

**MACS (Multilingual access to subjects)
A virtual authority file across languages.**

Genevieve Clavel-Merrin
Swiss National Library

Introduction: authority control comes of age

Shared authority files and co-operation in the development of national lists, both author and subject, have enabled libraries to share resources and improve access to their collections. As we move from national resource sharing to a more international approach (for example with the TEL (The European Library) project (<http://www.europeanlibrary.org>)) the difficulties of searching across collections have been underlined. Variations in author and corporate author headings across languages and subject headings in different languages have meant that frequently the search option of choice has been keyword. This brings results but its limitations are well—known.

Attempts are now being made to build on the significant work carried out in many countries to build and maintain subject headings languages and authority files. This is possible for a number of reasons, both organisational and technical: the increasing use of standard communication formats, the preparation and maintenance of well-structured manuals and the development of sophisticated distributed management and updating structures built upon the widespread availability of the networking environment enabling the institutions increased access to the centralised authority control systems developed previously. As a result the application of authority control to collections using the same subject heading languages is facilitated as libraries can exchange and copy records and thus improve (monlingual) indexing consistency across collections.

In addition, recognising that many subject heading languages follow similar construction and application principles (see the IFLA *Principles underlying subject heading languages*¹), the groundwork is laid to allow libraries to build on the solid infrastructure for authority control at a monlingual level in order to extend control and access at a multilingual level.

Whereas traditionally, bilingual/multilingual access was assured through multiple subject headings coded in the authority records, today many research initiatives seek to provide access through Internet based linking processes. Concepts such as language cross-linking, or interoperability in subject headings languages, which are common in today's research in multilingual access were limited ten years ago to language specialists.

MACS uses the stable authority control environment and this linking process to extend the use of subject heading languages on an international networked multilingual level.

Background to the MACS initiative

MACS (Multilingual access to subjects) was created in 1997 in response to a request by the Conference of European National Librarians (CENL) to find a solution to the problem of multilingual subject access to European databases. As a multilingual country, Switzerland is

¹ Maria Inês Lopes and Julianne Beall, ed., *Principles Underlying Subject Heading Languages (SHLS)*. Working Group on Principles Underlying Subject Heading Languages, IFLA Section on Classification and Indexing (München: K.G. Saur, 1999).

of course particularly interested in finding solutions, but it is significant that other countries were interested in this multilingual approach even if they have traditionally been considered monolingual. This serves to underline the international needs of users and the necessity to co-operate in collection building and access not only on a regional or national level but increasingly within the international sphere.

A working group was set up in order to discuss the issue and among many approaches discussed, the idea of establishing links between different subject heading languages (SHLs) was, for four of the participating libraries, a promising solution. These four libraries, the Swiss National Library (SNL), the Bibliothèque nationale de France (BnF), the British Library (BL) and the Deutsche Bibliothek (DDB) accepted the task of defining a concept and conducting a study to determine the feasibility of such a concept, details of which are given below. The MACS project was then set up with its main goal:

“to provide the means by which library databases can be accessed on a multilingual basis by the use of equivalent headings from subject authority files”. The approach taken was to create broad multilingual access that could be used as widely as possible by using indexing languages that are commonly used in most libraries in any given country.

This multilingual access is made possible thanks to the equivalence links established between the three Subject Heading Languages used in these libraries: SWD/RSWK for German, RAMEAU for French and LCSH for English.

Each of these Subject Heading Languages is widely used at a national and international level e.g. the use of SWD is widespread in Germany but also in Austria and Switzerland, RAMEAU is used in university, municipal and public libraries in France and is being taken up by other francophone countries, while the spread of LCSH in English-speaking countries and in international institutions is well-known.

The decision to use the approach of linking also took into account the fact that the libraries (French, German, English) have invested considerable time and effort in the creation and maintenance of each SHL (LCSH, RAMEAU and SWD) and which have been used extensively to provide subject access to millions of documents, both in the national libraries but also as noted above, among libraries throughout the countries concerned and elsewhere. Creating a new multilingual vocabulary, based on translation, would have been unrealistic and uneconomical. Using the resources available maintains access to documents already indexed and extends access across linguistic boundaries.

The study took a subset of headings from the three SHLs in the fields of Theatre and Sport, plus the most frequently used headings used to index documents in the Bibliothèque nationale de France. The group studied different mapping methods including the comparison of indexing of the same document in different subject heading languages before establishing links between the headings manually. Methods to solve one-to-many or many-to-many mappings were analysed and principles of co-operative link creation and management were established. The results of the study were encouraging and showed that linking headings was a feasible approach to cross-language access to documents in many though not all cases. A match or link was considered successful when a concept, represented by similar headings in the different SHLs which are matched manually (intellectually) returned almost equivalent results in subject retrieval.

Details of the feasibility study may be found at: <http://infolab.kub.nl/prj/macs>. The group recommended that a prototype be developed in order to test the management structure, link creation and multilingual searching.

Prototype

Following a restricted tender, the contract to develop the system was awarded to Index Data aps and Infolab Tilburg.

The system is available for guest access at <http://infolab.kub.nl/prj/macs>.

The prototype contains the following data:

- All headings from the subject areas *Sports* and *Theatre* for which terms had been matched in the feasibility study, plus 500 most used headings identified first in the RAMEAU database and their equivalent headings from the other SHLs amounting to around 3000 headings;
- 1200 links created between the concepts
- Bibliographic records containing these headings extracted from each of the 4 partners' databases in order to allow the system to simulate searches of the library catalogues using the Z39.50 protocol.

The prototype is a standards-based Web application independent of the partners' systems and the SHLs used. It has two modules:

- The ***Link Management Interface*** enables link creation and maintenance, with mechanisms for managing partial or complex links. The system is based on the principle of 'federative management' i.e. each SHL is autonomous with each partner being responsible for link management in their own SHL (shared responsibility for SWD between DDB and SNL) and no central editor. Partners may make proposals concerning other parts of a link but these must be confirmed by the appropriate partner. An annotation function enables partners to comment on links or proposals or to indicate where problems have occurred. The following principles are observed in link management:
 - All SHLs are of equal status; there is no pivot SHL. Instead, the system contains mappings that are considered as conceptual clusters, identified by an abstract (numeric) identifier only;
 - Hierarchical structures, thesaural relationships or non-preferred terms are not mapped or reproduced as part of the process of linking individual headings: only headings at the authority level are linked.
 - Where an equivalence cannot be found a proposed heading should stand alone in the system to represent the concept (for possible future mapping)

The ***User Search Interface*** enables end users to carry out subject searches in the language of their choice simultaneously across target catalogues (currently subsets of the partners' catalogues) using the Z39.50 protocol. The user enters a search term from one of the available Subject Headings Lists (currently LSCH, RAMEAU or SWD), chooses one or more target catalogues then may either launch a search across the targets or choose to browse a list of corresponding headings and their equivalents from which one or more clusters may be chosen before the search is sent out. Tests have shown that this latter approach leads to better results and puts less strain on target resources. From the search result set the user can access individual bibliographic records in any of the libraries by simply clicking on the title. In the prototype options are provided for a brief display or for a full display in the MARC format used by the selected library.

Tests

MACS has made searching and a read-only access to the Link Management Interface available since 2001. During that time, over 500 people from all over the world have signed up to test the MACS prototype. With comments received and tests conducted by the MACS

partners, refinement of the organisational model has resulted in the design of a new version of the MACS system. Despite improvements in the linking mechanisms, the basic task of establishing links will remain very time consuming. Luckily the partners can build on existing bilingual links in the RAMEAU authority files managed at the Bibliothèque nationale de France and loaded in the MACS system. Approximately 70'000 RAMEAU authorities contain a proposed LCSH heading and have been assigned a subject domain number. Work on adding the SWD headings to these links will start in 2003 and will be divided between partners in Germany and Switzerland according to domains. In the meantime, other methods to facilitate link creation will be studied.

Integrating MACS into real-life environments

The search results have shown that pertinent documents can be retrieved using links across subject headings and a move from using subsets of data to the partners' own databases is planned. In April 2002 tests began with the catalogue of the Deutsche Bibliothek, in Summer 2002 the Bibliothèque nationale de France carried out successful tests and the Swiss National Library will carry out a similar test in Winter 2002. Changes in the Search Interface will be necessary to ensure that users can obtain manageable queries and search sets and of course an increased number of links is essential. The interest of the library community in MACS is underlined by the fact that the national library partners in the TEL project wish to integrate the product in the TEL portal which in its prototype phase aims to offer access to the resources of 8 European national libraries. Portal suppliers will be asked to propose ways in which MACS could be integrated, at the simplest level by a link to the system or perhaps by taking over the MACS data itself.

Incorporating true authority management (see and see also references)

As initially planned, only headings at the authority level are linked. The partners recognise that it is not always easy for users to find the preferred terms for searching without the help of see-and see also-references. The goal is to incorporate an access to the national authority files as an option for users who wish to check headings and refine their search before sending it to the target databases.

In addition, no SHL remains static, headings may be altered in response to linguistic change or new discoveries. It is essential that such changes be carried over into the Link Management System (and also of course into the appropriate library catalogues). Where a straightforward replacement occurs it should be possible to update the Link Management System: this will require the SHL providers to be able to track and extract changed headings and their ID and the link Management System to load and match these changes.

Since the Link Management Interface contains the authority ID for the headings, the links for navigation and updates when approved headings are modified can be incorporated without the need to duplicate information.

The future

The immediate value for users of the Swiss National Library (SNL) is the real possibility that in the near future subject searches will be conducted in more than one language. As a federal institution, the Swiss National Library is committed to serving its users in the three official languages of Switzerland (German, French and Italian). The SNL's decision in 1998 to adopt

the SWD as its indexing standard came with the assurance that the French and Italian users would also profit from that indexing through a multilingual search tool. Going beyond national boundaries, other users of the SNL and of other national libraries will also find their searches facilitated by a multilingual gateway to their catalogs. The Deutsche Bibliothek views the project as a way to foster the value of the SWD and will make the MACS results available to its partners in Germany and Austria. The Bibliothèque nationale de France has already undertaken a survey to look into ways of extending the MACS linking work at a national level.

Extending the links to other Subject Heading Languages is a goal, though future partners should be aware that the investment in staff time can be considerable since at present linking is still analyzed and carried out manually. Until the operational system is up and running it is still difficult for the partners to estimate the exact time required for link creation and checking but preliminary estimates show that the time required per link may vary between 2 and eight minutes depending on the subject area and the complexity. Any potential partner therefore should have an infrastructure similar to those in France or Germany in which the management of the Subject heading Language is spread across participating institutions.

A virtual authority file?

At present MACS is restricted to the authorised headings from the participating Subject Heading Languages, and as Freyre and Naudi² have indicated could be considered in some respects to play the role of a *multilingual dictionary of Subject Heading Languages*, a ‘bridge between systems that are themselves designed to organise and name concepts’. The incorporation of a link to the other elements of the authority records within the partners’ national authority files will build on this to create a virtual multilingual authority file.

² Elisabeth Freyre and Max Naudi, “MACS: Subject Access Across Languages and Networks”, in *Subject Retrieval in a Networked Environment: Papers Presented at an IFLA Satellite Meeting Sponsored by the IFLA Section on Classification and Indexing & IFLA Section on Information Technology*, OCLC, Dublin, Ohio, USA, 14-16 August 2001. Dublin, OH: OCLC, 2001