The science of context: sharing knowledge among archives, libraries and museums descriptive systems

STEFANO VITALI

s_vitali@archiviodistato.firenze.it

First phase

- integration of descriptions of archival material, books and other art and cultural artefacts in common electronic catalogues thanks to the adoption of common descriptive models, data formats and descriptive rules, mainly bibliographic (ex: MARC format and AACR2 used to describe archival material in the American bibliographic networks) during the Eighties;
- many limits and drawbacks

Second phase: the Web

- the search for a common data format has been replaced by a new vision based on the development of more flexible architectures, capable of searching, retrieving and managing data from heterogeneous sources, through common communication standards, such as the Z39.50 protocol;

- largely adopted in library domain, the Z39.50 has been less popular in archives and museums domains (CIMI and Bath profiles)

- difficulty of establishing a real semantic interoperability between systems which are based on descriptive models governed by dissimilar, specific logic

# Third phase: metadata harvesting

- common gateways to discover and access digital resources produced by cultural institutions (archives, museums, libraries);

- based on metadata harvesting mechanisms, according to the OIA-PMH protocol;

- metadata adopted is the Dublin Core Metadata in one of its versions (simple or qualified).

- DC is used to describe the resources or as exchange and communication format of data originally recorded according to specific formats in data bases or in XML DTD or schema

Examples

- **Michael**:
  - describes "digital collections of museums, libraries and archives from different European countries"
  - consulting the descriptions, the user can acquire a basic knowledge of the content and characteristics of the digital collections and, if interested and if the collection is available on line, he/she can go straight to the related web site

- **BAM, Portal für Bibliotheken, Archive, Museen**:
  - gathers the descriptions of books, archival material and museum artefacts on a central server.
  - the metadata can be searched simultaneously by keywords through the open source search engine Lucene.
  - the short descriptions are linked to the catalogues or the information systems of the institutions participating in the project and to digital reproductions, if they exist.

Potential and limits

- This interoperability can help very much in making easier and more efficient the search for quality digital cultural resources on the Web

- It is worth being adopted more largely at regional, national and international level

- Rather than a real integration of information and knowledge included in archives, libraries and museums information systems, the model aims to assemble research results in lists, whose entries have the same kind of relationships as in a list generated by a Google query.

- Unfortunately the metadata adopted is semantically very poor and can only establish superficial relationships between the objects described in the systems.
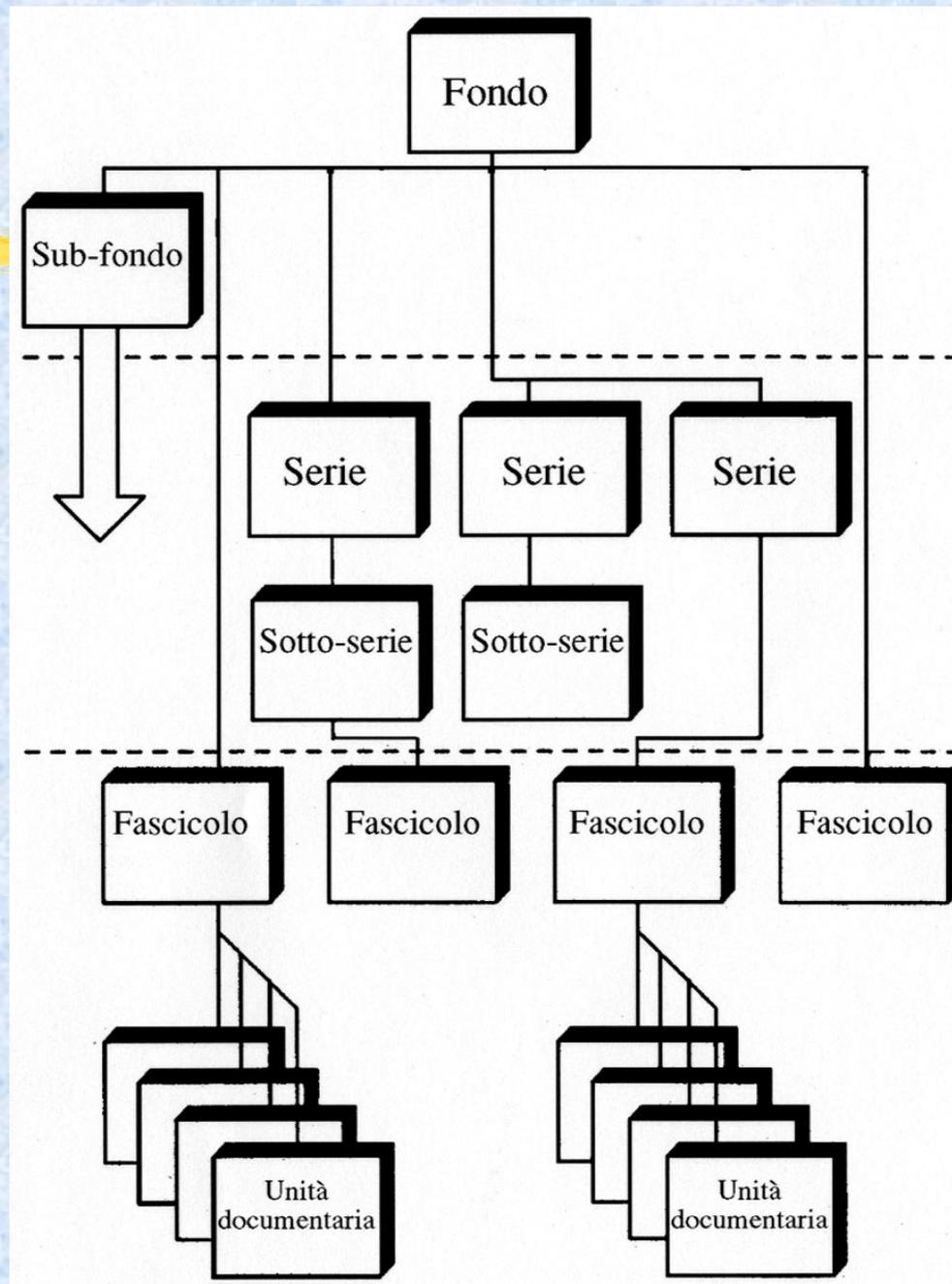
New representation models

- while an oversimplified description of digital resources "has been so extensively used and misused in recent years" in digital libraries and in across domains interoperability projects, more and more sophisticated representation models have been developed to be the basis for building information systems in archives, libraries and museum domains (FRBR, EAD-EAC, CIDOC-CRM, etc.);

- the semantic wealth of those models is always lost when interoperability projects are designed and implemented;

- sharing the knowledge, embodied in information systems, across domains in order to generate new forms and levels of knowledge

Sharing contextual information

shifting from a document-centric approach to the resources discovery processes across the different domains, typical of DC to a different approach based on the integration of contextual information about people, place, events and all the other elements which define the circumstances in which objects held by cultural institutions (archival material, books, art and other museum's artefacts) were created, transmitted over the time, used and interpreted.

Two methodologies

- the model suggested by ISAAR (CPF) based on autonomous and separate systems which describe the objects (archival material, books, art and other museum's artefacts, etc.) but, at the same time, share the authority files of people or other entities which have significant relationships with the objects themselves;

- the CIDOC CRM ontology largely focused on "contextual information": "historical, geographical and theoretical background in which individual items are placed and which gives them much of their significance and value". The convergence between information systems should be based on the "complex interrelations that exist between objects, actors, events, places and concepts in [each] field[s] of cultural heritage"

The experts who work in each domains (archivists, librarians, museum curators etc., but also information technology specialists) should acquire the attitude to learn from each other and to respect more their different points of view.

To say the truth, that seems very often more difficult than make the computer talk each other.

Thanks so much for your attention!