The Design of a DLS for the Management of Very Large Collections of Archival Objects

Maristella Agosti, Nicola Ferro and Gianmaria Silvello

Information Management Research Group (IMS)
Department of Information Engineering
University of Padua, Italy
Outline

- Pursue Interoperability: Digital Libraries
- Archives and Archival Descriptive Metadata
- The Conceived Distributed DLS Architecture
- Extension to Compound Digital Objects
- Conclusions and Future Directions
Outline

- Pursue Interoperability: Digital Libraries
- Archives and Archival Descriptive Metadata
- The Conceived Distributed DLS Architecture
- Extension to Compound Digital Objects
- Conclusions and Future Directions
Pursue Interoperability: Digital Libraries

Archives → Libraries → Museums

M. Agosti, N. Ferro and G. Silvello
The Design of a DLS for the Management of Very Large Collections of Archival Objects
OAI-PMH is a flexible and lightweight protocol for metadata harvesting. It is becoming the de-facto standard for metadata exchange in distributed environments.

OAI-PMH permits metadata harvesting between different repositories in a straightforward fashion.

OAI-PMH is based on two main components: Data and Service Provider.
Pursue Interoperability: OAI-PMH and DC

- Dublin Core, a tiny and lightweight metadata format is getting the preponderant mean to exchange information in a wide distributed environment.

- DC has been chosen as the minimum common denominator in the OAI-PMH environment.

- Libraries have been using for the couple OAI-PMH and DC since a relatively long time with good results.

- Relevant European initiatives which both enjoy the benefits of OAI-PMH are: The European Library portal, TELplus and DRIVER.
Outline

- Pursue Interoperability: Digital Libraries
- Archives and Archival Descriptive Metadata
- The Conceived Distributed DLS Architecture
- Extension to Compound Digital Objects
- Conclusions and Future Directions
The Design of a DLS for the Management of Very Large Collections of Archival Objects

Context - Hierarchy - Relationships
Archival descriptive metadata are the foremost digital resources shared by the archives. We must retain context and hierarchy information.

Variable granularity access to archival descriptions is an important requirement in order to share them in a distributed environment.

Archival Descriptive Metadata is Encoded Archival Description. EAD retains context and hierarchy but it discards variable granularity.
We proposed the couple DC and OAI-PMH as the means to enable the sharing of archival descriptive metadata and to map the EAD files in shareable metadata format.
Outline

- Pursue Interoperability: Digital Libraries
- Archives and Archival Descriptive Metadata
- The Conceived Distributed DLS Architecture
- Extension to Compound Digital Objects
- Conclusions and Future Directions
The conceived DLS architecture

- The goal of the DLS is to put archival resources together and must take into account the size and the structure of participating archives: also small and medium archives preserve unique and valuable pieces.

- DLS architectures in archival context must take into account two aspects:
  - The maintenance of archives autonomy;
  - The necessity of a coordination view to give an integrated vision of the archives participating the system.

- A DLS has to consider a large number of different archives distributed in a territory facing heterogeneity and interoperability issues.
The Design of a DLS for the Management of Very Large Collections of Archival Objects

- DLS-DEI: Data Exchange Infrastructure based on OAI-PMH.
- In order to retain context and hierarchy information Service Provider must harvest not only DC metadata but also the whole set organization of Data Providers.
DLS Architecture

- DLS-MM Management Level: composed by an application logic and a data logic.
- DLS-UI Presentation Level: presents two interfaces; the first is a general-purpose interface and the second is dedicated to specialized users.
Outline

- Pursue Interoperability: Digital Libraries
- Archives and Archival Descriptive Metadata
- The Conceived Distributed DLS Architecture
- Extension to Compound Digital Objects
- Conclusions and Future Directions
The proposed distributed architecture deals with the heterogeneity issue.

Many archives start to keep not only descriptive metadata but also digital objects; a very large DL is required for managing a wide number of heterogeneous archives and for governing high space demanding digital objects.

CDOs are digital objects that include information about context, provenance and relationships between the resources. CDOs are aggregation of different information combine together in order to shape a logical unique object.

Considering system interoperability, the use of CDOs is challenging.
Compound Digital Objects

- DLS-CDO is built upon the DLS-MM layer and exploits it to manage, share and expose CDOs.

- Data logic appertains to the archives side, whereas application logic is developed on the side of DL.

- DLS-MM links the metadata with the digital objects building CDOs that will be managed by the CDO application.

- CDOs are managed, shared and retrieved through metadata which are the foremost entities that enable interoperability.
Outline

- Pursue Interoperability: Digital Libraries
- Archives and Archival Descriptive Metadata
- The Conceived Distributed DLS Architecture
- Extension to Compound Digital Objects
- Conclusions and Future Directions
Conclusions and Future Directions

- We presented a DLS architecture able to share, collect and manage archival metadata in a distributed environment.

- Lightweight and scalable solution.

- We proposed an extension enabling the management of CDOs.

- Future work will involve new considerations about CDOs evaluating also the outcomes of international initiatives that are working on this field.

- In particular OAI-ORE (Open Archives Initiative – Object Reuse and Exchange) which is studying the CDOs and designing an effective way to expose them in the Web.