

INTEROPERABILITY.

A key concept for large scale, persistent digital libraries.

Interoperability is an essential feature for federated information architectures to work in heterogeneous settings and over time! However, use and understanding of the concept still are very heterogeneous: interoperability is conceived in an object-related or in a functional perspective, from a user's or an institutional perspective, in terms of multilingualism or of technical means and protocols. Moreover, interoperability is conceived on different abstraction levels: from the bitstream layer up to semantic interoperability. The briefing summarises some of the relevant vectors of thought, indicates related conceptual frameworks and places the issue in the strategic context of Europeana.

Interoperability and the European Digital Library Agenda

From the conception of the European Digital Library as part of the i2010 agenda the issue of interoperability has been dominant as in the following statement made by Commissioner Reding in 2005: “I am not suggesting that the Commission creates a single library. I envisage a network of many digital libraries – in different institutions, across Europe.”

As a consequence, interoperability was prominent on what then became the Europeana agenda in that

- Europeana is federating objects from distributed sources,
- Europeana is federating objects from heterogeneous sources with different community background,
- Europeana is part of a bigger framework of interacting global information networks including e. g. other 'Digital libraries', scientific repositories and commercial providers,
- Europeana relies as much as possible on standards and existing building blocks as well as be based on web standards,
- and therefore ‘interoperability’ is the technical heart of the European Digital Library vision!

Aware of the central nature of the concept the EC had set up a Working Group on Digital Library Interoperability active from January to June 2007 and which conceived a number of strategic recommendations based on the concept matrix outlined hereafter.

A Concept Matrix for Digital Library Interoperability

The ISO/IEC 2382 Information Technology Vocabulary defines interoperability as “the capability to communicate, execute programs, or transfer data among various functional units in a manner that requires the user to have little or no knowledge of the unique characteristics of those units.”

In a more systematic view, six aspects of interoperability can be distinguished:

1. **Interoperating entities** can be assumed to be the traditional cultural heritage institutions (libraries, museums, archives) offering digital services, or again the digital repositories (institutional or not), eScience and/or eLearning platforms or simply web services.
2. **Objects of interaction**, the entities that actually need to be processed in interoperability scenarios. Choices range from the full content of digital information objects (analogue/digitised or born digital) to mere representations of such objects – and these in turn are often conceived as librarian metadata attribute sets, but sometimes also are conceived as ‘surrogates’.

3. **Functional perspective of interoperation.** This may simply be the exchange and/or propagation of digital content. Other functional goals are aggregating digital objects into a common content layer. Another approach is to enable users and/or software application to interact with multiple Digital Libraries via unified interfaces (dynamic portals) or to facilitate operations across federated autonomous Digital Libraries. Others again seek to establish a common service architecture and/or common service definitions.

4. Linguistic interoperability (**multilingualism**) can be thought of in two different ways: as multilingual user interfaces to Digital Libraries (relatively well known) or as dynamic multilingual techniques for exploring the Digital Library object space. Three types of approaches can be distinguished in the second respect: dynamic query translation for addressing Digital Libraries in different languages, dynamic translation of metadata responding to queries in different languages or dynamic localisation of digital content.

5. **Design and user perspectives.** Interoperability concepts of a Digital Library manager differ substantially from those of a content consuming end user. A technical administrator will have a view very different from an end user providing content as an author. Different conceptions, again, will emerge from the perspectives of a digital content aggregator, a 'meta user' or a policy maker.

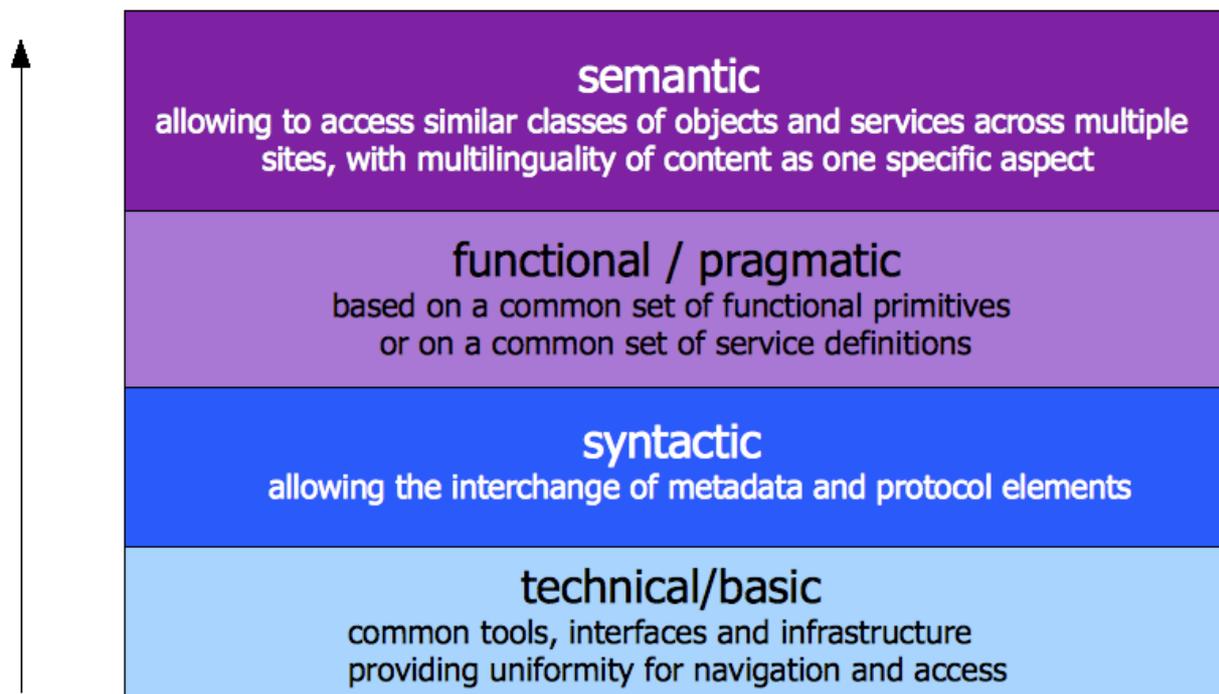
6. **Technological standards** enabling different kinds of interoperability constitute a major dimension with more traditional approaches geared towards librarian metadata interoperability such as Z39.50 / SRU+SRW or the harvesting methods based on OAI-PMH or again web service based approaches (SOAP/UDDI) and the Java based API defined in JCR (JSR 170/283) as well as GRID based platforms such as iRods.

Moreover, interoperability can be considered on different abstraction levels, and the distinctions to be made in this respect cut across all the other matrix dimensions. Within a continuum ranging from a very concrete to a very abstract perspective it is possible to distinguish the four layers of technical, syntactic, functional and semantic interoperability.

Interoperability Abstraction Levels

Interoperability can be considered on very different abstraction levels, and the distinctions to be made in this respect cut across all the other matrix dimensions. Within a continuum ranging from a very concrete to a very abstract perspective it is possible to distinguish four layers as in the illustration below:

Abstract



Concrete

References

- [1] EC Working Group on Digital Library Interoperability.

<http://bnd.bn.pt/seminario-conhecer-preservar/doc/Stefan%20Gradmann.pdf>

- [2] DELOS Reference Model.

http://www.delos.info/files/pdf/ReferenceModel/DELOS_DLReferenceModel_0.98.pdf

- [3] 5S: Streams, structures, spaces, scenarios, societies (5s): A formal model for digital libraries.

http://portal.acm.org/ft_gateway.cfm?id=984325&type=pdf&coll=GUIDE&dl=GUIDE&CFID=80483819&FTOKEN=14249910

- [4] OAI-ORE (data model):

<http://www.openarchives.org/ore/0.9/datamodel.html>

- [5] JISC Information Model:

<http://www.ariadne.ac.uk/issue31/information-environments/>

- [6] JCR:

<http://jcp.org/en/jsr/detail?id=283>

- [7] iRods:

https://www.irods.org/index.php/IRODS:Data_Grids%2C_Digital_Libraries%2C_Persistent_Archives%2C_and_Real-time_Data_Systems

- [8] Europeana:

<http://www.europeana.eu>

While all levels are relevant for the interoperability issue it should be clear that technical and syntactic issues of interoperability probably are far better understood than the upper two layers and that therefore current discussions and research activities concentrate on the intersection of the pragmatic and the semantic layers.

Selected Approaches to Digital Library Interoperability

The following approaches illustrate characteristics of the interoperability concept matrix.

The **Delos Reference Model for Digital Library Management Systems (DELOS)** is a unified model for architecture, functionality and technological components of digital libraries. The interoperating entities are 'Digital Libraries' in a sense that only accidentally corresponds to libraries in the sense of cultural heritage institutions. The primary objects of interoperation are digital 'information objects' defined as a resource identified by an information object identifier. The functional perspective is very generic.

The **5S-model** uses a set of fundamental abstractions – i. e. Streams, Structures, Spaces, Scenarios and Societies – as a basis for a unified formal theory of Digital Libraries (DLs). The interoperating entities of the 5S model are any structured collections of information objects. The primary objects of interoperation are structured bytestreams.

Object Reuse and Exchange (OAI-ORE) is a model for repository federation based on a data model for complex information entities as resource aggregations described by resource maps. The interoperating entities are content repositories. The objects of interoperation are scientific textual documents. The functional perspective is restricted to a few core functions for object retrieval, manipulation and personalisation. The design and use perspective is that of the scientific community maintaining and using scholarly document repositories.

The **JISC information environment** provides integrated access to the networked research and e-learning platforms in the UK. The interoperating entities are content repositories and platforms for e-learning. The functional perspective is to federate these by the use of an all-encompassing service model. The design and use perspective is that of the scientific community maintaining and using scholarly document repositories.

The **Java Content Repository API (JCR)** provides a universal API based on fine grained functional primitives. The interoperating entities are document/content management systems. The functional perspective is to enable direct communication of software instances based on formalised functional primitives that can be invoked using related function calls. The design and use perspective is that of software developers.

Long Term Preservation

In the perspective of this briefing, long term preservation is an interoperability issue along the time vector. The conceptual matrix presented above can be applied to identify the changes occurring over time in terms of entities, objects etc. and what needs to be done to compensate losses of interoperability due to such changes.