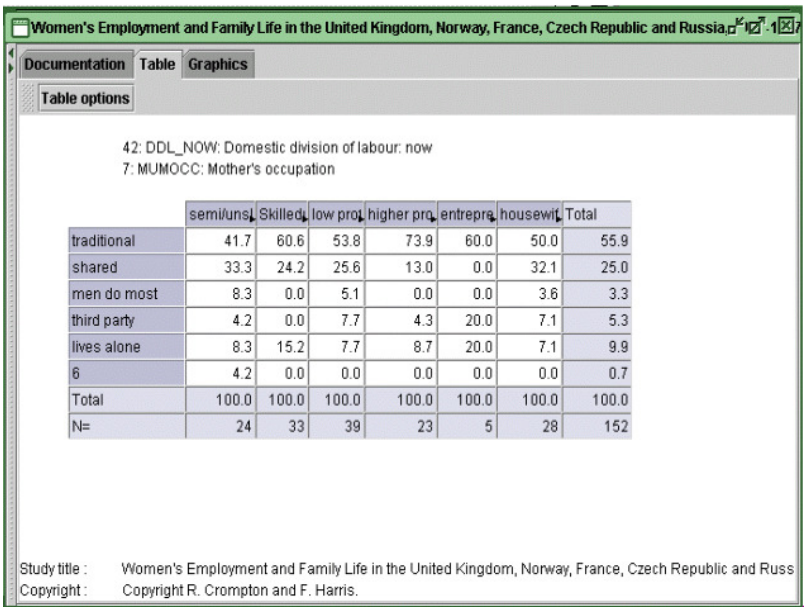


Figure: The dataset returned from a query, displayed by NESSTAR as a table.

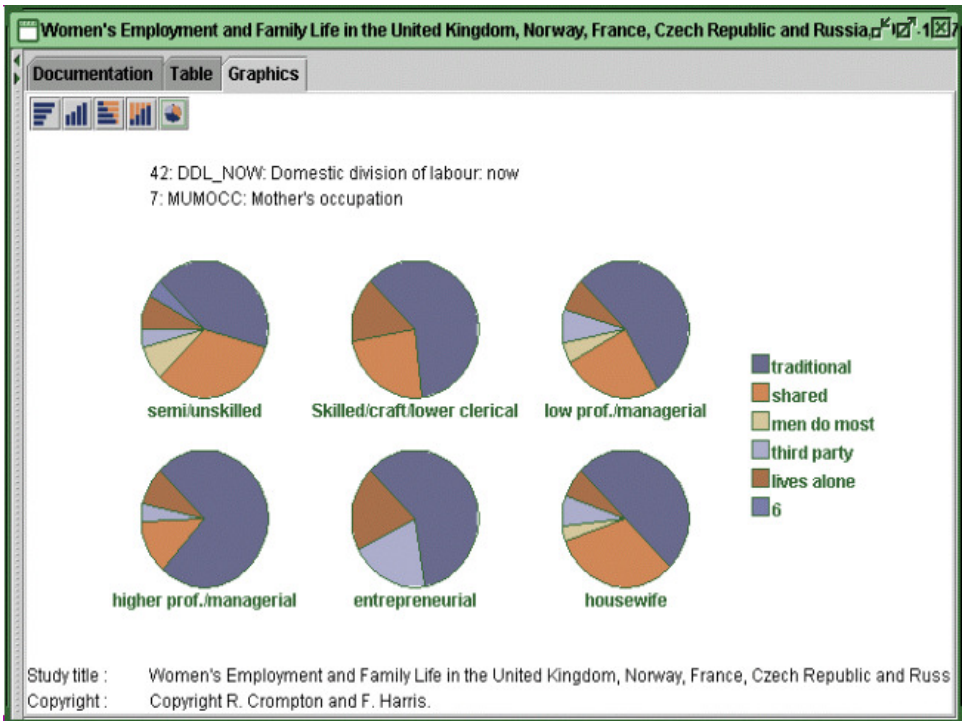


Figure: Graph of datasets returned through NESSTAR using the LIMBER TMS in support - an English dataset displayed in English.

Figure: Graph of datasets through NESSTAR using the LIMBER TMS in support - an English dataset displayed in another language. Different parts of the data presented are translated from different sources as shown.

5. Work Done

The conclusion of the project for each of its five major products is:

1) ELST: a multilingual thesaurus (English, French, Spanish, German) for the social science domain as a restricted vocabulary for indexing and accessing metadata entries. The thesaurus has been produced containing 49 hierarchies, incorporating 1456 preferred terms in an MS-Access database which writes out RDF or the thesaurus conventional alphabetic listing and has tools to examine the thesaurus and search using terms in one language to access metadata in another. Although produced, it is actively being used, and is therefore constantly under revision by UKDA. In particular, following the project further translations of terms into Finnish, Norwegian, Danish & Greek are planned, as are the inclusion of terms related through inexact translations in addition to the exact translations already included. Also the CESSDA group of European Data Archives have agreed to adopt ELSST as the European Controlled Vocabulary for Social Science, and will meet to define an adoption timetable in March 2002. The RDF Thesaurus Interchange Format is available for download to the reviewers and the public from the LIMBER web site, as is a fragment of ELSST in that format. Deliverables D4.1 and D4.2 have been produced describing the thesaurus, the translation methodology, the RDF Thesaurus Interchange Format and including a user guide.

Figure: RDF representation of the single entry for "Economics" in the ELSST Thesaurus.

```
<?xml version="1.0" encoding="UTF-8" ?>
- <rdf:RDF xml:lang="en"
    xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
    xmlns:rdfs="http://www.w3.org/2000/01/rdf-schema#"
    xmlns:thes="http://www.limber.rl.ac.uk/External/thesaurus-
    iso.rdf">
  - <rdf:Description rdf:ID="EN-N">
        <rdf:type rdf:resource="http://www.data-
      archive.ac.uk/Limber/thesaurus#TopConcept" />
    <thes:ClassificationCode>N</thes:ClassificationCode>
    - <thes:PreferredTerm>
      - <rdf:Description>
          <rdf:type rdf:resource="http://www.data-
        archive.ac.uk/Limber/thesaurus#Term" />
                                    <thes:inLanguageOf
          rdf:resource="http://www.data-
          archive.ac.uk/Limber/ISO639#en" />
        <rdf:value>ECONOMICS</rdf:value>
      </rdf:Description>
    </thes:PreferredTerm>
    - <thes:UsedFor>
      - <rdf:Description>
          <rdf:type rdf:resource="http://www.data-
        archive.ac.uk/Limber/thesaurus#Term" />
                                    <thes:inLanguageOf
          rdf:resource="http://www.data-
          archive.ac.uk/Limber/ISO639#en" />
```

<rdf:value>**POLITICAL ECONOMY**</rdf:value>
</rdf:Description>
</thes:UsedFor>
<thes:NarrowerConcept rdf:resource="**#EN-N50/59**" />
<thes:NarrowerConcept rdf:resource="**#EN-N98.20**" />
<thes:NarrowerConcept rdf:resource="**#EN-N35.50**" />
<thes:NarrowerConcept rdf:resource="**#EN-Q21**" />
<thes:NarrowerConcept rdf:resource="**#EN-N90.10.10**" />
<thes:NarrowerConcept rdf:resource="**#EN-T30**" />
<thes:NarrowerConcept rdf:resource="**#EN-L84.50.20**" />
<thes:NarrowerConcept rdf:resource="**#EN-N28/30**" />
<thes:NarrowerConcept rdf:resource="**#EN-N04.15.20**" />
<thes:NarrowerConcept rdf:resource="**#EN-N11.10**" />
<thes:NarrowerConcept rdf:resource="**#EN-N12**" />
<thes:NarrowerConcept rdf:resource="**#EN-N21/26**" />
<thes:NarrowerConcept rdf:resource="**#EN-N06/10**" />
<thes:NarrowerConcept rdf:resource="**#EN-N69.30**" />
<thes:NarrowerConcept rdf:resource="**#EN-J20.05**" />
<thes:NarrowerConcept rdf:resource="**#EN-N79/99**" />
<thes:NarrowerConcept rdf:resource="**#EN-S35**" />
<thes:NarrowerConcept rdf:resource="**#EN-N09.20**" />
<thes:NarrowerConcept rdf:resource="**#EN-N09**" />
<thes:NarrowerConcept rdf:resource="**#EN-N35/49**" />
<thes:NarrowerConcept rdf:resource="**#EN-N25**" />
<thes:NarrowerConcept rdf:resource="**#EN-Q48/59**" />
<thes:NarrowerConcept rdf:resource="**#EN-Q52**" />
<thes:NarrowerConcept rdf:resource="**#EN-N60/64**" />
<thes:NarrowerConcept rdf:resource="**#EN-N67**" />
<thes:NarrowerConcept rdf:resource="**#EN-N23**" />
<thes:NarrowerConcept rdf:resource="**#EN-N69**" />
<thes:NarrowerConcept rdf:resource="**#EN-N38**" />
<thes:NarrowerConcept rdf:resource="**#EN-N38.10**" />
<thes:NarrowerConcept rdf:resource="**#EN-N24**" />
<thes:NarrowerConcept rdf:resource="**#EN-N51**" />
<thes:NarrowerConcept rdf:resource="**#EN-N66**" />
<thes:NarrowerConcept rdf:resource="**#EN-N65/75**" />
<thes:NarrowerConcept rdf:resource="**#EN-S52**" />
<thes:NarrowerConcept rdf:resource="**#EN-R89**" />
<thes:RelatedConcept rdf:resource="**#EN-Q45/99**" />
<thes:ExactEquivalent rdf:resource="**#FR-N**">
<thes:inLanguageOf rdf:resource="**http://www.data-archive.ac.uk/Limber/ISO639#fr**" />
</thes:ExactEquivalent>
<thes:ExactEquivalent rdf:resource="**#DE-N**">
<thes:inLanguageOf rdf:resource="**http://www.data-archive.ac.uk/Limber/ISO639#de**" />
</thes:ExactEquivalent>
<thes:ExactEquivalent rdf:resource="**#SP-N**">
<thes:inLanguageOf rdf:resource="**http://www.data-archive.ac.uk/Limber/ISO639#sp**" />
</thes:ExactEquivalent>
</rdf:Description>

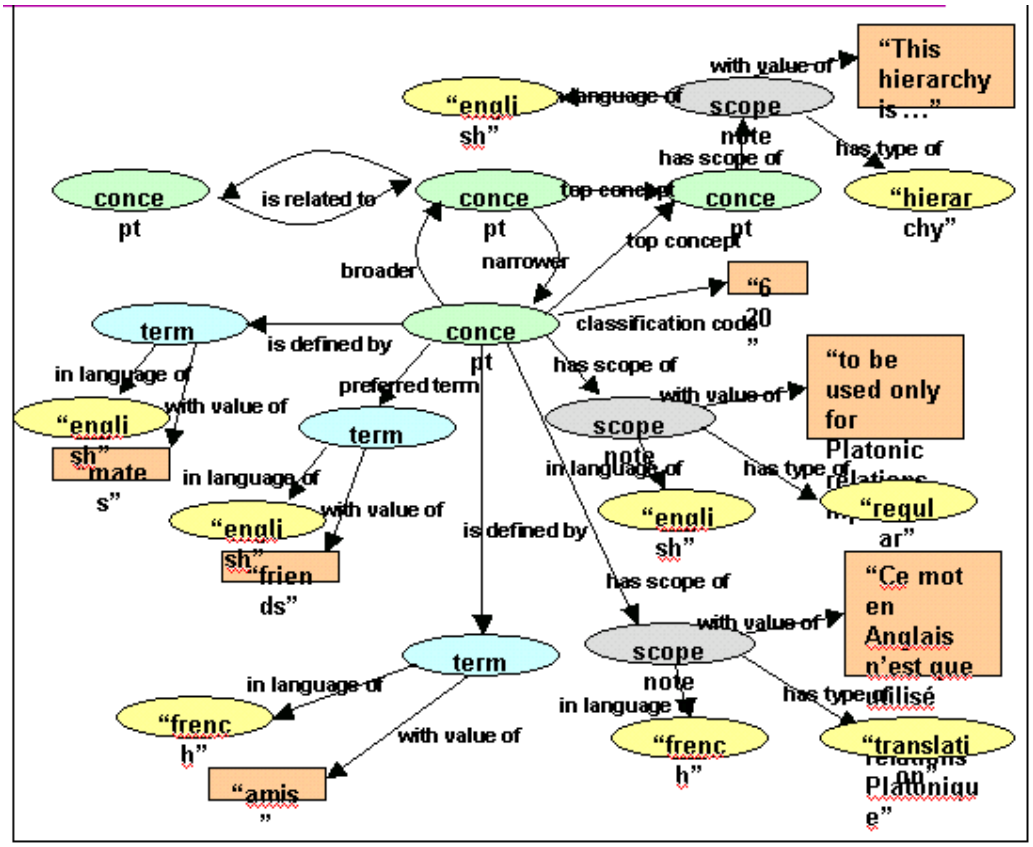


Figure: A graphical view of the RDF Thesaurus Interchange Format representation of the concept "Friend".

2) A metadata representation and format for the metadata entries for data sets held in social science data archives. The existing DDI XML metadata format V1.0 which is the standard used by Social Science archives across the world has been modified to include multiple language entries, keywords down to the question level, and pointers to restricted vocabularies. This proposal has been adopted by the DDI committee as V1.1. A proposal for a wider scope of metadata definition to include data cubes rather than just questionnaire data has been accepted by DDI and will be included in V1.2 for release in 2002. The longer term objective of an object oriented data model has been accepted by DDI in principle based on the RDF proposal from LIMBER and the detailed representation is under consideration as V2.0 of the standard for release in 2003. Deliverable D3 has been produced describing the metadata environment.

3) An internationalised user interface to a data archive system to access metadata, and to the multilingual thesaurus management system localised to English, French, Spanish and German languages. The existing NESSTAR data archive tool has had user interface files produced in the target languages, which were included in the public release of June 2001. The NESSTAR/FASTER client has been modified to use the LIMBER TMS to allow searching by keywords in each language, including the translation of queries between languages. The

NESSTAR/FASTER client is available for download by the reviewers from the LIMBER web site. A 15 minute demonstration is available for the review. WP7: Intermediate design deliverable D7.1 submitted March 2001, final deliverable D7.2 delivered December 2001.

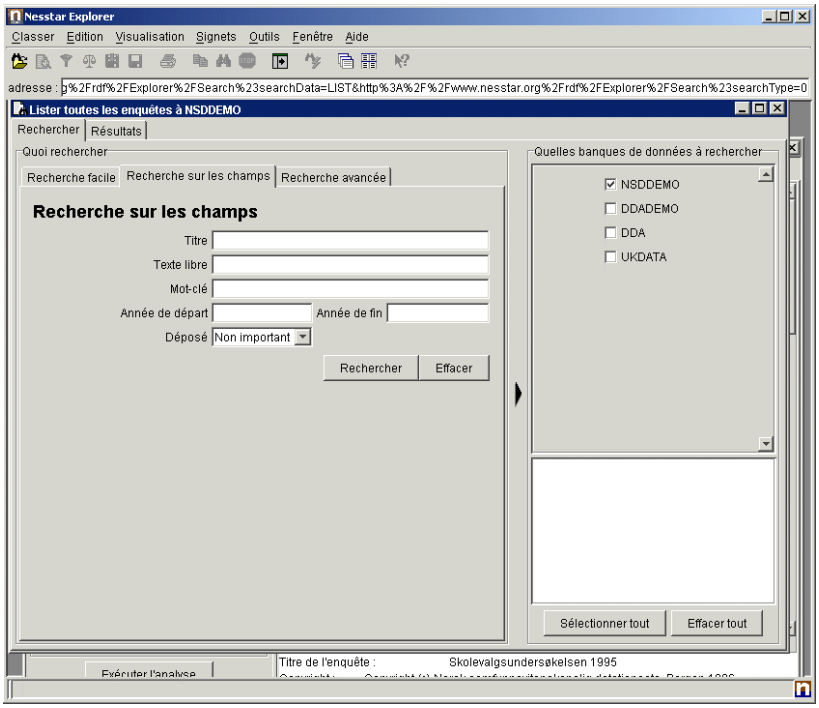
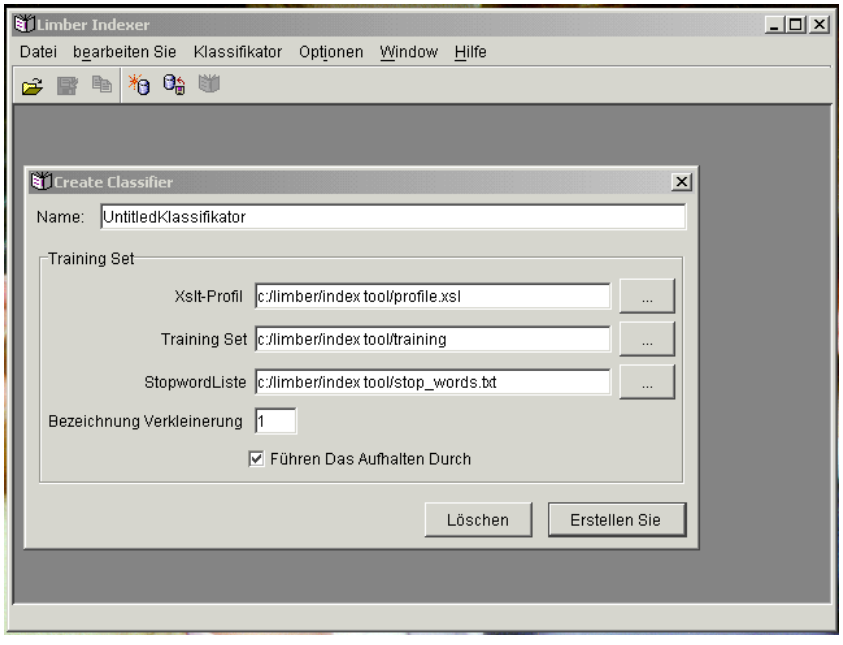


Figure: The NESSTAR search interface localised to French.

4) A text categorisation tool to automatically index metadata entries with keywords to allow keyword search on the metadata. The tool has been implemented and evaluated. Due to installation problems it is not currently available for download. A 5 minute demonstration of the semi-automatic indexing capabilities will be available for the review. WP8: Intermediate design deliverable D8.1 submitted March 2001, final implementation details and user guidance deliverable D8.2 submitted December 2001.



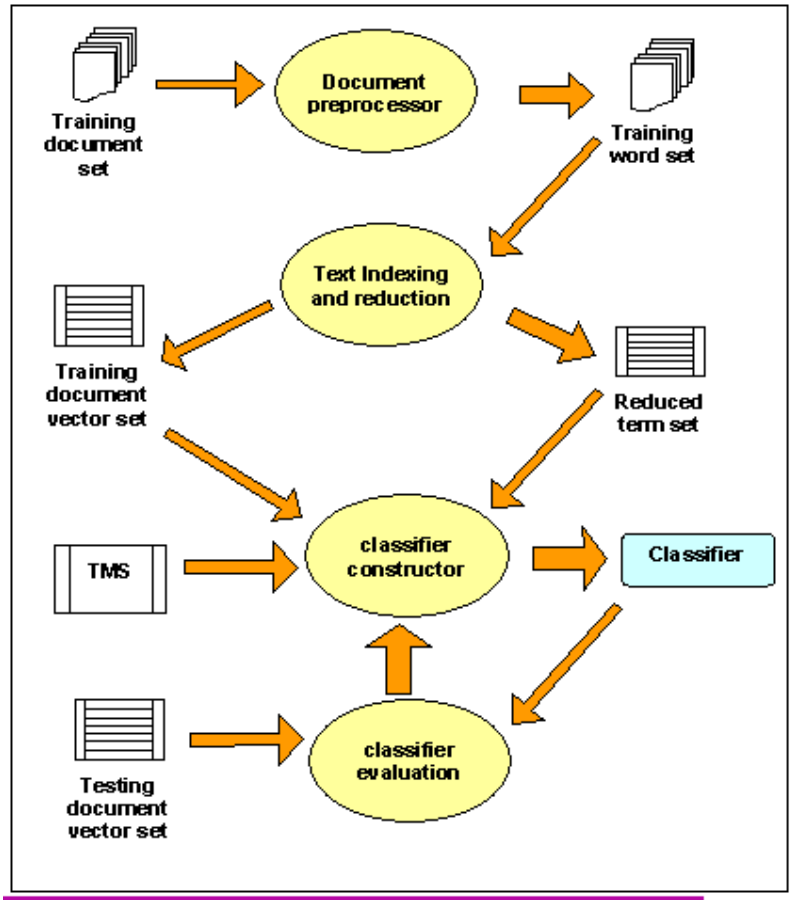
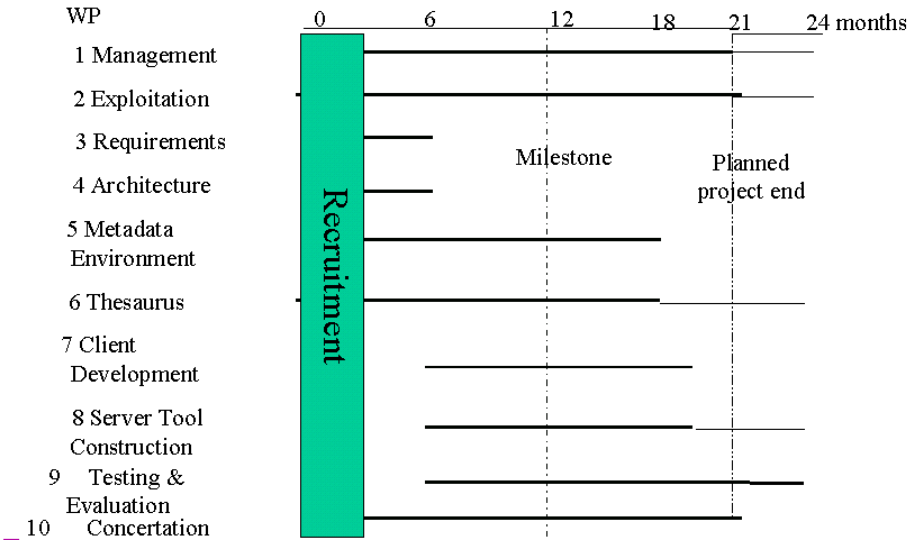


Figure: The flow of training and test sets of data to produce a classifier using the LIMBER indexing tool.

5) The Thesaurus Management System (TMS) to store and serve the thesaurus from a server to a client, and with an API to allow access from other clients. The implementation has been completed based on a previously existing TMS (SIS) which was as the basis of the LIMBER TMS. The user and administrator clients using the server are available for demonstration at the review, in addition to the NESSTAR/FASTER client which also uses the server. WP7 & WP8: Intermediate design deliverables for client and server D7.1 and D8.1 submitted March 2001, final deliverables D8.1 and D8.2 submitted December 2001.

5.1 Conformance to the contract



Some of the milestones named in the contract have changed their names and scope as "a prototype of multilingual query, retrieval and explanation tools" has become the multilingual NESSTAR interface to the thesaurus management system and the existing NESSTAR data archive tools; the "prototype of construction and maintenance tools for data archives using metadata representations" have become the indexing tool. Both NESSTAR and the Cheshire repository have been developed in parallel to the LIMBER project through the EU funded IST project FASTER and an NSF funded project at Berkeley, respectively. The resulting system still provides the overall functionality required by the LIMBER contract.

5.2 Deliverables

All paper deliverables are available from the LIMBER web site for download - http://www.limber.rl.ac.uk/

5.3 Evaluation and User Involvement

Evaluation activities in WP9 have included a workshop at the international IASSIST Conference in Amsterdam in May 2001 where the international user community were presented with ELSST, and descriptions of the supporting tools, and used the thesaurus and evaluated it. A second workshop addressing a wide audience of UK and international information retrieval specialists was held at the University of Essex in September 2001 where all the tools were presented and users could evaluate them. Detailed evaluation of performance and technical aspects of the tools have also been undertaken and are reported in D10.

Users representing the European social science data archives, and some of their end users through a user group associated with the project have participated in the requirements collection, in refining translations, and were involved in both evaluation workshops.

As dissemination and publicity activities, the consortium have presented conference and workshop papers to existing conferences in the area of usage: social science data archives and their users, government data producers; and in the following areas of technologies: software internationalisation, metadata, thesaurus construction and standardisation. Over 30 activities have been undertaken with 17 publications available on the public pages of the LIMBER web site.

5.4 Standardisation Activities

The standardisation activities in social science metadata (DDI), thesaurus construction (UNECSO and NKOS) and web based metadata (W3C RDF) have all had presentations of the requirements and technologies for the project to ensure that the work of the project is both consistent with their developments, and that we are leading them.

5.5 Project management

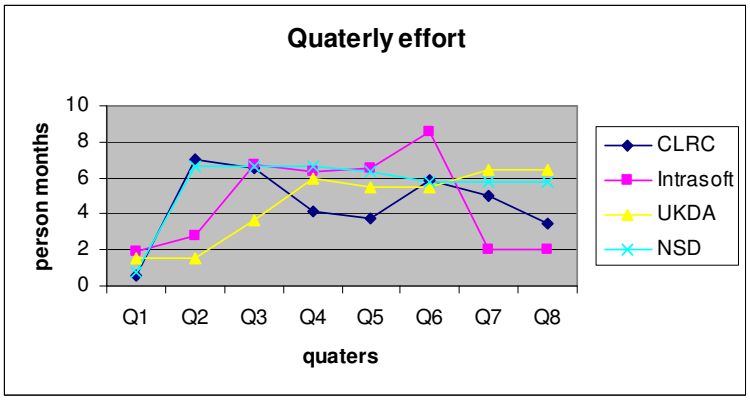
The major project management issue has been staffing problems. UKDA and CLRC have both suffered significantly since April from sick staff. One from each organisation have undergone major open heart surgery with the result that neither have been able to continue work on the project. The person responsible for translation at UKDA also left 6 months before the end of his fixed term contract, leaving the work there to be picked up by Anne Etheridge who took over as site project manager following Peter Littlewood's retirement on medical grounds. It was reported at the last review in April that the project was probably 3 months behind schedule, and these illnesses have caused further problems. Consequently, in addition to the two contract amendments requested in Feb 2000 that were finally completed by the CEC in August 2001, an extension of 3 months was requested and granted to the project until Dec 2002.

The second management problem has been cash flow, in that until the granting of the contract amendments in August 2002, Intrasoft International could not be given any funding, or have any of their cost statements processed. Once these amendments were granted, all payments that could be made, were including the advance payment and the payment of the first claim in August 2001.

The result of these management problems is that UKDA are about 2.5 person months under-resourced by the end of the project, although other partners have slightly exceeded their effort allocation and some travel costs have been higher than budgeted so the project should be approximately on budget as a whole.

	Total	Limit	remaining
CLRC	36.4	36	-0.4
Intrasoft	36.92	36	-0.92
UKDA	36.365	39	2.635
NSD	44.11	44	-0.11
total	153.795	155	1.205

Effort expenditure at the end of the project.



Effort expenditure profile by partner during the project.

6. Users

The consortium has established a user group with representatives of the major European social science data archives – UK, Netherlands, Germany, France, Spain, Austria, Finland, Denmark, Greece.

The user group are keen to adopt the thesaurus and supporting tools in their own archives, although the resources available to them for this vary considerably.

User group representatives and some of their end users have participated in the requirements collection by attending a workshop in Colchester (UKDA) in March 2000. Representatives have also been active in refining translations. They have also been involved in evaluations at the workshops arranged for 15th May in Amsterdam, and September in Colchester. In addition, representatives from some countries are also keen to write their own user interface translation files in their own languages, and to translate the thesaurus into their own languages not covered by the LIMBER project, given the example set by the project.

The European organisation CESSDA is mainly represented on the user group, but they have also had access to the project through their own meetings and are keen to adopt ELSST (as reported above) and are considering the general adoption of the LIMBER & FASTER technologies beyond the UK and Norwegian archives that are currently using them.

7. Objectives

The objectives of the project were to:

1) Research and develop a metadata model and representation (in XML and RDF) for social science datasets to allow their integration within and across data archives. This has directly been addressed by the RDF metadata proposals to DDI.

2) Develop and evaluate a multilingual thesaurus to be used to index and access social science datasets in data archives. This has directly been addressed by the development of the ELST thesaurus from the English HASSET one.

3) Develop and evaluate a multilingual query tool drawing on the multilingual thesaurus to allow queries to dataset archives to be made in several languages. This has been directly addressed by the development of the interface to the NESSTAR query tool to the multilingual thesaurus management system.

4) Develop and evaluate a multilingual retrieval tool drawing on the multilingual thesaurus to allow keyword and phrase translation in retrieved datasets into several languages. This has been directly addressed by the development of the interface to the NESSTAR query tool to the multilingual thesaurus management system.

5) Develop and evaluate a multilingual tool to assist queries and retrieval with explanations that draw on the metadata to guide the users in their task. This has been directly addressed by the development of the multilingual thesaurus management system, and the interface to it, that are used by its own stand alone client, and by the NESSTAR query tool to the multilingual thesaurus management system. The use of both clients shows the generic nature of the interface provided.

6) Develop and evaluate tools to support the construction and maintenance of datasets in an archive using automatic indexing drawing on the multilingual thesaurus. This has been directly addressed by the development of the indexing tool for metadata.

7) Evaluate an XML metadata server using the metadata model and representation for social science datasets and the multilingual thesaurus. We have evaluated alternative servers to the existing Cheshire one, but none provides better functionality.

It should be noted that the indexing tool, TMS server and clients, the RDF Thesaurus Interchange Format are generic technologies not tied to any outside LIMBER and not tied to any domain. While the Metadata proposals to DDI and the ELSST thesaurus are tied to the domain of Social Science, and the developments of the NESSTAR/FASTER client to support LIMBER is clearly tied to that data access system the methods used and lessons learned from developing these can be addressed to other domains and tools. Therefore although it is easier to describe the LIMBER system in concrete terms as linked to the NESSTAR/FASTER data access system and using the ELSST thesaurus to access Social Science metadata stored in the DDI format it remains a generic technology for supporting multilingual access to data through the use of multilingual thesauri.

8. Problems Encountered

The project has had problems with staff illness that has caused delays. A three months extension was requested and granted to allow for these.

The project has not found participation in clustering and concertation activities easy, although we did correspond with the CLASS support action on Cross-Lingual

Knowledge Managment, but received no replies. There has clearly been a great deal of interaction with the FASTER project since both it and LIMBER are developing the existing NESSTAR data access system tools.

Intrasoft have been replaced by Intrasoft International following a contract amendment. The same staff are working on the project in Athens, but the move was part of a corporate restructuring by Intrasoft as described above under management issues.

9. Planning

Outline changes or development of plans (development, verification, validation, exploitation, dissemination, promotion etc.).

The consortium decided not to produce a project brochure as stated in the exploitation workpackage description in the TA, since the web site contains the same information and is updated, while brochures quickly age.

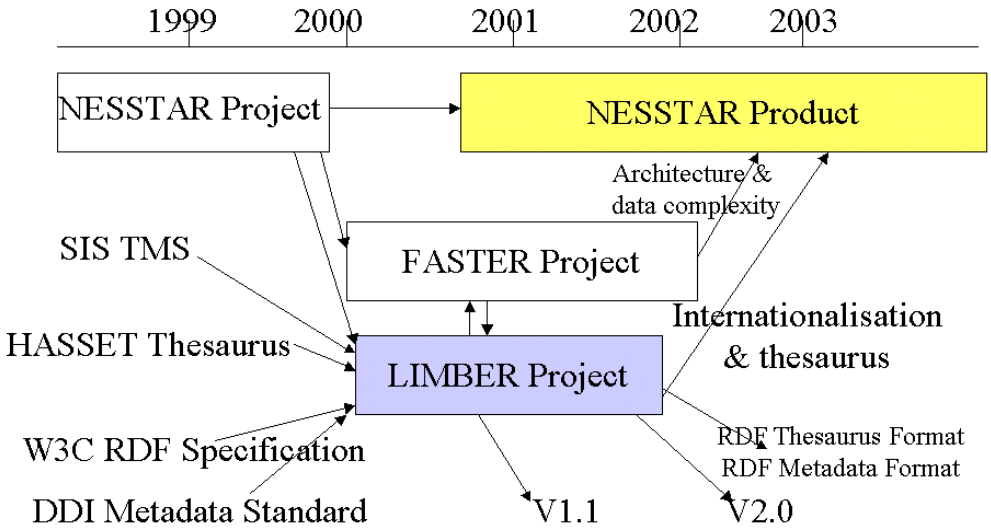
Risks identified and actions to prevent these:

1) No significant take-up of RDF. The ELSST thesaurus and the metadata standard use RDF as a representation language. If RDF does not have sufficient take up, these can be trivially translated into alternative XML based representations since RDF is built on XML and the rich models behind these define them thoroughly.

2) Failure of social science community to assimilate NESSTAR tool. NESSTAR has been used as the data archive tool for LIMBER, but the thesaurus management system has a loose rather than a tightly coupled interface to this to allow it to operate as a stand alone tool or as part of alternative data archive tools. Considerable effort has been put into involving the social science archive community in LIMBER and NESSTAR to ensure their adoption and assimilation of them.

3) Even if RDF is adopted for restricted vocabularies/ontologies and metadata descriptions, the power of this approach to data access can only be fully seen when multiple domains use interoperable standards. The move towards interoperable descriptions across domains is slower than was hoped when the proposal was written in 1998. The consortium has been active in W3C and the GRIDS initiative to promote interoperable take up, and has persuaded the IST programme to make the semantic web a topic for the next call for projects. These moves may produce more momentum behind interoperability.

10. Exploitation



The LIMBER project did not arise originally from nowhere, but had a history on which it built, and it does not finish off the work with it - the work and technologies developed continue after it.

A TIP was produced in July 2001, and a final TIP in Dec 2001.

A consortium agreement has been signed and is available on the LIMBER web site for review.

For exploitation each of the five products will be exploited independently, as well as collectively.

The whole multilingual access system will be available to the NESSTAR company to exploit along with the existing metadata and data access mechanisms to provide multilingual international access. A contract with NESSTAR has been drafted and apart from some minor debate about the exact licensing terms should be signed very soon. A draft of the contract is available on the LIMBER web site for review.

The thesaurus, thesaurus management system and indexing tool will each be independently exploitable. The TMS is an improved version of the SIS TMS whose original owners are the University of Crete/FortH. Intrasoft International have a long standing IPR agreement with them covering its adoption by LIMBER for research purposes and its long term commercial exploitation. Intrasoft International expect to commercially exploit the TMS as part of their usual systems consultancy business.

The indexing tool is the product most close to a research topic. The text categorisation techniques included are fairly standard, but in this area considerable tuning of specific algorithms is required to the training, test and real data sets, as well as to the exact usage patterns. It is planned that CLRC will take the indexing tool further into other research projects, and incorporate it into the European GRID technologies as well as the W3C semantic web initiative. These outlets will provide potential uptake, adoption and hopefully assimilation of the technology, however, they will not provide

immediate financial return. Financial return is only realistic from consultancy and training course activities for these technologies. In that case the return is hard to partition down to this particular technology since it will be one of several incorporated in those activities.

The metadata standard has been placed in the public domain (through DDI) to maximise adoption and assimilation by the social science community – any commercial exploitation would limit adoption, and therefore limit the adoption of the other tools dependent on it.

The RDF Thesaurus Interchange Format has been made available to the public on the LIMBER web site. It will be taken into another IST project with W3C starting in April 2002 to be used in a demonstrator. It is planned to further its standardisation through W3C and other bodies.

The long term economic and societal impacts are hard to predict with any precision. Clearly, the adoption the NESSTAR technology incorporating the LIMBER multi-lingual interface by the European social science data archives will provide access to European social science data to public and private decision makers and planners which should provide a basis for evidence based decision making. The impact of this could be significant.

As globalisation of the knowledge based economy increases multilingual international access to data is clearly required and provides a market for multilingual thesaurus management as a key technology. Intrasoft International currently have considerable European consultancy work which is expected to require this technology, and provide a realistic market sufficient to justify investment.

Metadata standards and restricted vocabularies have been promoted as a solution to the heterogeneous database access problem for 20 years, but they have had little significant impact yet due to lack of standardisation and assimilation, despite considerable, disparate adoption. W3C has been successful in having HTML and XML adopted as standards by the community for information and data description. The generic adoption of metadata standards, restricted vocabularies through the W3C RDF standard could cause dramatic changes to the global economy, although not as significant as the 3% per annum US growth in GDP resulting from e-commerce based on HTML and XML technologies in the late 1990's. There is considerable interest in RDF, demonstration systems have been produced in various areas, and books are in press, but given the real industrial memory of expert systems and previous knowledge representation language standardisation efforts, assimilation barriers may exist that will not be overcome – it would be naïve to claim otherwise.

11. Project Publications.

- Brian Matthews and Michael Wilson, Multilingual metadata to access social science data, Proceedings of the Data Management Workshop, Daresbury Lab., 23-25 February 2000. Presentation materials.
- Brian Matthews and Michael Wilson, Multilingual metadata to access social science data, ERCIM News, January 2000.
- Ken Miller,"Framing the Future II" - UKDA workshop - LIMBER presentation - 27-28th march 2000.
- Ken Miller,"IASSIST 2000 - Data in the digital library conference" - LIMBER presentation - 7-10th June 2000
  N.B. IASSIST - International Association for Social Science Information Service & Technology
- Ken Miller,"DDI committee meeting" - committee member - agenda item RDF (Resource Description Framework)
  N.B. DDI - Data Documentation Initiative, producers of the XML codebook metadata standard used by CESSDA (Council for European Social Science Data Archives) members. (8th June 2000)
- Ken Miller,"ECASS Workshop" - LIMBER presentation - 24th June 2000.
  N.B. ECASS - European Centre for Analysis in the Social Sciences.
- Miller, K., LIMBER: A blind date with data, UK Data Archive Bulletin, No 74 May 2000, 6-7.
- Jan-Dec 2000, Public Annual Report
- Ken Miller, Brian Matthews **Having the right connections: the LIMBER project**; *Journal of Digital Information* (2001) (http://jodi.ecs.soton.ac.uk/)
  Keywords: metadata, thesaurus, multilinguality, data archives, RDF
- Ken Miller,NKOS workshop - ECDL 2000 - LIMBER presentation - 18-20th September 2000, Lisbon. N.B. NKOS – Networked Knowledge Organisation Systems and ECDL – European Conference for Digital Libraries.
- Ken Miller,“New Horizons” – LIMBER presentation and display boards at UKDA Launch of New Services – 10th November 2000, London
- Ken Miller,DDI workgroup meeting" - Metadata/ RDF meeting - 26th November 2000, Washington N.B. DDI - Data Documentation Initiative, producers of the XML codebook metadata standard used by CESSDA (Council for European Social Science Data Archives) members.

- Ken Miller,UN/ECE Work Session on Statistical Metadata (METIS)" - Thesaurus/LIMBER presentation - 28-30th November 2000, Washington.
- B.M.Matthews, K. Miller, A.Ramfos, J.Ryssevik, M.D.Wilson, Internationalising data access through LIMBER, in D.L.Day and l.m.dunckley (eds) Designing for Global Markets 3: proceedings of iwips2001, pgs 129-142, Open University: Milton Keynes. Presentation materials
- Ken Miller, UNESCO Thesaurus Working Party Meeting 21st February 2001, held at ULCC
- Ken Miller, LIMBER Workshop, IASSIST Conference, Netherlands 15th May 2001
- Myriam García Bernabé , Se busca: The Perfect Language. LIMBER as a 'metaboundary tool', IASSIST Conference, Netherlands, May 2001
- Matthews, B.M., Miller, K. and Wilson, M.D. (submitted) A Thesaurus Interchange Format in RDF. Submitted to the Semantic Web Conference 2002.

12. Deliverables Produced

ID No. (1)	Name/Title/Description	Version (2)	Date completed	Current status (3)	Classification (4)
D1	Requirements	1	31/7/2000	A	P
D2	Architecture	1	31/7/2000	A	R
D3	Metadata Environment and Report	1	30.11.2001	A	P
D4.1	Thesaurus Description	1	31/3/2001	A	P
D4.2	Multilingual Thesaurus (Report only)	1	30.11.2001	A	P
D7.1	Client Design	1	31.3.2001	A	R
D7.2	Client Implementation Notes	1	30.11.2001	A	R
D8.1	Server Design	1	31.3.2001	A	R
D8.2	Server Implementation Notes	1	30.11.2001	A	R
D10	Evaluation Report	1	21.12.2001	A	P

(1) the deliverable identification used in the project's Technical Annex
(2) version information required to identify the most recent version of a deliverable
(3) status of a deliverable can be D = draft (internal circulation only to date), A = accepted, R = rejected (for modification), W = waiting for comment from Commission.
(4) classification P = Public, R = Restricted, C = Confidential

13. Other Project 'Products'

ID No. (1)	Name/Title/Description	Version (2)	Date completed	Classification (3)
	Private and Public Web Site	n	Dec 2001	P
	Consortium Agreement	1.0	Nov 2001	R
	Dissemination and use Plan	1.1	July 2001	R
	Technology Implementation Plan	1.1	July 2001	R
	TMS Server & Client Software	1	Dec 2001	R
	ELSST Thesaurus Linguaware	2	Dec 2001	R
	RDF Thesaurus Interchange Format	2	Sep 2001	R
	NESSTAR/FASTER Client Software	3	Dec 2001	P
	Indexing Tool Software	2	Dec 2001	R
	Final Project Report	1	Feb 2002	P
	Final TIP (D11)	2	Dec 2001	R

(1) the deliverable identification used (where appropriate)
(2) version information required to identify the most recent version; (3) classification P = Public, R = Restricted, C = Confidential