Contribution ID: **61** Type: **not specified**

The SIAR Project

The SIAR (Sistema Informativo Archivistico Regionale) is a project which aim is to develop a distributed Digital Library System (DLS) for sharing archive metadata; these are maintained in several archives spread across the Italian Veneto Region. The Veneto Region archives belong to different kinds of institutions and in this context, we have to satisfy a strong requirement for cooperation and interoperability: the autonomy of all these institutions has to be preserved as well as their way of managing and organizing the archives. Furthermore, we have take into consideration the structure of the archives that is strongly hierarchical; throughout this structure it is possible to infer the context information of the archival documents and the meaningful relationships between the documents. In the digital environment the archives and their components are described by the use of metadata; these need to be able to express and maintain such structure and relationships. The standard format of metadata for representing the complex hierarchical structure of the archive is Encoded Archival Description (EAD), which reflects the archival structure and holds relations between documents in the archive. On the other hand to maintain all this information an EAD file turns out to be a very large XML file with a deep hierarchical internal structure. Thus, accessing, searching and sharing individual items in the EAD might be difficult without taking into consideration the whole hierarchy. On the other hand, users are often interested in the information described at the item level, which is typically buried very deeply in the hierarchy and might be difficult to reach.

We have considered each of these requirements and issues to design the SIAR system. On one side we have guaranteed the local bodies management autonomy of their archives and we have built-up a regional coordination so that we can have an integrated global vision of the local archives that participate to SIAR. On the other side we have designed a methodology to overcome the issues concerning the exchange of metadata with a large hierarchical structure, such as the EAD metadata format. The main tool we adopted to share archival metadata is the Open Archives Initiative Protocol for Metadata Harvesting (OAI-PMH); by the means of OAI-PMH local archives, acting as Data Providers, can export their metadata in any XML formats without any change in their internal organization and the Veneto Region, acting as a Service Provider, harvests these metadata providing advanced services such as a public access and a full-text search over them.

Throughout OAI-PMH we have also addressed the issues related to the access and exchange of EAD files. We have proposed the "NEsted SeT for Object hieRarchies"(NESTOR) framework defining two set data models based on organizations of nested sets, which enable representations of hierarchical data structures alternative to the tree. This framework used in conjunction with the set organization of OAI-PMH permits to manage and share archival metadata adding new functionalities to the protocol without any change to its basic functioning. With the couple OAI-PMH and NESTOR, we can set a hierarchical structure of items as a well-defined nested set organization that maintains the relationships between the items just as a tree data structure does and moreover we can exploit the flexibility of the sets exchanging a specific information subset while maintaining the integrity of sthe data. In this way we are able to decompose the EAD files or any other complex metadata formats into an organization of nested sets containing small and shareable metadata files, such as the Dublin Core; consequently we can enable the exchange of archive metadata throughout OAI-PMH, without taking into consideration the whole hierarchy and at the same time maintaining their full informational power.

Primary author: Mr SILVELLO, Gianmaria (University of Padua)

Presenter: Mr SILVELLO, Gianmaria (University of Padua)