Archival Information Systems in Italy and the National Archival Portal

Stefano Vitali

Director of Archival Supervising Office for the Emilia Romagna Region, Italy

1 Archival Description and Information Systems

The glossary of the International Standard for Archival Description (General), drawn up by the International Council on Archives for the development of archival information systems, defines archival description as "an accurate representation of a unit of description and its component parts, if any, by capturing, analyzing, organizing and recording information that serves to identify, manage, locate and explain archival materials and the context and records systems which produced it."

This definition summarizes the fundamental problems which need to be tackled when developing archival information systems which in a digital environment are the equivalent of the paper finding aids such as guides and inventories, which traditionally are prepared to enable access to and consultation of materials held in archival institutions. The definition underlines the fact that archives are first of all complex "objects". Indeed they are made up of a collection of entities and relations which link them to each other. These relations create strong and specific bonds - "determined" as some scholars of twentieth century archive systems have defined them to stress the fact that these bonds are generated by the common origin of the entities - which even in the digital environment cannot be ignored. It is precisely the nature of these links which distinguishes archives from other "objects" in the realm of cultural heritage (e.g. books and works of art) which in general are perceived as individual and unrelated entities.

Archives (or "fonds" as they are also called) are in fact made up of series which in turn can be organized in sub-series which are formed of archival units (files, registers and so on) which have a homogeneous nature and can in turn be divided into sub-units containing items such as letters, reports, contracts or even photographs, drawings, audio-video recordings and so on. This implies that each of these entities can only be correctly identified and interpreted in relation to the entity they belong to and from which they inherit certain characteristics. This obviously influences the manner of their representation and the retrieval of these representations. The prevailing solution, in the development of digital systems of archival description, has been to represent these relations with hierarchical metaphors which collocate each entity in a vertical relationship of subordination with the entity it belongs to, i.e. with the "father" entity (e.g. a series with the fond it belongs to). The hierarchical representation is further complicated by the fact that the entities that belong to the same father - and are therefore related to each other by a horizontal-type relationship - need to be represented according to a significant sequence which reflects the position that they have in the logical and/or material order of the archive (e.g. in a company archive the series of the articles of incorporation will be located before the series of the deliberations of the board of directors). Both the vertical relations and the horizontal ones, established according to a

pre-established sequence, form the complex archival context of a determined entity and this context contributes in a fundamental manner to its identification.

However, a similarly fundamental role in archival description is played by other types of contexts, which in a certain sense are external to the archives themselves. Archives are in fact historical entities, which like sediment slowly are accumulated within a certain space-time context, generally speaking, as the outcome of the practical activity of certain subjects (corporate bodies – such as institutions or organizations - families and persons).

The context in which a certain archive is created is therefore in itself an essential part of the descriptive system of archives. Each archive therefore should be related with one or more creators who presided over its accumulation and their history, functions, activities, etc. should be described. In addition, when archives held by a number of institutions are described in the same information system, also these institutions have to be described to help users to locate the archives described. The so-called archival institutions, together with archival materials and creators therefore constitute the essential entities which in general make up information systems of archival description in accordance with the three standards issued by the International Council on Archives over the last decade. These standards are: the International Standard Archival Description (General) or ISAD (G), regarding the description of archival fonds and their common components; the International Standard Archival Authority Records (Corporate Bodies, Persons, Families) or ISAAR (CPF), regarding the preparation of authority records for creators; the International Standard for Describing Institutions with Archival Holdings (ISDIAH), which instead is dedicated to describing archival institutions (Fig. 1).

In many systems of archival description, however, the context has a broader meaning and therefore the entities described and related to each other and, in either a direct or indirect way, with the archival materials are more numerous and include the political-institutional settings and jurisdictions where the creators operated, or the previous archival institutions where the archival materials where kept. Other entities can be added to these, such as the description of finding aids existing for a certain fond, bibliographic references, other information resources, etc. with the outcome of creating relatively complex systems (Fig. 2).
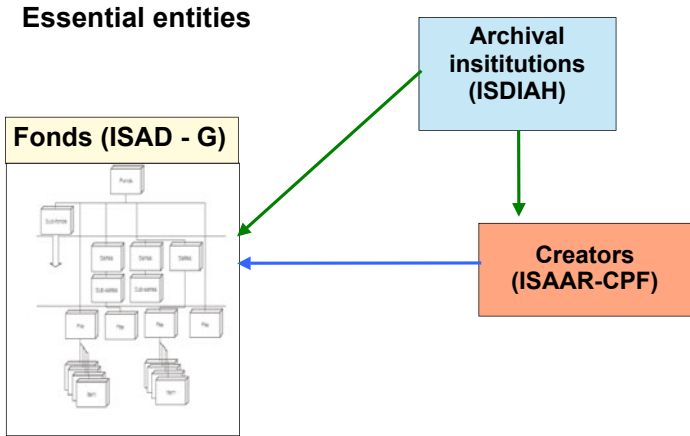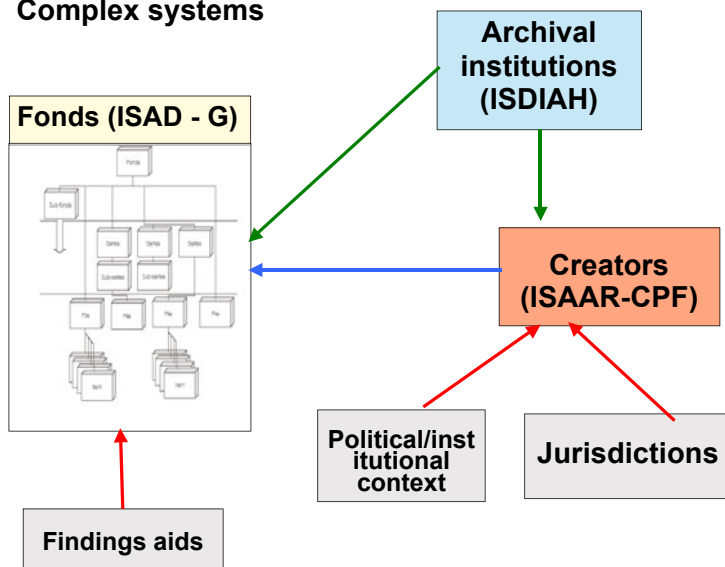

Fig. 1.

Complex systems

Fig. 2.

In terms of the nature of the information stored in the these systems, it should be noted that the archival description is typically a collection of structured data (e.g. dates), semi-structured data (e.g. the physical or logical extent of the archival material, i. e. its quantity, bulk, or size), of narrative texts which may be quite large (e.g. the institutional history or biography of the creator). These data are often uncertain, problematic and attributed according to certain criteria. Precisely due to this problematic nature, the sources of the data has to be indicated too and placed in a historical context. All of this obviously will have an influence on the way this information itself will be represented, treated and retrieved within the system.

2   Archival Information Systems and the Catalogue of Archival Resources of the National Archival Portal[1]

Over the last decade numerous systems of archival description have been developed in Italy, each with characteristics similar to those mentioned above. These systems were developed by the State Archival Administration at the national level (e.g. the General Guide of State Archives [2], the State Archive Information System or SIAS[3], the Unified

---

[1] This section is a re-working of the speech made at the 14th "Archivwissenschaftliches Kolloquium" of the Marburg Archive School (1-2 December 2009) and will be published in part in the "IBC" journal.

[2] Available in its first version at http://www.maas.ccr.it/h3/h3.exe/aguida/findex and in the new 2009 version at URL: http://guidagenerale.maas.ccr.it/.

[3] http://www.archivi-sias.it/.

System of the Archival Supervising offices or SIUSA[4], the Multimedia Historical Archive of the Mediterranean [5]), by some of its local branches (e.g. the State Archives of Florence[6], Milan,[7] Bologna[8], Naples[9], Venice[10]), by some of the Regions (e.g. Lombardy[11] Emilia-Romagna[12], Piedmont[13] or Umbria[14,]), by other local bodies (e.g. the Historical Archives of the Province of Trento [15]), by individual cultural and non-cultural institutions (e.g. the "Giorgio Agosti" Piedmontese Institute for the History of the Resistance and Contemporary Society [16], the Giangiacomo Feltrinelli Foundation[17], the Senate of the Republic[18], the Chamber of Deputies[19] and many others) or by a collection of "federated" bodies (e.g. the Network of Institutions of the Resistance[20] or the project Twentieth Century Archives[21]).

Alongside the development of these systems there has also been a growth in the need to establish links, data exchanges and increasing levels of interoperability between

---

[4] http://siusa.archivi.beniculturali.it/.

[5] http://www.archividelmediterraneo.org/portal/faces/public/guest/.

[6] See the Information System of the State Archive of Florence, or SiASFIi, URL: http://www.archiviodistato.firenze.it/siasfi/.

[7] See the on-line Guide of the State Archive of Milan, URL: http://archiviodistatomilano.it/patrimonio/guida-on-line/.

[8] See the Archival Heritage of the State Archive of Bologna, URL: http://patrimonio.archiviodistatobologna.it/asbo-xdams/.

[9] See the Archival Heritage of the State Archive of Napoli, URL: http://patrimonio.archiviodistatonapoli.it/xdams-asna/.

[10] See the on-line Guide of the State Archive of Venice, or SiASVe, URL: http://www.archiviodistatovenezia.it/siasve/cgi-bin/pagina.pl.

[11] See the archival section on the Lombardy cultural heritage portal, URL: http://www.lombardiabeniculturali.it/archivi/.

[12] See the IBC Archives, URL: http://archivi.ibc.regione.emilia-romagna.it/ibc-cms/.

[13] See the Guarini web archives, URL: http://www.regione.piemonte.it/guaw/MenuAction.do.

[14] See .DOC - Information System of the Umbrian Archives, URL: http://www.piau.regioneumbria.eu/default.aspx.

[15] See the on-line inventories on the site of the Historical Archives of the Province of Trento, URL: http://www.trentinocultura.net/catalogo/cat_fondi_arch/cat_inventari_h.asp.

[16] See ArchOS. Integrated System of Archival Catalogs, URL: http://metarchivi.istoreto.it/.

[17] See the on-line Archives of the Foundation, URL: http://www.fondazionefeltrinelli.it/feltrinelli-cms/cms.find?flagfind=quickAccess&type=1&munu_str=0_6_0&numDoc=95.

[18] See the Archives Project of the Senate of the Republic, which gathers together digitalized descriptions and materials, including fonds preserved at other institutions, URL: http://www.archivionline.senato.it/.

[19] See the site of the Historical Archive of the Chamber of Deputies, URL: http://archivio.camera.it/archivio/public/home.jsp?&f=10371

[20] See the Guide to the Historical Archives of the Resistance, URL: http://beniculturali.ilc.cnr.it/insmli/guida.HTM.

[21] See the Archives of the Twentieth Century. Collective memory on-line, URL: http://www.archividelnovecento.it/archivinovecento/.

them and much has been debated on how to achieve this[22]. One of the outcomes of this debate has been the project developed by the Directorate General of the Ministry of Cultural Assets and Activities which intend to develop a catalog of archival resources, or CAT, within the National Archival Portal dedicated to joining the initiatives of the various institutions which in Italy hold archives (State archival administration, Regions, local independent bodies, universities, cultural institutions, etc.) and to promote knowledge of Italian archival heritage among a wide national and international audience.

Through the development of common access and the provision of concise information regarding national archival heritage, the CAT aims at being a tool for connecting the existing systems without replacing any of them, but rather making them more visible and enhancing their specific characteristics. An operation of this kind can only be carried out because the systems developed in recent years, despite the diversity in the software used and the different aspects of the descriptive formats, share the same conceptual model and a common reference to the international archival standards mentioned above.

The CAT, therefore, shall outline a general map of the national archival heritage able to provide initial orientation for researchers and direct them towards more detailed information resources present in the systems participating in the National Archival Portal. It will include descriptive records of archival institutions, fonds or archival aggregations, finding aids and creators. It will be populated and updated through procedures which favor methodologies of *harvesting* of data from participating systems based on the OAI-PMH protocol. However, other methods of importing data shall not be excluded, as for example the *upload* of an XML file in a specific area of the system according to predefined formats or the direct input of the data in the CAT via web templates. Thanks to this plurality of techniques the aim is to obtain the participation in the portal of even less technologically equipped archival institutions.

The identification of the descriptive elements for fonds and other archival aggregations, for creators and for finding aids has been based on the idea of subsidiarity between systems, and makes reference primarily to those descriptive elements considered mandatory in the international standards, with the integration of few others which are for the most part considered essential in our archival tradition. Concision in any free-text fields shall be ensured by the provision of a number of maximum characters. The data which shall be used to populate and update the CAT data base, as a rule imported by the systems participating in the project, shall be published without any modification. Each of these records shall contain a direct link to the corresponding record present in the system of data provenance. The description of archives will be also connected to those of any digital reproduction project made accessible on the Portal.

Since it cannot be ruled out that more than one description from different systems of the same archival complex, finding aid or creator may converge in the CAT, each imported record is linked to a record describing the relevant system of origin so as to provide context for the origin and characteristics.

---

[22] See for example *Verso un Sistema Archivistico Nazionale?* Special issue of *Archivi e Computer*, XIII (2004), 2, edited by the author of this paper.

For creators, however, a further effort towards standardization is planned with the aim of offering the user higher quality information. In fact the progressive development, by special editorial staff throughout the country, of an authority file of creators, which shall become not only the main access point for searching and navigating in the CAT but also a reference point at the national level for the identification of corporate bodies, persons and families and the formulation of their names. This authority file could also be the reference point for systems participating in the portal in the preparation of their own descriptions of creators, when they do not consider it appropriate, as would be desirable, to directly entrust it with the overall management of the descriptions of the creators. Lastly, it could function as an interface and connection with similar authority files present in the catalogs and descriptive systems of other sectors of cultural heritage, such as the National Library Service.

3   The Definition and Character of the SAN Standard

The architecture of the CAT, its contents and the standards of communication with the participating systems were designed thanks to intense work of comparison and debate which went on right throughout 2009 within working groups appointed in the context of the State-Local Authorities Joint Technical Committee for the Definition of Archival Standards. The working groups saw the important participation of different institutional and geographical entities. In addition to the choice of the descriptive elements for the various entities included in the CAT[23] and the definition of the formats and protocols for importing data into the CAT database, the working groups have developed methodologies for drawing up the authority files of creators. During the first few months of 2010 the formats for metadata of the digital resources shall also be issued, which will be made available on the National Archival Portal, according to methodologies not dissimilar from those used for the CAT. The digital archive of the Portal should contain thumbnails of the images and essential information, which allow the user to search among the digital resources available, make a preliminary selection and then be directly addressed to the harvested systems for quality viewing of the digital reproductions of archival documents.

For exporting the descriptions from the existing systems to the CAT, an XML exchange format has been developed. It has been named "SAN exchange format". It is based on three schemas, each of one includes a subset of elements the most diffuse standards at the international level, i.e. Encoded Archival Description for archival aggregations and for finding aids and Encoded Archival Context for creators (Corporate Bodies, Persons, Families)[24], recently issued. A special exchange format in XML has been developed for the initial population of the records regarding the creators, which when up and running will be produced by the editorial staff of the Portal.

In the development of the XML schema to which the systems should conform in the generation of export files to the CAT, an approach has been adopted which can be

---

[23] See the Technical Subcommittee for the Definition of Metadata (…), *Tracciati descrittivi del CAT: soggetti*, URL:
http://ims.dei.unipd.it/data/san/metadati/docs/2009-04-17_Documento-conclusivo-sui-tracciati-CAT.pdf (provisional address)

[24] See the documentation on the relative site, URL: http://eac.staatsbibliothek-berlin.de/

described as "record centric". Each description of an archival aggregation (first level or subordinate), finding aid or creator to be exported from the systems into the CAT will always have a corresponding record in the export-import file. The relations between the various entities, also with regard to archival aggregations, shall therefore be made explicit through the indication of the identity code of the other connected entities in the export-import file.

The XML schemas of the four entities (archival institutions, creators, archival aggregations, finding aid), integrated with control information required for the correct execution of the import procedures, were lastly grouped together in an overall import-export format, which is available to all managers of archival systems who desire to participate in the National Archival Portal and contribute to the efforts of building a single access point to the Italian archival resources present on the Web[25].

The open and cooperative model for the definition of the SAN standards has already achieved broadly satisfying results. The consolidation of these results will derive from their broad use and shall constitute the precondition for further steps forward in defining exchange formats which embrace other, more complex aspects and components of archival description.

---

[25] The overall schema and the illustrative documentation are provisionally available, respectively, at the following URLs:
http://gilgamesh.unipv.it/cat-import/cat-import.xsd and
http://gilgamesh.unipv.it/cat-import/cat-import.html - id6.