

Connecting with users : Europe and multilinguality

Dr. Britta Woldering
Assistant for International Relations to the Director General
Die Deutsche Bibliothek (German National Library)
woldering@dbf.ddb.de

Abstract:

This paper introduces to the new Internet service The European Library, provided by the Conference of European National Librarians (CENL), and gives an overview of activities in Europe for multilingual library services, developed and tested in various projects: TEL-MEMOR, MACS (Multilingual Access to Subjects), MSAC (Multilingual Subject Access to Catalogues of National Libraries), Crisscross, and VIAF (Virtual International Authority File).

Connecting with Users: Europe and Multilinguality

“Language is the foundation of communication between people and is also part of their cultural heritage. For many, language has far-reaching emotive and cultural associations and values rooted in their literary, historical, philosophical and educational heritage. For this reason the users’ language should not be an obstacle to accessing the multicultural heritage available in cyberspace. The harmonious development of the information society is therefore only possible if the availability of multilingual and multicultural information is encouraged.”ⁱ

The Internet is not monolingual and non-English content is growing rapidly as well as the non-English native speaking Internet users. The language share has completely changed in the last ten years: In 1996 the “online population” (=people having access to the Internet) was 40 million English native speakers and 10 million non-English native speakers while in 2005 the “online population” increased to 280 million English native speakers and as much as 680 million non-English native speakers.ⁱⁱ

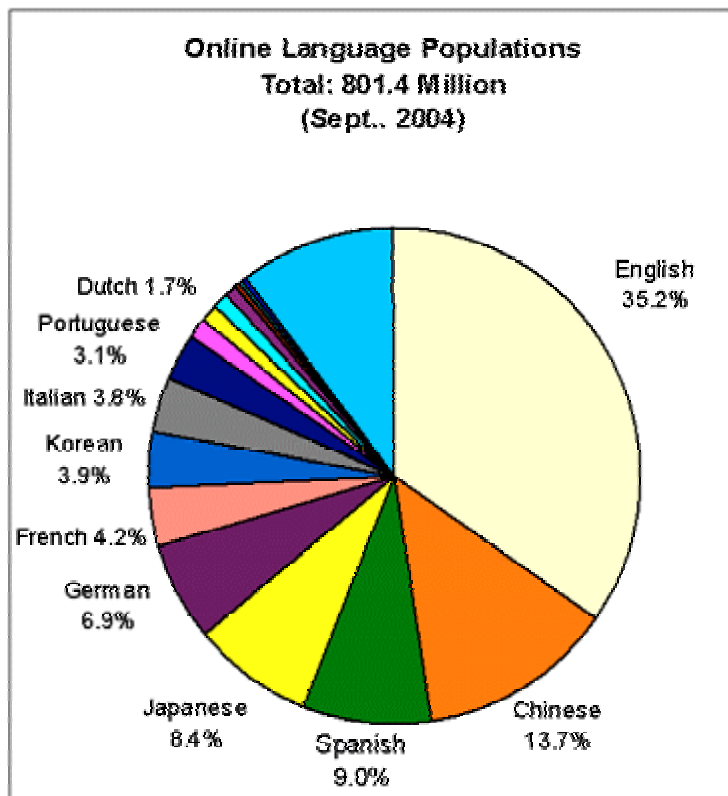


Fig. 1: Online Language Populationsⁱⁱⁱ

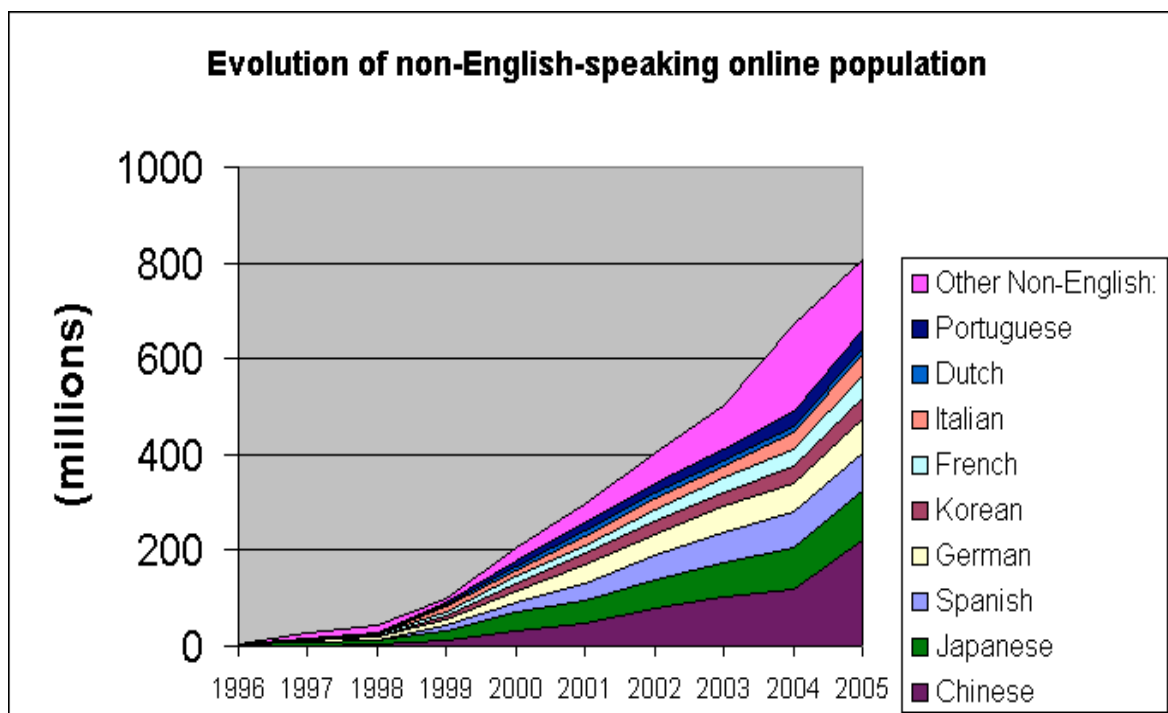


Fig. 2 : Evolution of non-English-speaking Online Population

Also the Internet user profile has changed radically: from primarily academic use to widespread commercial, leisure, educational, entertainment etc uses. Both developments - the change in the language share and the change of the user profile - mean increasing pressure for providing access to information without language or cultural barriers and create a need for multilingual information access. Libraries have never been monolingual, neither in their acquisitions nor in their user community. This is even more true since libraries are offering their catalogues and services via Internet.

Looking at the statistics of online linguistic populations given above, of course the Asian languages are growing rapidly, but also European languages gain a larger share in the worldwide online population. Some more statistics: based on the membership in the Council of Europe, Europe consists of 46 countries with a population of about 800 million people and 74 official languages^v (of which 24 languages are based in the countries of the Russian Federation).

How do we face this challenge of multilinguality in Europe?

Conference of European National Librarians (CENL)

Every national library of the member states of the Council of Europe is eligible to become a member of the Conference of European National Librarians (CENL, www.cenl.org). 45 libraries from 43 member states are members in CENL at the moment. CENL began in 1987 as an informal meeting once a year of a few western European national librarians but soon it became clear that these meetings for exchange of experiences, discussion of topics relevant for all national libraries and for tightening the relationship and cooperation on a personal level was very fruitful and the group grew steadily. After the breakdown of the communist world, eastern European countries gradually became members of the Council of Europe, and CENL also opened up and invited the national librarians to join. In 1999 CENL became a foundation under Dutch law and adopted its statutes.

The members continue to meet annually and strengthen their cooperation by appointing working groups on topics relevant for all national libraries and by carrying out projects together. Actually there are three working groups in force: one being in regular contact with the Federation of European Publishers (FEP / CENL working group), one dealing with performance evaluation and benchmarking for national libraries, and the third one working on persistent identifiers and keeping contact with the DOI Foundation, where CENL is corporate member.

CENL is an ideal platform for national libraries to find partners for projects and quite a number of projects with the participation of European national libraries were carried out (see "Related projects" on the CENL website). There are several funding models: Some of the projects were nationally funded, some were funded by the libraries themselves and some were funded by the European Commission. Actually four projects are carried out which will be introduced in more detail below: *The European Library* (TEL) service offering integrated catalogue search in the OPACs and collections of twelve and eventually of all European national libraries; *TEL-ME-MOR*, a project aiming at taking the national libraries of the ten new member states of the European Union on board The European Library and at raising awareness of opportunities for the new member states to participate in European projects within the Information Society Technologies (IST) programme of the European Commission; *MACS*, a project aiming at multilingual access to subjects by linking the English, French and German subject heading languages, and finally *MSAC*, a project mapping the verbal equivalents of the Universal Decimal Classification (UDC) in eight eastern European languages to the English version of the UDC.

Gabriel

In 1994 CENL decided that some form of online bulletin board for their libraries would be very useful as a means of keeping up-to-date with current developments and activities. The Internet offered various opportunities to create such a service. Taking the initial idea of a bulletin board a step further, it was agreed that a World Wide Web site on behalf of the entire membership of the CENL would not only be useful for the libraries themselves, but it could also provide the worldwide Internet community with an online guide and a single point of access to the networked services of the various national libraries.

The prototype version of Gabriel was developed and edited at the British Library and made available on four different servers located across Europe, and proved to be very successful. Within its first year, the service reached about 1,000 hits per day. In September 1996, after detailed information about all the participating libraries had been added, the directors of the national libraries in Europe decided that Gabriel should be launched as the official WWW service of CENL.

The mission of Gabriel was to provide information about Europe's national libraries, their collections and their services in order to facilitate access to them, and to foster the development of new services based on a shared infrastructure. The standardised Gabriel information pages were offered in three languages, English, French and German, provided by each library and edited by a language working group of native speakers.

When Gabriel was successfully established, the idea came up to enhance the service: not only offering standardised information about Europe's national libraries but also to offer an integrated search facility for the catalogues and collections of the CENL libraries. This was the birth of the idea to build a European virtual library.

The European Library: from project to service

As the technology and the metadata formats in use in the CENL community were (and still are!) different, it was not clear whether an integration of the systems to form a virtual European catalogue would be feasible. It was decided to carry out a feasibility study and a project proposal was submitted to the European Commission. The proposal was approved and in February 2001 The European Library (TEL) project started with 66 % funding from the European Commission. The core work packages of the project were the testing of the interoperability of the systems of the project partners and the development of a metadata application profile for TEL. A third work package elaborated a business model for the prospective service. TEL ran for three years and was successfully completed in January 2004.^{vi}

Project partners were the national libraries of Finland, Germany, Italy, the Netherlands, Portugal, Slovenia, Switzerland and the UK along with ICCU, the national cataloguing institute of Italy, and CENL.

The project results were presented at the CENL Annual Meeting and were that promising that the CENL Board agreed to support the transfer of the project results into a new service: The European Library, the public face of CENL. The beta version of The European Library service was launched in March 2005, deliberately not as a big bang but in order to leave the test environment and to be tested and commented by “real” users.

Functions

The European Library service consists of three parts: Search, libraries and treasures. The search part offers an integrated search and retrieval facility for the catalogues and collections of the full members of The European Library. The European Library service is organised and funded by membership. Only full members are providing their metadata for the search facility, at the moment twelve national libraries. Besides the OPACs of each national libraries a large variety of special collections is offered for searching. The default selection of catalogues comprises all the OPACs but the user is offered the possibility to chose from the complete list of all catalogues and collections the ones s/he wants to search in. At the moment more than 120 collections are available. A simple search form is offered directly on the homepage, but an advanced search is also available. The results are shown in an integrated list by order of the libraries. From the full data record the user is lead to an information page of the providing library for further information on how to order or receive the material s/he found.

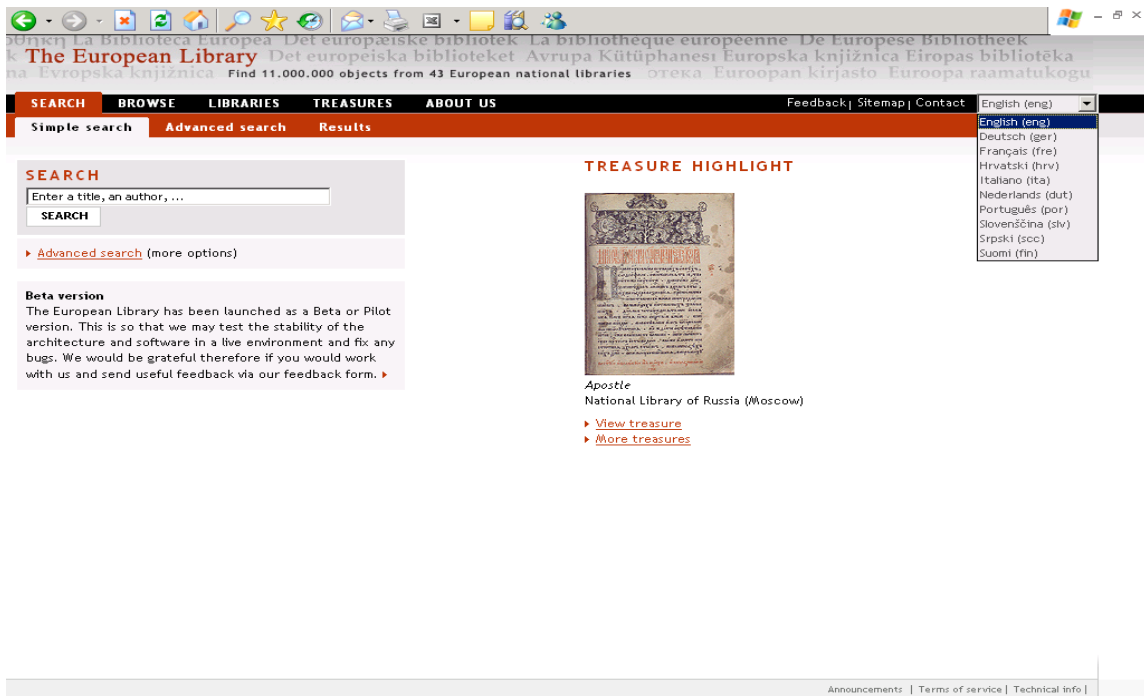


Fig. 3: Homepage of the European Library (www.theeuropeanlibrary.org)

The “libraries” part of The European Library is the former Gabriel service which was integrated completely. The user is offered a list of all CENL countries and clicking on one of the countries leads him / her to the standardised information pages of the respective national library. These information pages comprise history of the library, description of the collections, links to the OPAC and other online services as well as to the homepage of the library but also general information like address, phone numbers and contact persons.

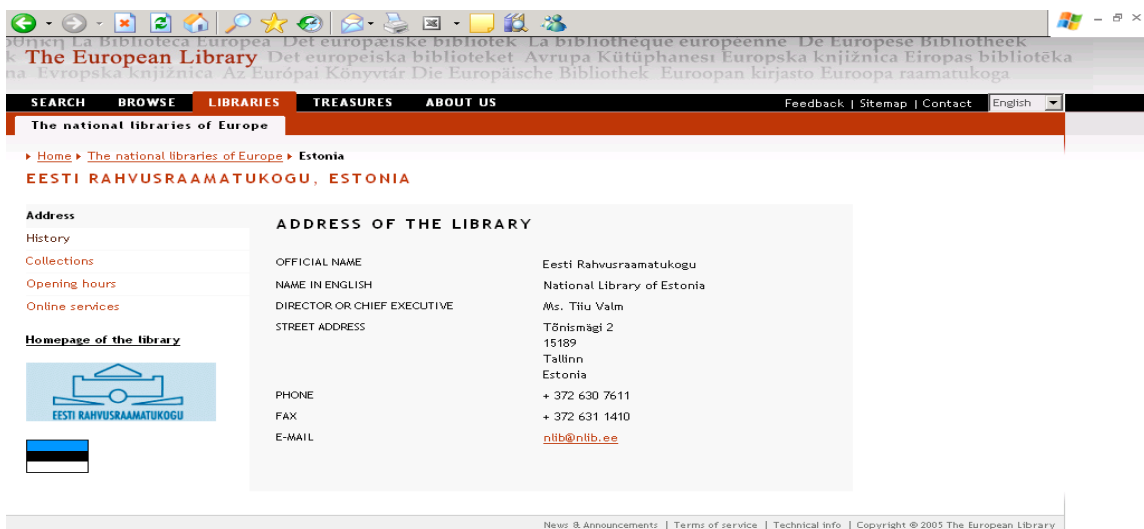


Fig. 4: “Libraries” part of The European Library

The treasures part of The European Library is a virtual collection of scans of special pieces from the collections of the CENL libraries – no full texts but only one picture per item accompanied by a short description in English. The treasures appear on the homepage at random: every time the page is opened another treasure is shown.

Multilinguality

Back to multilinguality: The European Library is a multilingual service in many aspects. First, the *search interface* is offered in all languages of the full members offering their metadata for the integrated search. At the moment The European Library has 12 full members and provides the interface in their 10 official languages: English, German, French, Dutch, Italian, Finnish, Slovenian, Serbian, Croatian and Portuguese. Every new piece of text – help texts, error messages, temporary messages etc – is provided as English master text which then has to be translated by appointed contact persons in each full member library into their own languages.

The European Library also aims at true multilinguality regarding the input of search terms and the display of the results. This means a permanent struggle with *character sets* to guarantee that differentiated input is possible and that the results are displayed properly with all umlauts, Háček, accents and other diacritics there are in European languages.

Multilingual search is not possible at the moment but there are several ongoing projects regarding multilingual search whose results will be implemented in The European Library in due course. An introduction to these projects will be given later in this paper.

The information pages on the national libraries in the “*libraries*” part of The European Library are provided in English, French and German – as the Gabriel service did. It is being discussed whether the pages will also be translated and maintained in Russian which would be very convenient for eastern European users. Every CENL library is obliged to submit standardised information for that part of The European Library in English. If the linguistic skills are available at the partners’ libraries, the submission of the information pages in German and French are appreciated. Native speakers in English, German and French edit the texts and guarantee the correctness of the texts and the updates.

Last but not least the multilinguality of the service is also reflected in the *logo*. The three words “The European Library” were collected in all languages of the CENL members and forwarded to the designers for use in the creation of the logo: Different languages appear and all three scripts in use in the CENL community were used in the logo.

Further challenges regarding multilinguality of The European Library

The European Library aims to keep the multilinguality of the search interface but it is not yet clear if this could be maintained with finally 45 partners and about 50 official languages.

Another challenge where a decision has to be taken is the question of whether to offer a *helpdesk*. Undoubtedly it would be a great service to offer a central questionpoint for information about and from European national libraries. But managing, maintaining and staffing a helpdesk is probably the hardest thing about a website. The European Library offers a feedback form on every page meant to provide the possibility to give feedback on the service, on things that do not work or are missing and to help The European Library Office to improve the service. People are making good use of this form but not only for giving feedback but also to ask questions such as How do I find 2nd World War newspapers? How

do I order the Botanists Almanac? It is bad enough if this is in one language, but covering a helpdesk in several languages requires the employment of linguists to respond. The European Library will check the possibilities of automated or decentralised solutions such as Questionpoint, but would then have to call again on the resources of the national libraries. In the course of 2006 a basic decision whether to establish a helpdesk at all will be taken.

Organisation and funding

Finally a short description of the business model for The European Library, the organisation and funding of the service. The European Library service is owned by CENL. As contracts have to be signed, e.g. for staffing The European Library Office, for hardware, for software licences etc, it was necessary to have a proper corporate entity as owner. On the other hand the service should not be owned by one of the members. CENL being the owner has the advantage of all CENL members being “on board” from the very beginning of the service. But as it was clear that not all CENL members were ready to provide their metadata right from the start, it was decided to divide the membership into “basic” and “full”. All CENL members are basic members of The European Library. They provide information on their libraries and services and appoint a member for the editorial group, maintaining the “libraries” part (as they did before for the Gabriel service). Basic membership is included in the CENL membership fee, nothing has to be paid extra.

Full members are the project partners, expanded by the Bibliothèque nationale de France who joined right after the end of the project phase, and the national libraries of Austria, Croatia and Serbia, who joined in summer 2005. Above submitting information for the “libraries” part, full members provide their catalogues and digital collections for the integrated search facility of the European Library. They pay an annual The European Library membership fee and appoint personnel for different working groups: the commitment for being a full member is not only financial but also in providing human resources.

The day-to-day work of transforming the project results into a real service, of developing and maintaining the website and the portal, of marketing, supporting the full members, and of implementing strategic decisions taken by the management board is done by a small team of five persons, based at the National Library of the Netherlands: The European Library Office.

Every full member is represented by its director in the management board, responsible for strategic decisions and the control of the development of the service. As the management board only meets twice a year and will grow constantly, an executive group of three persons from the management board was elected to come together more often by telephone conference and to be able to make quick decisions if needed.

As The European Library Office only does the work which is better done centrally for the service, it has to be supported by the full members and makes use of working groups, staffed by the full member libraries:

- The contacts group, coordinating all activities for The European Library within their libraries, submitting strategic recommendations to the Management Board, and producing new ideas for the service;
- The technical working group, dealing with technical questions and metadata, taking responsibility for ensuring that The European Library is working to industry standards, for future technological strategy and development, reviewing and approving all major architectural changes and specifications, and making best practice recommendations to the Office and help in their execution;

- The languages working group which carries out all translations necessary for The European Library and takes care of correct spelling and syntax of the “libraries” section in English, French and German;
- And the editorial group, maintaining the information pages in the “libraries” section.

Full members are eligible to appoint personnel for every group while the basic members only staff the editorial group.

Organisational Structure

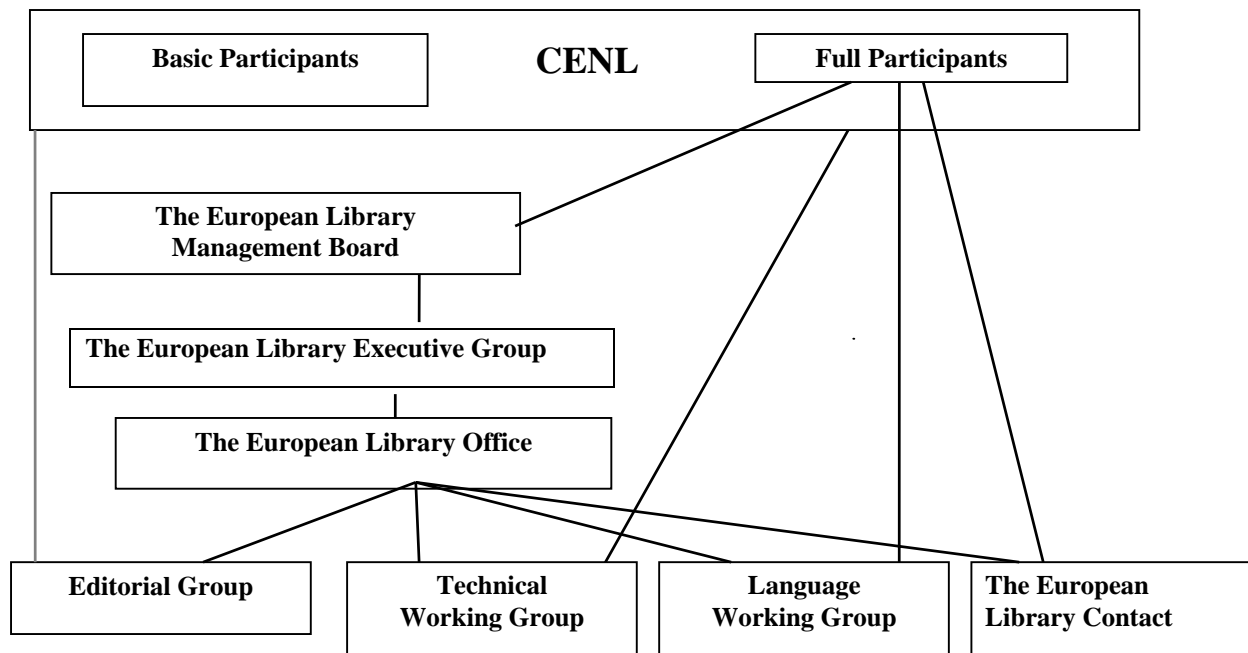


Fig. 5: Organisational Structure of The European Library

Extension of The European Library

The extension of The European Library is planned in two ways: Extension by new partners and extension by functionality, especially with regard to multilinguality. Both aspects are covered by the ongoing EU-funded project TEL-ME-MOR and by a project proposal for the “European Digital Library (EDL)”, submitted to the European Commission in November 2005.

TEL-ME-MOR

TEL-ME-MOR^{vii} is funded in the course of the last enlargement of the European Union: there are special funds for the integration of the ten New Member States^{viii} (NMS). The project started on 1 February 2005 with the overall goal of stimulating and facilitating the participation of organisations from the NMS of the European Union in projects funded within the Information Society Technologies (IST) programme, particularly in the areas of Cultural Heritage and Learning. TEL-ME-MOR also supports nine of the ten national libraries of the NMS to integrate their catalogues and collections into The European Library (Slovenia, though also a NMS, was project partner in TEL and is already integrated). The aim is to

increase access by researchers, scholars, teachers, students and individual citizens across Europe to essential European information resources held by European national libraries and being accessible via The European Library.

In this project multilinguality again is treated as an important issue: Naturally the search interface of The European Library will be offered in all official languages of the NMS, but furthermore the project website is provided not only in English (which is the default) but also in all official languages of the project partners. As the project aims especially at awareness building and dissemination on EU activities and opportunities, it is thought to be of paramount importance that there is no language barrier, neither on the project website nor on the project events. Several information workshops are carried out regionally and all public relations activities are done in the national languages.

By the end of the project in January 2007 The European Library will have gained nine more full partners.

European Digital Library

A further expansion in terms of full membership is envisaged in the project “European Digital Library (EDL)” which was applied for at the European Commission under the aegis of CENL. The project aims at integrating the remaining national libraries of the EU and EFTA^{ix} states as full members into The European Library, namely Belgium, Greece, Iceland, Ireland, Liechtenstein, Luxembourg, Norway, Spain and Sweden.

But the project also pursues the goal to take The European Library a step further towards its vision and mission:

Vision: Provision of equal access to promote world-wide understanding of the richness and diversity of European learning and culture

Mission: The European Library exists to open up the universe of knowledge, information and culture of all Europe’s national libraries

At present The European Library mainly provides bibliographic records of the holdings of the partner libraries. Accessing the material admittedly is difficult as every library has different using conditions and most of the material is “traditional”, which means printed books or journals and not digital resources. Hence the “world-wide understanding of the richness and diversity of European learning and culture” and to “open up the universe of knowledge, information and culture of all Europe’s national libraries” is provided, but very indirectly.

In a user survey The European Library carried out in summer 2005 it became clear, that many users would expect direct access to the resources of the partner libraries – may this be a realistic expectation or not, we have to deal with that.

In September 2005 the European Commission published the communication “i2010 Digital Libraries”^x where it unveils its plans for European digital libraries for the years to come. The core statement of this communication is that the material in Europe’s libraries and archives should be made accessible to all by making it available on the Internet. The key areas for action are digitisation, online accessibility and digital preservation. As this is exactly what The European Library needs to fulfil its mission – more digital and digitised material being accessible online and The European Library acting as a single access point for the material of

European national libraries – CENL submitted a project proposal for the “European Digital Library” under the actual eContent plus programme to prepare for activities to come under the i2010 programme from 2007. One work package will deal with how digitisation activities in the CENL member states can be coordinated on European level to avoid duplication of work, to share best practice and standards, and to offer a portal to European digital cultural heritage. Furthermore, this work package will work on the identification of existing digital collections (newspapers, celebrities etc.) which could be merged virtually to a European collection, e. g. of newspapers, celebrities etc. This could serve as a starting point of co-operation and might serve as a model to create a framework for co-ordination in the future.

And last but not least there will be a work package on multilinguality, aiming at developing the multilingual capacity of The European Library network. The provision of multilingual interfaces is important for pulling users into the service. But the real multilingual service would comprise multi- and cross-language search mechanisms using controlled subject vocabularies, classification systems and name authority files. Furthermore this work package will seek to extend this investigation to full-text and non-controlled data such as titles and abstracts, reflecting the increasing quantities of full text material available in digital format.

We hope the project proposal for the European Digital Library will be approved so that we can start to work on the topics outlined above from summer 2006.

Before coming back to multilingual issues again, let me summarise the growth of The European Library in terms of partners:

- *Founding members:* national libraries of Finland, France, Germany, Italy, Netherlands, Portugal, Slovenia, Switzerland, United Kingdom (9)
- *New Members in 2005:* Austria, Croatia, Serbia (3)
- *New Members via the TEL-ME-MOR project:* Czech Republic, Cyprus, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia (9)
- *New Members in 2006:* Denmark, Russia-Moscow (2)
- *New Members via the European Digital Library project:* Belgium, Greece, Iceland, Ireland, Liechtenstein, Luxembourg, Norway, Spain, Sweden (9)

= 32 of 45 CENL members will be full members by the end of 2007.

Activities on Multilingual Issues

Which activities on multilingual issues are already underway to develop mechanisms which allow multilingual and cross-lingual searching? For multilingual subject search there are three ongoing projects: MACS, Crisscross and MSAC. For multilingual personal name search, there is an ongoing project using mappings of name authority files: VIAF.

MACS (Multilingual Access to Subjects)^{xi}

Most of the national libraries use some kind of subject indexing system, either a classification or a subject heading language. Classifications are language independent and do not cause any problems regarding multilingualism. But many national libraries use subject heading languages for subject indexing.

If a patron wants to do a subject search, he usually does not have any problems within his own country. But if he wants to use foreign libraries, a subject search becomes impossible if he does not master the foreign subject heading language. This problem is even more evident

in countries with more than one official language like Switzerland. Switzerland has four official languages (German, French, Italian and Rhaeto-Romanic) but the literature of the Swiss National Library is only subject indexed in German. Swiss people who are not that fluent in German have thus difficulties in doing a subject search in their own national library.

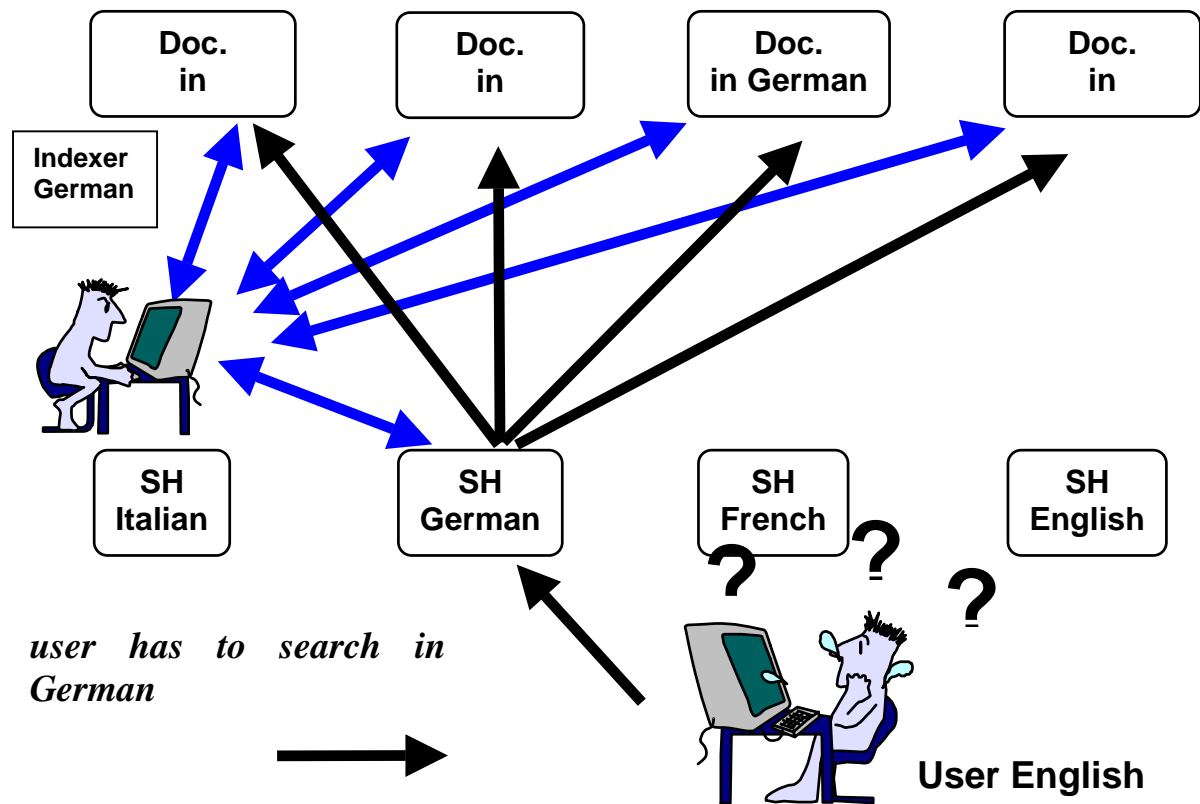


Fig. 6: Monolingual Subject Search

The picture above illustrates the problem. In a German or Swiss library catalogue documents are indexed in German, regardless of the language of the documents themselves. An English speaking user would have to master the German subject heading language to do a subject search where he hopes to find documents in English.

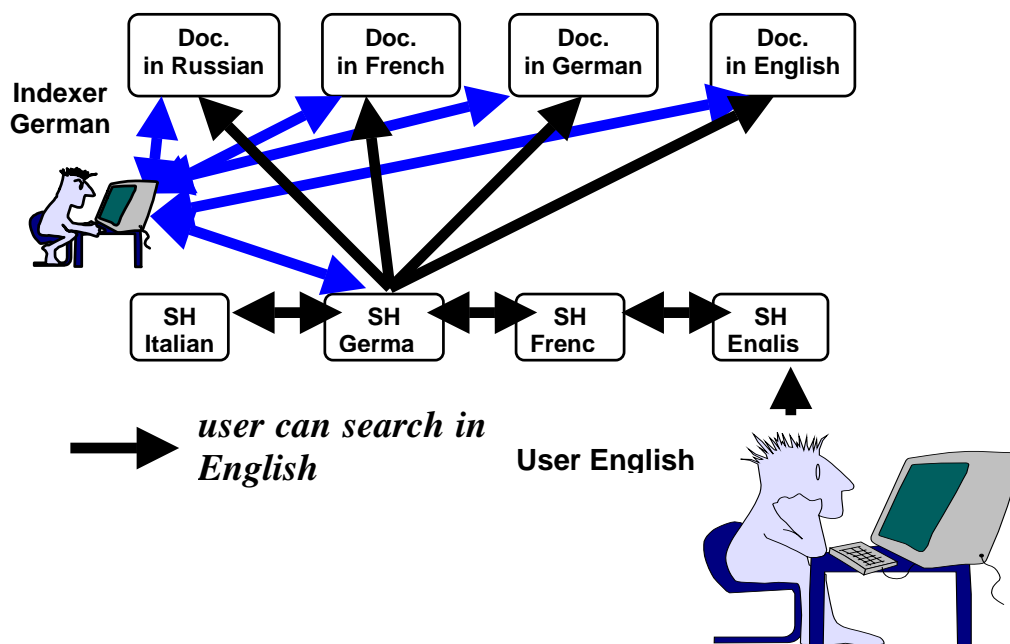


Fig. 7: Multilingual Subject Search

The solution: If the existing subject heading languages would be connected, every user could use his own language to do subject searching, regardless of the language the documents are indexed in. The user would enter English search terms, the system would transfer this search term via links into the appropriate indexing language and send a query with the proper term to the database. It became clear, that a multilingual subject heading system is a necessity but it was also evident that the invention of a new multilingual subject heading language would be too expensive with regard to costs and personnel. Furthermore, all national libraries possess millions of documents which are already indexed, and no one could imagine seriously that these would be re-indexed.

So the solution would have to build on the existing systems and the idea was to create links between the subject heading languages. This was tested in a feasibility study and a pilot project with the national libraries of Switzerland, Germany, the United Kingdom and France for the subject heading languages in English, French and German. The feasibility was proved: searches via the linked subject heading systems gave back relevant results.

There are some basic principles in MACS: The languages and subject heading languages are all equal, there is no pivot language. No new multilingual thesaurus will be created but the system builds on the existing subject heading languages. Between the terms of the subject heading languages equivalences will be established intellectually, not automated. It is NOT a translation of the thesauri! These links are maintained in an extra, external link management database so that the subject heading languages remain autonomous. In the MACS system only the mappings and the numeric identifiers are kept.

The Canadian National Library carried out a similar mapping exercise between the Library of Congress Subject Headings and the French Rameau – about 70.000 subject headings were linked. For transferring the MACS project into a real service these links were taken over,

updated and edited by the Bibliothèque nationale de France and have now to be supplemented by the German subject headings.

Crisscross

The addition of German subject headings to the English-French links will be done in a project called Crisscross which has just been approved by the German Research Foundation and started in December 2005. But in this project more will be done than just finishing what was started in MACS.

The Dewey Decimal Classification (DDC)^{xii} was translated into German and from 2006 certain parts of the holdings of Die Deutsche Bibliothek will be indexed with DDC. The DDC is the most used classification scheme world-wide. Even in countries where the DDC is not actively used, DDC enriched bibliographic data are available via copy cataloguing. To provide the possibility of not only cross-lingual but also cross-system search, the idea was to add DDC numbers to the German subject headings to create a link between the classification and the subject heading language. This will be done in the Crisscross project for all data sets of the German Schlagwortnormdatei (except for personal names and corporate names).

The goal is to provide the user with linked subject index systems so that he can do subject searches without caring about the language or knowing how the different systems work in detail. An integrated subject search in heterogeneously indexed databases will be possible: The user enters a search term in English, this term will be transformed into the equivalent in German and French, is linked via the German subject heading language to the DDC so that the query is forwarded to distributed databases in English, German, French and with the DDC number, and all results indexed with one of the indexing languages will be given back. This is especially convenient in an online environment, searching databases from different institutions world-wide. For Germany this applies for example for the Virtual Subject Libraries which are developed for special subject areas, seamlessly connecting traditional library material with electronic media from different sources, and their common portal *vascoda*^{xiii}.

MSAC (Multilingual Subject Access to Catalogues of National Libraries)

The MSAC project has an additional approach. The goal of the project is to establish a multilingual subject approach to the catalogues of eight participating libraries (Czech Republic, Croatia, Estonia, Latvia, Lithuania, Macedonia, Russia, Slovenia) via the Universal Decimal Classification (UDC) which is widely used in these countries. The English version of the UDC classification scheme is used as core set where all the other language link their equivalents to. When this is done, a user can enter a search query in one of the eight languages or in English and will receive results from all catalogues using the UDC scheme.

As this project is not externally funded but has to be done in addition to the everyday work of the partner libraries, progress is very slow but promising.

A mapping between DDC and UDC exists, so that the outcome of the Crisscross and the MSAC projects could be linked again. Furthermore there are other national libraries showing interest in joining MACS and adding e. g. Italian and Spanish subject heading languages. In the future we will have to discuss further mappings and linkings, and tests would have to be carried out on how scalable the mappings are. But it is for sure that there is potential in linking different systems and thus providing language and system independent search mechanisms to the users.

Name authority files and multilinguality: VIAF (Virtual International Authority File)

Regarding search there are not only problems with different languages and different scripts but also with different transcription schemes for these foreign scripts. This is especially evident for personal names.

I will give you an example on how we struggle in Europe with transcribed names from the Chinese: There is a classic Chinese philosopher called Kong Zi. The name under which this philosopher is known best in Germany is “Konfuzius”. In the German name authority file the main entry for Kong Zi is “Kong, Qiu” with lots of “used for” relations. If the user searches with “Konfuzius” in a German library catalogue, s/he will find all documents of or about Kong Zi, regardless of the variation of the name used in the document itself. But if the user submits a search with this most common version of this name in German “Konfuzius” in the British Library catalogue, s/he does not get any records, though the BL of course owns many documents of Konfuzius. This means that in a distributed search environment like in The European Library the user might not find any records on a certain author though the libraries own documents, just because s/he does not know how the author’s name is written in that particular language. This is a problem which The European Library has to meet to offer refined and successful distributed searches over all catalogues connected via TEL. There are already some projects going on aiming at finding methods of how to make the best use of existing name authority files in the partner libraries, e. g. VIAF.

Within the framework of their [ICABS](#) (IFLA-CDNL Alliance for Bibliographic Standards) responsibilities Die Deutsche Bibliothek and the [Library of Congress](#) together with [OCLC](#) not being an ICABS partner itself intend to prove and to realise stage to stage a concept of the [IFLA Cataloguing Section](#) for international co-operation in the field of authority data. The concept of VIAF includes linking of national authority files virtually through a crosslink file, thus creating an international authority file. Authority records are created and stored locally within the single national authority files. By means of the crosslink file VIAF they can be searched in the same way as in a physically unified international authority file and are available for user search, international data exchange and the co-operative data use.

Matching and linking the German personal name authority file Personennamendatei ([PND](#)) and the Library of Congress Name Authority File ([LCNAF](#)) to a VIAF for persons, the project proves that the concept of VIAF can be realised also with huge amounts of data. At the same time the project develops the functionalities necessary for the VIAF thus delivering the basis for its further development.

The project is subdivided into 4 stages:

1. Matching of the files, identification of identical records, linking of identical name authority records in a crosslink file. To support the identification the name authority records are enriched by attributes originating from the associated title records.
2. Building a VIAF server
3. Establishing an update-workflow between the VIAF and the authority files involved by using the OAI protocol
4. Providing a user interface, implementation of Unicode, activation for user search

During the present phase of stage 1 of the project OCLC matches Library of Congress Personal Name Authority records to the German Personennamendatei. For this purpose the name authority files are enriched by data elements originating from the associated title

records of Library of Congress, Die Deutsche Bibliothek and BibliotheksVerbund Bayern (Bavarian union catalogue), and matched with them. A Trigram-based algorithm is used for the comparison of text strings. As soon as a matching name record is found, a VIAF record is created. The VIAF record contains at a minimum the ID number of the authority file and the main entry personal name from both sources, thus referencing back to the source records. For the purpose of testing in stage 1 a VIAF test database has been established.

Coming back to the example above: Given that a VIAF was created across Europe, if a user searches in The European Library for “Konfuzius”, he would also receive results from the British Library though “Konfuzius” does not exist in their name authority file. But as all name authority files available in European national libraries are linked and searchable, the “used for” entry in the German file would link to the appropriate British authority file and deliver titles from the British catalogue.

Conclusion

Above four fields of The European Library where multilinguality plays an important role were outlined:

- User interface
- Helpdesk
- Subject search
- Personal names

We have a solution for the user interface which was already partly operational in the Gabriel web service and which we hope will be scalable enough to manage all the languages needed eventually. The testing in the MACS project has been completed, to 70.000 English-French links the German equivalents will be added within the next two or three years to build the base for an operational multilingual subject search service. The German subject heading language Schlagwortnormdatei will be enriched with DDC numbers so that a convenient verbal search can be carried out in data indexed with DDC – in German, but as the German terms will be linked to English and French, these languages will also be available for searches for DDC indexed titles. Further mappings are conceivable, for example between DDC and UDC, respectively linked to verbal terms in different languages, like in the MSAC project. In different projects we have done some testing regarding making use of the different name authority files for better search results. These projects are still ongoing but promising that there will be a solution for an operational service. And finally all these developments have to be implemented into The European Library and be extended by more partners, hence more languages.

Multilinguality is a true challenge and a lot of work – but we think that it is of paramount importance for information services in order to truly connect with users.

Endnotes

ⁱ UNESCO's Recommendation on multilingualism and access to cyberspace, <http://unesdoc.unesco.org/images/0013/001329/132965e.pdf>, p 19

ⁱⁱ <http://global-reach.biz/globstats/index.php3>

ⁱⁱⁱ Source: <http://global-reach.biz/globstats/index.php3>

^{iv} Source: <http://global-reach.biz/globstats/index.php3>

^v List of official languages (in German):
http://de.wikipedia.org/wiki/Liste_der_Amtssprachen

^{vi} Homepage of the TEL project: www.europeanlibrary.org

^{vii} Homepage of the TEL-ME-MOR project: www.telmemor.net

^{viii} The ten new member states: Czech Republic, Cyprus, Estonia, Hungary, Latvia, Lithuania, Malta, Poland, Slovakia, Slovenia

^{ix} EFTA: European Free Trade Association, international organisation of Iceland, Liechtenstein, Norway and Switzerland. www.efta.int

^x Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: i2010 Digital Libraries (COM(2005) 465 final);
http://europa.eu.int/information_society/activities/digital_libraries/doc/en_comm_digital_libraries.pdf

^{xi} Homepage of the MACS project: <https://ilmacs.uvt.nl/pub/>

^{xii} www.oclc.org/dewey/

^{xiii} Vascoda: www.vascoda.de