

Interoperability in The European Library: ‘the devil is in the details’

Sjoerd Siebinga
The European Library Office
the Netherlands

1. Background: What is the European Library

The European Library (<http://www.TheEuropeanLibrary.org>) is a hybrid portal, which provides a unified access-point of the combined resources of Europe’s national Libraries. From its inception, it has aimed to provide a pan-European cross-collection search which would otherwise have been impossible. To achieve this, overlapping techniques/protocols are used to connect from the portal to the different libraries and return results from the users queries to the browser. The main techniques/protocols used are OAI-PMH, SRU, Z39.50 and AJAX (Asynchronous Javascript And XML). Currently, 246 collections are available through the portal from 22 different libraries.

To make these different libraries interoperable the TEL Application Profiles (henceforth TEL-AP) were designed, one for applications and one for collection descriptions. The TEL application profile for applications can be seen as the largest possible set of common metadatafields amongst the libraries, that trigger a function in the portal. Each collection made available via the European Library has its own TEL Application Profile for collection descriptions, to allow the user to select their own subsets. The TEL-AP approach is perfectly suitable for a basic disclosure library resources, but when further integration is desired a more refined method of querying the library sources is required. The European Library office is currently reviewing the option of developing a server-side query-building-engine which composes target-specific queries and clusters the target responses. This new strategy moves away from the current architecture in which the integration of the targets is done in the browser. However, the transition is not as great a departure from the original philosophy as it might appear at first glance. The SRU/Z39.50 gateway, for example, already mitigates much of the library-side diversity.

Even though, we are in the early stages of research, it might still be interesting to go into some of the nitty gritty problems we need to tackle in order to create a more fine-grained level of interoperability between European National Libraries. In the following sections, I will present some of the main focus areas that need to be processed before the query can be sent to the target.

2. Interoperability through adaptive query-construction

One of the most time-consuming consequences of this approach is that the full breadth of the detailed differences between the collections need to be charted and added to the collection descriptions. We have chosen to use a collection view rather than a library, because we have found that even collections within a single library can have wide ranging differences. Even if they appear identical in the specs, hidden features such as normalisation rules or different use of the datefield can still give unexpected results. In the following list, a number of these focus areas are listed.

- normalisation rules
- character encoding
 - * if unicode-variant, are composed or decomposed diacritics used?
- main language, auxiliary language
 - * is the spelling system used consistent and brought up to standards?
- access-protocol
- recordscheme
 - * has there been any non-standard use of metadatafields, e.g. ‘*fourteenth century*’ in date-field

- What controlled vocabulary has been used and on which fields
- Which metadatafields can be seen as free text fields
- Are the identifiers used persistent, or in any case suitable for deep-linking.
- etc.

In the query building component as we currently envision it, we aim address the following issues in the near future:

- Target adaptive normalisation
- better handling of composed vs. decomposed unicode glyphs
- better handling of target-specific non-unicode character encoding requirements
- Multi-lingual subject access
- Mapping of various cataloguing schemes to provide unified access for the user on record level
- Multi-lingual search in free text fields (from any to all)

3. Conclusion

Although much can be accomplished via computational interoperability, full scale accurate coverage is an extremely difficult ideal to work towards. Making library catalogues remains a manual exercise with all its inherent inconsistencies. And even when one takes these inconsistencies aside, the sheer number of variables that need to be taken into account increases exponentially with every collection that is added. Therefore, the interoperability within the European Library, especially controlled vocabulary, will always remain a work in progress. That being said, the user benefit is unmistakable. In-depth access to large sets of collections via the European Library portal, will provide the user with new vistas and cross-sections for research to Europe's cultural heritage. Through continued efforts of collaboration from various fields of expertise, it is my opinion that, this ideal can be approached.