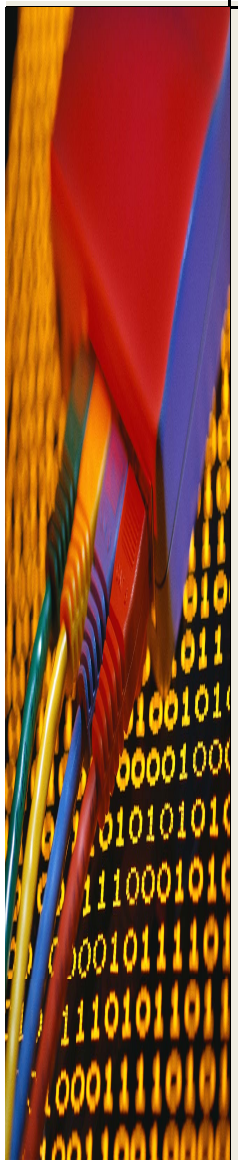
DC-NET WP3



New services workshop

ELABORATE/PROCESS: tools for scientific analysis. The needs for performing analysis on content

Christophe DESSAUX - Ministère de la Culture et de la Communication

Priorities identification

1.Networking

2.ICT Research

3.Realisation of innovative services and pilots (prototypes inserted into real working environments)

4.Network of excellence, competence centres, laboratories and innovation areas.

Source : Italian working group (WP1) and EARNEST European study

Priorities identification

ICT Research

ICT Research area aims to develop technologies that are of specific interest for the cultural heritage sector, such as storage/preservation, advanced search engines, multilingualism, techniques for the access to and the visualisation of scientific data, etc.

Priorities identified by WP1 :

– Digitisation and digital documentation of tangible cultural heritage
– Quality of digital objects, services and uses
– Digital repositories and digital preservation
– Digital Library infrastructures
– Virtual research environments and knowledge infrastructure ecosystems

Priorities identification

Realisation of innovative services and pilots (1/2)

The EARNEST study shows that e-Infrastructure is accepted in general as a major facilitator for research and teaching. It appears that
- the expectations of network users seems not longer to focus only on the basic needs like providing connectivity but also on provision of more complex network services.
- as the networking has passed from the technology innovation to implementation, in which there is a need to enhance services for better support of the researchers needs.

Five major areas are identified:
–bandwidth and quality of service;
–high-quality collaboration services;
–authentication and authorisation infrastructures;
–support and training for using new Internet technologies as they arrive;
–raising awareness in the AHSS community of how new services developed by the NRENs could transform their work, and also raising awareness in the NREN community of the problems that are being approached within AHSS.

Priorities identification

Realisation of innovative services and pilots (2/2)

The main research priorities identified by WP1 are:

– Studies on the requirements of the new services

– Evolution of the MICHAEL service to provide a unique access point to the digital collections of museums, libraries and archives. Integration with national portals

– Authentication and authorization infrastructure

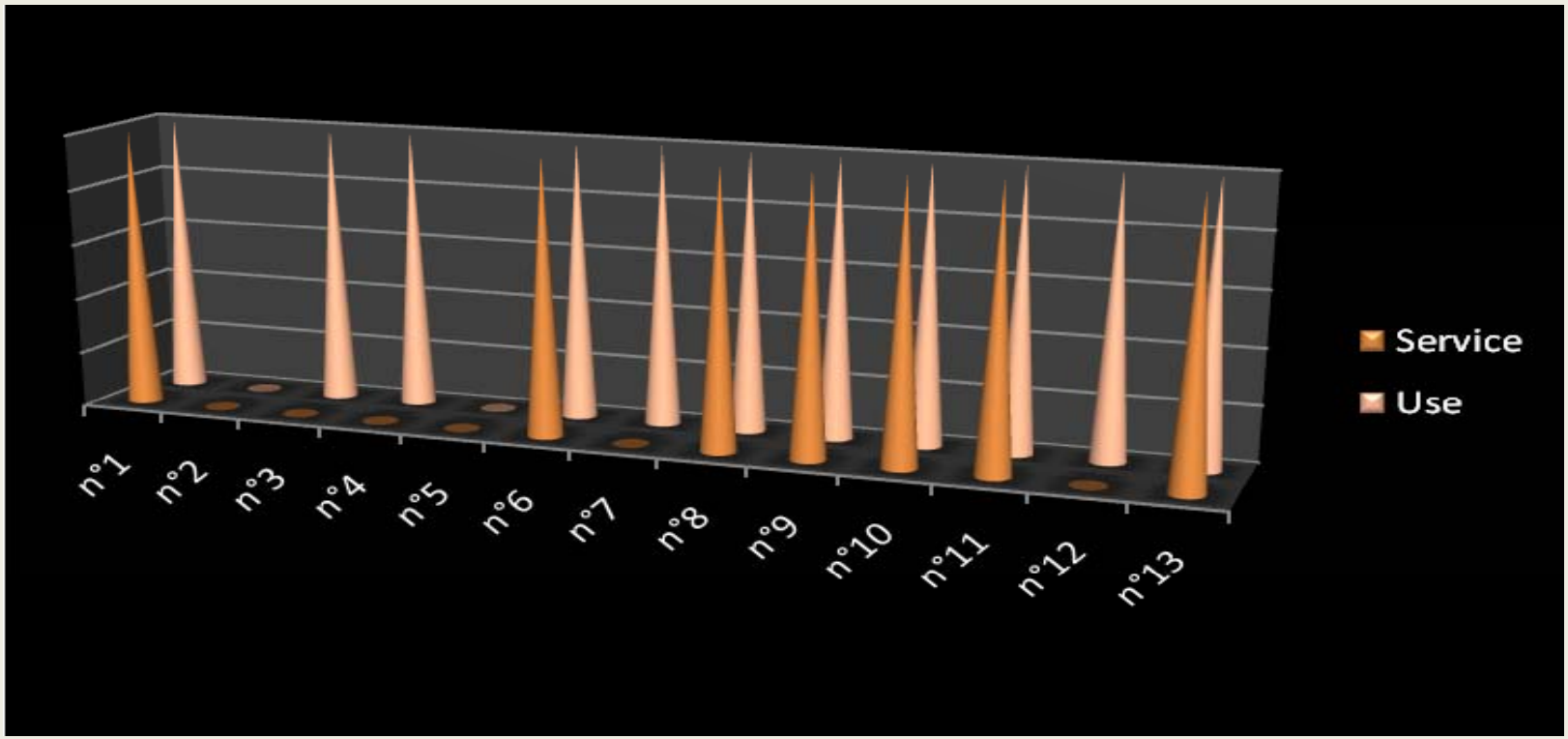
– Implementation of pilots in the sector of the digital audiovisual heritage.

– Realisation of pilots that integrate scientific data coming from laboratories for the analysis of tangible heritage (e.g. chemical analysis, physics characterisation, geological investigations, data from the earth observation, etc.) with innovative services for the visualization of data.

– Virtual reality (3D, immersive technologies, etc.)

– eLearning, eCollaboration

Data Bases creation



Service coverage concerning database creation

Questions

- **Why such a lack of services?**
- **Are some services too specific to be offered by generic e-infrastructures?**

Best practices

- **Identify generic needs in this amount of specific analysis tools.**
- **A broader survey and a more detailed analysis could be necessary.**

A proposal by a researcher

We would be interested in having an **aggregation** and **semantic research** platform where it would be possible to upload, transform and do complex searches on metadata of museum objects, museum collections and related documents and information coming from different (international) parties and in different formats.

The aggregation platform should provide tools to map different data formats to a common format so to allow comparative research between metadata.

A proposal by a researcher

The **aggregation management platform** should provide:

- A metadata mapping tool to map one XML schema to a common standards or to another format. The solution could be based on a XML tool like Stylus Studio, XMLSpy, OxygenXML or other software to create XSLT using graphical mapping as required. It should also be possible to transfer the data after it has been harmonized (data aggregation in and out)

- Most probably data could be delivered by the partners in CSV format to start as their collection management software only allows export in that format. Offer a software tool to convert CSV to a tabular XML. This could be combined with the XML tools mentioned here for.

- It should also be possible to upload thesauri form different sources and formats. It would be useful to have the possibility to upload a simple CVS and be able to transform this into xml SKOS. A thesaurus mapping tool should allow interconnecting all these different sources together. Furthermore the platform described above should also allow semantic enrichment processes by giving the possibility to connect (multilingual) data to multilingual thesauri and ontology's information.

A proposal by a researcher

Semantic enrichment is adding specifications to a word or name, such as an explanation of the meaning, an indication of context, or a pointer to a word or name, or a pointer to a concept which is described elsewhere by means of an URI or URL. In this way, the term becomes semantically loadable and identifiable. It can then be used as a linking device between 'objects' that carry the term in their description or metadata.
After this process a computer will be better able to interpret the data, and will arrive at more meaningful search results which could lead to new insights and relations between objects.
(..)
There are request from scientific researcher to have an environment where they can work together on improving data and exchange information. The research environment would benefit greatly to have access to such a platform for its own research projects. It would be interesting that a project could get an own space on the platform where only those that are defined as research partner should get access to upload and edit data and thesauri on that part of the platform.

The technology for realizing this semantic web capabilities is the so called Resource Description Framework (RDF), designed to express in the simplest possible way something about something (a resource). Therefore the tool should allow converting the aggregated data to RDF stores controlled by pipelines or another mechanism to enrich the base data. This will also require the thesauri data to be converted to RDF in order to be used for this enrichment when needed. All the aggregated collection data from partners will have to go through the same process. At the end of this enrichment process the data will be ready to use all these RDF stores for comparative research and data interchange.

New digital services proposed

- **Database creation environments with possibility of sharing : aggregation / collaborative databases**

- **Database analysing tools specialized for DCH : 2D, 3D, audiovisual items and metadata advanced research tools - semantic research**