

The European Library (TEL): Access to European national library collections

Maja Žumer

National and University Library, Slovenia

Theo van Veen

Koninklijke Bibliotheek, The Netherlands

Introduction

TEL (The European Library) is a collaboration between a number of European national libraries under the auspices of CENL (Conference of European National Libraries). The main goal is to establish a single access point to their holdings for users in Europe in worldwide.

The vision of TEL stems from the unique characteristics that the national libraries offer. These include:

- the comprehensiveness of the collections held by each of the national libraries, a feature that is greatly enhanced when the national libraries unite and pool their resources;
- the national libraries' holdings of outstanding historical and special collections, offering a richness and diversity that cannot be found anywhere else;
- that one of the most important reasons for an Internet user to use the services of the national libraries is the reliability of information and the high quality;
- the fact that national libraries are the institutions that are the most dedicated in their own countries to the long term preservation of the intellectual output, both traditional and digital, of their own country; when the various national libraries unite in their service offerings, the worth of this feature is greatly enhanced;
- that as many of the resources in the national libraries are unique, the libraries should invest in creating added value to the digitised documents: the national libraries are in many cases the only source of information for a user;
- that in a united Europe the national libraries are in a unique position to show the cultural diversity within each country and in all the participating countries acting in unison;
- that the national libraries are aware of the European identity of their collections alongside their national identity.

TEL is largely being funded by European Commission as an accompanying measure under the cultural heritage applications area of Key Action 3 of the

Information Society Technologies (IST) research programme. It started officially on February 1st, 2001 and will run for 30 months (until July 2003).

The project participants are:

- The British Library (Co-ordinating Partner),
- Die Deutsche Bibliothek (Germany),
- Koninklijke Bibliotheek (Netherlands),
- Helsinki University Library (Finland),
- Swiss National Library,
- Biblioteca Nacional (Portugal),
- Biblioteca Nazionale Centrale Firenze (Italy),
- Istituto Centrale per il Catalogo Unico (ICCU) (Italy),
- Narodna in Univerzitetna Knjižnica (Slovenia),
- Conference of European National Librarians (CENL)

Scope and goals

TEL is a feasibility study, but there is strong commitment to develop an operational service based on the findings and development within the TEL project. The goal is to offer integrated discovery and access to the collections, products and services of European national libraries, aimed at users worldwide, who want a powerful and easy way of finding European cultural resources. With future TEL service participating libraries will widen the access to their born-digital or digitised resources, but also to non-digital resources. The first emphasis will be the access to the legal/voluntary deposit (subject to rights and permissions) and digitised treasures. Secondary preferences include licensed published materials and themed collections (possibly as a result of co-operative collection building). The collections offered will be of any subject, but the primary focus will be on:

- humanities
- law
- social sciences
- economics
- business
- natural sciences

National libraries plan to target the user groups they presently serve:

- the 'informed citizen'
- academics/researchers
- higher education students

but also

- all users of libraries, archives and museums
- booktrade

Generally TEL partners wish to primarily address the following:

- European nationals in order to access/learn about other European countries
-

- European nationals in order to develop a sense of European identity and culture
- Non-Europeans to learn about Europe

The main results are expected to be the development and testing of open standards, working methods and practices that can readily be adopted by all national libraries to work as a seamless partnership.

The goals of TEL can be summarised in the following areas (which are also reflected in the workpackages):

Publisher relations (Workpackage 1)

- Common approach on negotiating legal and voluntary deposit agreements with publishers.
- Establishing co-operative approaches to business, licensing and copyright matters.

Business plans and models (Workpackage 2)

- Development of business plans and models.
- Market research.

Metadata development (Workpackage 3)

- Development of common standards for metadata to support wide scale access to digital, off-line digital and non-digital materials.
- Collection-level metadata
- Concertation work on multilingual access to support a model for a multilingual service.

Interoperability testbeds (Workpackage 4)

Search in distributed databases using Z39.50 and http/XML, including

- testing of interoperability (between Z39.50 targets and http/XML server),
- search functionality,
- presentation of results,
- portal functionality,
- other services (authentication, connection to local services...)
- evaluation of performance.

All information about TEL is available at the TEL website
www.europeanlibrary.org.

Highlights

There are several interesting new developments within TEL, particularly:

- Hybrid combination of central indexing and distributed searching
- Introduction of new search and retrieve protocol (SRU)
- Integration of multilingual and name authority services
- Implementation of OpenURL
- Metadata registry
- Collection level descriptions

Overview of infrastructure

An overview of the infrastructure is shown below. The TEL operational service will be a central portal but during the project there will separate portals for Z39.50 and SRU access. In the overview below the ellipse around both portals illustrates this.

Some TEL participants will provide metadata in XML via the OAI-protocol. These records will be indexed in a central index and the portal will search and retrieve them via the SRU-protocol. Other participants will make their metadata accessible via Z39.50 or SRU. The metadata will offer links to the digital objects and services, either directly or indirectly via link services using OpenURL's or URN-resolvers.

Additional services might be offered, for example multi-linguality (translation of specific subject headings) or name authority services allowing the use of search results from a name authority database as input for subsequent searches in TEL.

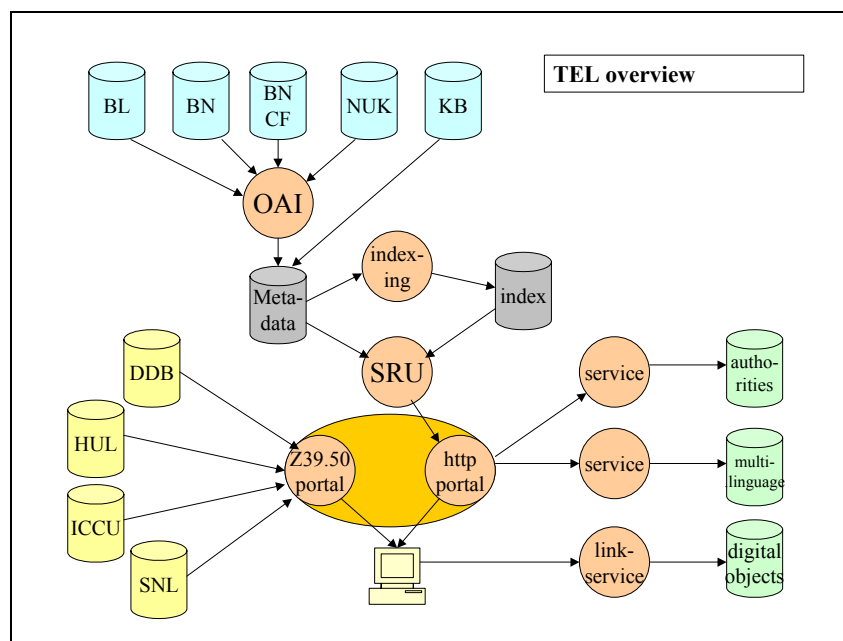


Figure 1. Overview of TEL infrastructure

Developing the TEL datamodel

It was considered very desirable to adopt existing metadata standards where possible. The first choice was for XML and Dublin Core as a starting point. To allow for extra services and functionality it was however acknowledged that more terms would be needed. The DC-Library Application Profile (DC-Lib) was identified as the most obvious candidate to form the basis of the TEL datamodel. The concept of *application profiles* has emerged within the Dublin Core Metadata Initiative as a way to declare which elements from which namespaces are used in a particular application or project. Application profiles are defined as schemas which consist of data elements drawn from one or more namespaces, combined together by implementers, and optimised for a particular local application.

The general approach has been to define the functionality needed to underpin the services and publication types as currently envisaged on the one hand and identify the metadata required to enable that functionality on the other. Special attention will be paid to digital collections and collection level descriptions (CLD). The functional requirements could then be analysed against DC-Lib to see what gaps existed in the element set. From these results we can determine whether DC-Lib is sufficient for TEL or whether it will be necessary to define a TEL Application Profile which will incorporate DC-Lib with additional elements.

To develop the TEL datamodel the following concept has been adopted. There are a number of functions and services that TEL wants to offer and in order to do so specific metadata are needed. On the other hand, once a metadata record is retrieved, the TEL portal will offer some functionality based on the occurrence of specific terms in the metadata. This concept can be visualised in a functionality matrix with functions at one axis and metadata at the other axis.

The main functions that we distinguished are grouped as follows:

- **Search and resource discovery**
All searchable fields contribute to this function
- **Record retrieval**
Record retrieval for harvesting and remote indexing requires a record identification.
- **Identification of resources**
Needed to find and access resources. Mainly elements for identification like ISBN or URI play a role.
- **Description**
Helps the user in a decision to obtain the object. Examples are description, abstract and thumbnails.
- **Linking services**
Linking services help locating objects or services for the object. Mainly identifier elements but also additional terms play a role.
- **Multilinguality service**
A service translating user input or received metadata into different languages to create new queries. Most textual metadata may play a role.
- **Authority service**
A service to find main entries for names and subjects. Creator and textual subject or classification data play a role.
- **Collection level services**
The functionality that helps to find and identify collections, link to those collections or broadcast queries to selected distributed collections.
- **Authorisation**
Access may depend on the service, type of publication, year of publication, publisher, the user etc. Restrictions are indicated by terms and conditions or access rights.
- **Administration**
Functions that keep track of usage, based on, for example, subject or publisher.
- **Hard and software requirements**
Specific metadata to inform users of the requirements on their workstations or detect whether the user's workstation is capable of accessing a publication type. Especially when preservation activities play a role.
- **Navigation**
This functionality concerns navigation between journal and articles, between records from the same author or with the same subject etc.
- **Copy cataloguing**
Metadata may be re-used by other libraries for cataloguing.
- **Miscellaneous**
Off-line ordering, ILL and other services. Mostly accessed directly via URLs.

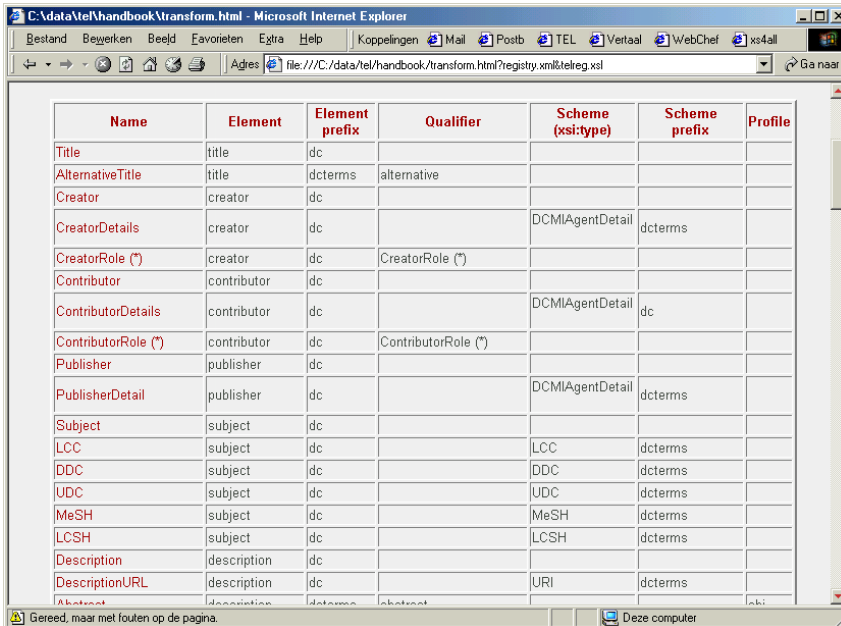
This overview is considered to be rather complete but not all functions are equally important in TEL.

Now rather than showing the functionality matrix, a few metadata elements are picked out as examples to explain the mapping process. First the term identifier. When this element is available we can use it to link to the actual object. For a journal it might be used to offer access to an external service for accessing

the full text of its articles and for a book it may be a link to a local bookshop where the book may be ordered from. But for all of these functions the term identifier doesn't say whether it is an ISSN or URL etc. We need to know more to offer the right functions related to the identifier. Another example is the element "description". If description is an abstract a translation service can be offered, if it is a thumbnail it may be used as inline image. But again we cannot tell that without refining the element "description".. So we need to introduce refined terms by means of qualifiers and encoding schemes according to the DCMI guidelines.

Although DC-Lib was identified as a valuable starting point for the TEL application profile it does not contain all the elements that TEL will need. Even if it sufficed for now it would not be sufficient in the future when new functionality is introduced. We therefore need to create a TEL application profile which may contain elements that are not part of a DCMI namespace at the moment.

This process of mapping resulted in a functionality matrix on bases of which we could define a first version of the TEL Application Profile, which can be used by TEL partners and dataproviders to search for elements and check their characteristics in order to help them in converting their metadata for TEL. There are different views possible and an example of a global overview is shown below.



Name	Element	Element prefix	Qualifier	Scheme (xsi:type)	Scheme prefix	Profile
Title	title	dc				
AlternativeTitle	title	dcterms	alternative			
Creator	creator	dc				
CreatorDetails	creator	dc		DCMIAgentDetail	dcterms	
CreatorRole (*)	creator	dc	CreatorRole (*)			
Contributor	contributor	dc				
ContributorDetails	contributor	dc		DCMIAgentDetail	dc	
ContributorRole (*)	contributor	dc	ContributorRole (*)			
Publisher	publisher	dc				
PublisherDetail	publisher	dc		DCMIAgentDetail	dcterms	
Subject	subject	dc				
LCC	subject	dc		LCC	dcterms	
DDC	subject	dc		DDC	dcterms	
UDC	subject	dc		UDC	dcterms	
MeSH	subject	dc		MeSH	dcterms	
LCSH	subject	dc		LCSH	dcterms	
Description	description	dc				
DescriptionURL	description	dc		URI	dcterms	
Abstract	description	dcterms	abstract			

TEL metadata registry and metadata formats.

It is not envisaged that TEL will stop at the end of the project but will continue to evolve afterwards. The impact of this is that the development effort will not be solely focussed on the TEL test-bed, but a more generic approach will be followed. This will allow future TEL functionality to be taken into account. It also raises the possibility of another approach to metadata specification. This option is to create a TEL registry of metadata that would allow the addition of new metadata elements that are in use by the partners.

As seen now, The European Library project is a system for access to all types of collections and materials owned by the European national Libraries. In the future it may be opened to other types of memory institutions and, if so, the issue of semantic interoperability will become an important aspect of the development. The flexible structure of Dublin Core and the different sectoral interpretations of how to describe a digital object could be an obstacle to interoperability. In Dublin Core there are no rules of the kind we employ in libraries for how the values in metadata are constructed and a unitary search does therefore not guarantee the localisation of all types of digital objects.

The creation of a TEL profile based on DC-Lib it is also important to define a model and an ontology as a starting point for the development of vocabularies relating to different applications in the Cultural Heritage sector. Libraries own and catalogue materials that are also owned and catalogued in different types of institution (Archives and Museums for example). It is important to define the correspondence between terms, functions and concepts in the systems describing that material.

There are may be several ways of dealing with this:

1. Promote a common standard schema, independent of what different data providers are using internally. This would entail the definition of one comprehensive mapping table using which all information providers could convert their metadata to a single TEL-schema.
2. Introduce a TEL metadata registry that contains metadata from existing metadata standards, but which can be extended with local metadata. The Library application profile will be the "main" entry, but as soon as new metadata are introduced for which there is no existing element in DC-Lib application profile, or other profiles accepted by TEL, then the introduction of new elements would be allowed. In this context, the TEL Registry is seen as a system that facilitates the procedures involved in allowing the TEL application profile to evolve in line with increased functionality and extension to different types of institution. In this it is slightly different from the concept of a registry as currently used by DCMI.
3. Use the TEL indexes mainly to find all resources sharing the *same index entry* and - for the other objectives - the user will be redirected towards the real catalogs. In this context, Dublin Core presents itself as a metadata pidgin for digital tourists who must find their way in this linguistically diverse landscape.

The first option is preferred but we may need the second option to realise the first one: the registry will define a common standard schema, but building the schema in a decentralised and incremental way is enabled. Data providers would be allowed to add new metadata elements but at the same time all providers can monitor the developing schema and raise objections to inappropriate metadata elements. The third option is a last resort option for very specific cases to provide metadata elements from the original record for which there are no corresponding metadata elements in the registry.

The TEL portal would use the metadata elements from this schema/registry for such actions as display, translate, generate a link or generate a new search. This

TEL registry would therefore contain information on how TEL will handle these metadata.

The registry will contain at least:

- element name
- name space
- originating metadata standard
- labels for presentation in different languages
- flag to indicate that it should be presented in the full presentation
- flag to indicate whether the element should be used as clickable link for searching
- element name that it maps to (in case it comes unconverted from another metadata standard)

This list may grow in the future as the usage of different metadata elements in practical situations is extended. When new collections from national libraries enter the TEL-system the flexibility of such a registry will facilitate compliance to the TEL system. In order to submit proposals for new metadata elements a submission form will be available for that purpose to the TEL participants and dataproviders.

Collections

TEL is a collection of collections and as such special attention is paid to the handling collections. A collection is considered to be an aggregation of objects. These objects are described in databases or websites. A very important issue in this respect is the fact that targets for searching the descriptions of these objects do not map one to one to collections. A target may be for example a library catalogue, but the objects of a collection might be described by a subset of such a target. In TEL we use collection descriptions as a means to find the targets for searching. In this respect it is interesting that a collection description is considered to be a normal metadata record and is treated as such. When the user finds a collection description the portal will use this description to offer one of the corresponding potential services, such as linking to the collection website or make this target one of the portal's search targets.

Data providers can notify TEL about new collections via a collection description form. After an acceptance procedure the collection may become available as part of the primary collection database and become target for searching or harvesting.

SRU test interface

A test interface has been developed for testing the SRU access to the central index and database and to test the metadata that were harvested via OAI. This testinterface showed to be a very valuable piece of software. It has the functionality of a portal giving simultaneous distributed access to SRU resources but it runs in the browser so there is no central portal involved. It can be used as a tool to demonstrate and test the complete TEL functionality. Because it lacks all layout features it is a good basis for a personalised interface. As it is only a webpage with XSL and Javascript it is accessible to everyone and the interface can be adapted by everyone who knows a little bit of XSL and Javascript.

A very important aspect of this interface is that terms that are not in the application profile are recognised as such. This allows signaling of new terms introduced by dataproviders and allows the user interface to be adapted to an evolving application profile and new functionality, but it also allows extensions with terms that are only of local importance.

During the presentation it will be demonstrated how different functions and services are based on the occurrence of specific metadata. Also the navigation within and between collections will be demonstrated.