

## Report of meeting with Dieter Van Uytvanck – CLARIN-ERIC and Annelies van Nispen – The European Library 29 November 2012

CLARIN works for the humanities and social sciences. CLARIN is a networked/distributed research infrastructure. Each country has its own CLARIN-centre(s) with digital repositories.

In the Netherlands there are many collaborations with cultural heritage and science institutions

An important task is to describe all data via metadata and make them interoperable. The data is very diverse, but always has a language component. A lot of Language and linguistics data is held in the repositories.

CLARIN has done research on metadata standards. Dublin Core and OLAC were used, but the general feeling was that these metadata schemas were not specific and rich enough.

Metadata components are reusable building blocks to make data interoperable. Researchers can use these building blocks to create their metadata schema and create interoperability. The CMDI standard (<http://www.clarin.eu/cmdi>) is used within all Clarin-centres.

The Data Category Registry (<http://www.isocat.org>) is the concept database. It might be good for the European Library to investigate if these concepts could be used. The language codes are linking to the ISO CDB codes<sup>1</sup>, but otherwise not much is linked to (online) authority files/controlled vocabularies. This is actually a "hot topic". The library world has much invested in authority files and controlled vocabularies. This is a possible area for collaboration between CLARIN and the European Library. One of the most wanted is a list of Academic Organizations, CLARIN has been informed the Deutsche National Bibliothek has such a list for Germany and is looking into this.

The [Open Skos](#) vocabularies approach developed by former colleague Hennie Brugman is probably being adopted and scaled to a European scale  
A meeting with the metadata and tech team of TEL and CLARIN early next year to share knowledge on this topic.

The Virtual Language Observatory (<http://www.clarin.eu/vlo/>) is a very important component in CLARIN for access to the repositories and metadata curation. Currently it is planned to add a search/visualization component based on geolocation and time.

Dieter was at a presentation where the Europeana/the European Library search facilities via time and place/geolocation (via [E4D](#) / [GeoTemCo](#)) were demonstrated. CLARIN is very interested in these features and would like to share knowledge with the technical team of TEL/Europeana.

We could also explore the possibilities about exchange of metadata. The API of the European Library could be used for this. But how this is done and what collections

---

<sup>1</sup> As the ISO CDB entries are not resolvable we are currently investigating to link to OpenSkos entries instead.

are interesting for mutual exchange could also be on the agenda for a meeting early next year to further discuss possibilities for cooperation. We also discussed harvesting via OAI-pmh or even .zip A start could be to integrate the API's, e.g. under the search box: "you can also search the catalogue of all European National Libraries ...."

CLARIN is working with the concept of [Virtual Collections](#).

This enables the researchers to link research data, objects, etc. These virtual collections can be linked to services that CLARIN has developed and will develop in the future, eg. automatic analysis, lemmatization, automatic translations or transcription. But you can also think about a classic book that can connect to a repository where the video's with eg. interviews are stored.

CLARIN has been in touch with DANS to discuss enhanced publications and would be interested in a (European) working group on enhanced publications to establish standards and agreements on what a "standard" enhanced publication should look like.

CLARIN is working on enriching metadata.

Also citation mechanisms and cross-referencing could be on the agenda.

All repositories are set up with [Persistent Identifiers](#) and use [CMDI](#).

CLARIN is currently working on a Full content search/

[Federated content search](#). The European Library has search functionality for Fulltext.

CLARIN & TEL would like to exchange knowledge and investigate the possibility to use/implement each others search tools. CLARIN support SRU. CLARIN is currently working on a prototype, this will be presented at a conference in June 2013. The European Library is invited for this conference.

Annelies proposed the possibility that the tech team could test the prototype and give some feedback. This would allow knowledge exchange.

CLARIN is also working on [Federated identity management](#). CLARIN develops personalized services and access connected to an identity identification that will allow researchers to get access to all European repositories and research environments. Users should be able with a single-sign-on to have access to research data and services.

This is also of possible interest for the European Library

We agree there are many possible areas for collaboration and knowledge exchange. We propose to have a meeting jan/feb 2013 to further discuss this.