Creator Description
Encoded Archival Context

Daniel V. Pitti
Institute for Advanced Technology in the Humanities
University of Virginia

Introduction

Encoded Archival Context (EAC) is an ongoing initiative within the international archival community to design and implement a prototype standard based on Extensible Markup Language (XML) for encoding descriptions of record creators. The primary developers of this prototype standard are members if the international archival community. The description of individuals, families, and organizations that create records is an essential component of the documentary evidence of human activity. Identifying record creating entities, recording the names or designations used by and for them, describing their essential functions, activities, and characteristics, and the dates and places they were active is an essential component of the management of archival records. Creator description facilitates both access to and interpretation of records.

Description of creators is also essential in bibliographic systems, and in museum documentation, and thus EAC may be of interest to other cultural heritage communities as well. As custodians of the records upon which biographies and organization histories are based, archivists are well-placed to develop a standard that will assist in the fulfillment of their professional responsibilities, and at the same time lay the foundation for building international biographical and organization history reference resources.

Records

Archival records are the evidence of people, acting individually, in families, or in formally organized and named groups. From a strictly archival perspective, records are the byproducts of people living their lives, or carrying out official duties or responsibilities. Archival records are the results of human functions and activities. Records document the conduct of business. As evidence of activities and official functions, they frequently have legal and historical value. Record, broadly speaking, encompasses both the narrower archival definition, but also all artifacts, whether created as byproducts, or as intentional products. "Anything made by human art and workmanship[1]" is thus a record: books, articles, movies, sound recordings, paintings, sculptures, collections of natural objects, and so on.

---

[1] This is the lead definition of "artifact" (or in British spelling, "artefact") in the *Oxford English Dictionary Online (3rd ed.): http://dictionary.oed.com/*

Creator and Record Description

Most standards development work to date has focused on the description of records or resources. While there are notable exceptions, this is true of both the archival community, in the development of Encoded Archival Description (EAD), as well as standards development efforts in other communities. The best-known example is the Dublin Core initiative, which concentrated on basic description of resources to facilitate their discovery. The library community has long had standards for both the description of bibliographic objects, as well as for uniquely identifying individuals, corporate bodies, and conferences. The library community's description of creators, however, is concentrated on controlling names, and not on detailed description documenting people and organizations bearing these names. In other words, library authority control standards serve bibliographic or resource description by controlling the headings or entries used therein.

Archivists have a professional responsibility for developing a standard for creator description. Archival records function as both legal and historical evidence. It is essential that the context of the creation of records be documented. In order for users to evaluate, understand, and interpret records, they need to know the circumstances that surrounded the records creation and use. Essential in the documentation of this context is recording information about individuals, families, and organizations. In particular, such creator description needs to document the name or names used, biographical or historical information about them, and information concerning their activities and responsibilities.

In addition to being professionally motivated to develop a standard for creator description, archivists are also in a unique position for doing so. While libraries, museums, and archives are all responsible for the preservation of the documentary evidence of human activities, archives in particular are responsible for the official records and personal papers that are considered the primary evidence on which biographical and historical description is based. While archivists are by no means the exclusive authors of biographies and organization histories, they are the custodians of unique resources and they have a special responsibility for developing adequate contextual information about those resources. This professional responsibility places archivists in a strong position to coordinate and develop a standard for creator description.

General Characteristics and Methods

In general, institutions and individuals are willing to develop and adopt standards if the benefits outweigh the sacrifices. Standardization of creator description offers both economic as well as professional benefits. Anyone familiar with authority control and creator description knows that it is an expensive undertaking. Description of individuals, families, and organizations frequently

involves detailed and challenging research, followed by careful composition of the description. This expense is in addition to the description of records and other resources. It is quite frequently more time consuming than the description of the records themselves. Records with a common provenance are frequently dispersed within or shared by more than one repository. In such instances, the creator research and description done by one archive, if based on a standard, can be shared and enhanced by other repositories, thereby distributing the costs. A standard also offers the potential for importing descriptive information from sources outside of the archival community, and adapting and enhancing such information to meet descriptive objectives.

In addition to economic benefits, a standard for creator description will provide professional benefits. A semantic and structure standard will facilitate the accurate representation of creator description that will enable effective access to and description of archival records. A vexing and ongoing challenge is that individuals, families, and organizations frequently conduct business under different names. Both archivists and public users frequently have difficulty in locating records simply because the name used to document the provenance of records is not the same as or significantly differs from the name with which they are familiar. A creator description standard will provide a means to uniquely identify creating entities and to document all of the names used by the entity. Further, a creator description standard will facilitate effective documentation of the critical characteristics of creator entities, and indexing the characteristic information can further enhance access.

In addition to more effectively achieving long held professional documentation and access objectives, a standard for creator description will facilitate building international, national, regional, and institutional biographical and historical databases that can serve as resources. Through links to record descriptions, creator descriptions can serve as a gateway to records. Creator descriptions can also function as an independent resource for users seeking information about individuals, families, and organizations.

While there are significant national and institutional differences in archival descriptive practice, creator description is inherently a major component of that practice. Responsibly and effectively preserving and providing long-term access to records necessarily involves documenting the context in which records were created. Central to this context documentation is the identification and description of creating entities.

While there are exceptions, the vast majority of archival description has been based on provenance. All of the records originating from one individual, family, or organization are preserved as a unit, and described collectively. Given the provenance-based approach to organization and description, there is a one-to-one correspondence between the archival descriptive unit, and a creating entity. Such a one-to-one correspondence makes it logical to document both the creator

and the records created in the same descriptive apparatus. Thus creator description has traditionally been an integral component of archival description.

The opportunities for improving archival practices and services presented by computers and network technology have inspired archivists to engage in a new analysis of archival description. The challenge of effectively and economically representing description in computers has forced a rigorous analysis of the logic and structure of the description. This analysis is leading to increasing differentiation and formal definition of the components of description and the relations between components.

While traditional archival description documents creators, records, and functions in one descriptive apparatus based on provenance, markup and relational database technologies are inspiring archivists to envision new systems that use distinct apparatus for each component, and that then dynamically interrelate them to form a complete archival descriptive system. Archivists increasingly recognize that the single apparatus is inflexible and inefficient in dealing with complex records.

While it is possible to establish the provenance of most records, it is common for records to be of mixed provenance, or records of the same provenance to be dispersed. Providing creator or context information in such common situations requires repeating information in more than record description. When records with a common provenance are dispersed in different repositories, it frequently means that expensive creator research and description is duplicated. The relations between functions, creators, and records also present problems. Within groups of records with a common provenance, it is frequently possible to identify groups of records that document the same activities and function. But functions and activities are not fixed in one organization or person. They frequently are shared by two or more creators, or transferred from one creator to another. In a descriptive system based on provenance, sharing or transferring functions leads to the descriptive and physical separation of records documenting the same function or activity.

Relations between records, creators, and functions and activities are dynamic and complex, and not fixed and simple. Creators are related to other creators. Records are related to other records. Functions and activities are related to other functions and activities. And each of these is interrelated to the others. Markup and relational database technologies enable developing flexible and dynamic descriptive systems. By developing dedicated semantics and structures for describing each descriptive component and their complex interrelations, we can build descriptive systems that are far more efficient and effective than those we have realized in print.

Developing a descriptive system for creators related to systems for describing records and functions and activities will enable the creator description to play

more roles than providing context for the origination of records, as essential and central as this role is for archival description. Creator descriptions can function as a first and important step in the discovery of records, as well as discovery of related creating entities, and functions and activities. Pursuing any relation, will reveal new constellations of relations, and so on. Independent of relations to other entities, creator descriptions can function as biographical and historical resources. Relational and markup technologies thus offer us the opportunity to develop a flexible, dynamic, sustainable descriptive system that is far more useful than traditional print finding aids.

EAC-Prototype Standard for Creator Description

The effort to develop a standard for creator description is taking place within the context of related initiatives within the international archival community. In 1996 the International Council on Archives (ICA) published *International Standard Archival Authority Record for Corporate Bodies, Persons and Families* (ISAAR(CPF)).[2] Under the auspices of ICA's Ad Hoc Commission on Descriptive Standards, work on this structural standard was initiated in 1993. While ISAAR has served as the point of departure for the EAC efforts, the initial design of EAC has not been constrained by it. ISAAR is currently under review, and several members of the ICA committee reviewing ISAAR are represented in the EAC initiative and the initial work on EAC will inform the review of ISAAR.[3] At the same time that ICA was developing ISAAR, there was an American effort to develop an SGML-based prototype standard for archival records description (or finding aids). This initiative eventually developed into an international effort, and resulted in the release of version 1.0 of Encoded Archival Description (EAD) in 1998. EAC is intended to extend and complement EAD. EAC will support the descriptive needs of the archival community, specifically in the creation, maintenance, and publication of creator description.

In addition to the international archival community, the EAC effort will be useful in the broader context of libraries and museums, formal authoring and publishing of biographical and organization histories, and the amassing of large genealogical databases. While the EAC working group anticipates interrelations between EAC descriptions and genealogical information, it should also accommodate existing authority, biographical, and historical data. There are several possible sources of authority and biographical and historical data. The Library of Congress Name Authority File contains approximately five million records.[4] There are also a large

---

[2] International Council on Archives (ICA) published *International Standard Archival Authority Record for Corporate Bodies, Persons and Families:* http://www.ica.org/isaarf.html.
[3] See Appendix A for a list of participants in the EAC initiative. Members of the EAC working group participating in the ICA review of ISAAR are noted.
[4] This data is taken from the Library of Congress Cataloging Distribution Service Web page: http://lcweb.loc.gov/cds/mds.html#na. The five million records also include geographic names, uniform titles, and series headings. EAC is not intended to accommodate these.

number of national biographical and historical initiatives, both public and private. While there will necessarily be rights issues associated with some of the existing data, it is anticipated that significant portions of the existing data will be converted to EAC and imported into one or more EAC databases.

Wendy Duff (University of Toronto) and Richard Szary (Yale University) first proposed an effort to develop an encoding standard for creator or context description in 1998. With the assistance of the author, and with funding from the Digital Library Federation in the United States, they organized a meeting held at Yale University in 1999. The effort was slow to continue until 2000, when, with the encouragement and assistance of Anne Van Camp of the RLG, funding was secured from the Gladys Krieble Delmas Foundation, and a meeting was organized and convened at the University of Toronto in March of 2001. The organizers attempted to identify and select international recognized archival description experts and supporting technologists as participants. In addition to selecting recognized experts, the organizers also sought participants with experience in working collaboratively and cooperatively in the development of standards and best practices.

When the working group met in Toronto, its initial efforts were devoted to a free and open exchange of ideas, a general methodological framework as well as a detailed list of principles and objectives to guide the design. The working group explicitly acknowledged that standards are intellectual and technical products, but also inherently political products. Cooperation and consensus are absolutely essential, and participants required a willingness to collectively create and shape ideas. A successful standard needs to embody agreement sufficient to be generally useful for the development of national and institutional systems of creator description, and the exchange of descriptive information between systems. At the same time, the standard must accommodate national and institutional differences. With this design constraint in mind, the working group is attempting to agree on as much as possible, while leaving room for differences that do not undermine shared objectives. Such a process is necessarily iterative. Each set of objectives needs to be provisionally implemented and the prototype standards evaluated with respect to objectives. (A complete list of participants is provided in Appendix A, and the detailed EAC design objectives and principles, the *Toronto Tenets,* can be found in Appendix B.)

Structure and Semantics

EAC development is only in its initial stages. An Alpha version of the DTD has not yet been released, but the developers have made significant progress in defining the structure and semantics of the prototype standard, and have begun early experimentation with descriptive data. The following description of the structure and semantics is provisional. While the overall structure has achieved a

measure of stability, many of the details need additional analysis, elaboration, and testing.

Each EAC document contains two mandatory elements, the <eacheader> and the <condesc>-context description. The <eacheader> contains data used in the control of the creator description, and to provide the context of the description. The <condesc>-context description encompasses the description of the creator. Both the <eacheader> and <condesc> contain specific elements to support the functional intentions of the parent element.

The <eacheader> contains the following subelements:

- **<eacid>-eac identifier.** Contains unique identifier for the descriptive document within the owning system. Accommodates both machine- and human-readable[5] versions of the identifier. Required.

- **<mainhist>-maintenance history.** Contains one or more <mainevent>-maintenance events that document creating, importing, updating, and deletion of the description. Each maintenance event contains the name of the person or system responsible for the event, date, and description of the event. Each <mainevent> has a MAINTYPE attribute to accommodate one of four possible values: create, update, import, or delete. <date> and MAINTYPE are machine-readable. Required.

- **<languagedecl>-language declaration.** Contains one or more machine- or human readable declarations of the language of the description. Optional.

- **<ruledecl>-rules declaration.** Contains one or more machine- or human readable declarations of the content rules used in the creation of the description. Optional.

- **<sourcedecl>-source declaration.** Contains one or more machine- or human readable declarations of the sources for the information used in the description. Optional.

- **<authdecl>-authority declaration.** Contains one or more machine- or human-readable declarations of authorities from which either descriptive categories or values are taken. Optional.

---

[5] In the description of the elements, machine-readable means that the information can be used to provide a traversable link to a resource or resources, or can be used to link related information within the <eac> through attributes of type ID and IDREF. All of the declarations in the <eacheader> play both roles: lead to resources, and are (or can be) associated with descriptive information in the <condesc>. Human readable means that the same or related information is supplied in a form that is intelligible to a human being.

In addition to the subelements, the <eacheader> element also contains several attributes. The TYPE attribute is used to designate the type of creator described in the EAC document: corporate, personal, or family. The STATUS attribute is used to designate the editorial status of the description, draft, edited, or deleted. The ENCODINGANALOGSYS is used to designate the system in which there are semantic analogs for descriptive values used in the EAC document. This attribute is used in conjunction with the EA-encoding analog attribute, available on EAC descriptive elements. The value given in an EA attribute is the analog designation for the containing element used in the system identified in the ENCODINGANALOGSYS.

Four other attributes associated with the <eacheader> are used to designate authorities for values used in the description or rules for formulating such values. These have default values:

- **LANGENCODING** Code values for language of description. Defaults to ISO 639-2b

- **SCRIPTENCODING** Code values for script used. Defaults to ISO 15924

- **DATEENCODING** Rules for formulating normalized date values: Defaults to ISO 8601

- **COUNTRYENCODING** Code values for designating countries. Defaults to ISO 3166-1 a2

- **OWNERENCODING** Code value rules for repository or owner codes. Defaults to ISO 11551

The <condesc>-context description comprises the description of the creating entity. Similar to the <eacheader>, <condesc> has several complex subelements used to describe different features of the entity:

- **<identity>** Complex structure containing the name or name used by the entity over the course of his, her, or its existence. Required

- **<eacrels>** Contains references to descriptions of related individuals, families, or organizations. Optional

- **<resourcerels>** Contains references to descriptions of related archival, bibliographic, or museum resources or records. Optional

- **<funactrels>** Contains references to descriptions of related functions or activities. Optional

- **<desc>** Contains formal description of entity characteristics as well prose or chronological list biographies and histories. Optional

The most complex element in the EAC DTD is the <identity>. In addition to needing to accommodate one or more names used for or by the entity, <identity> must accommodate two or more parallel names in different languages or scripts. In countries where there is more than one official language, such as Canada, names of corporate bodies are frequently provided in more than one language.

The <identity> contains the following elements:

- **<legalid>** Legal identifier for the individual, family, or organization. Optional.

followed by one or more from each of the following pairs of elements. Within each pair, the <*grp> can be intermixed with the <*head>, but the pairs are exclusive:

- **<persgrp> or <pershead>** one or more personal name-groups, for containing one or more parallel personal name headings; or one or more personal name headings.

- **<corpgrp> or <corphead>** one or more corporate name-groups, for containing one or more parallel corporate name headings; or one or more corporate name headings.

- **<famgrp> or <famhead>** one or more family name-groups, for containing one or more parallel family name headings; or one or more family name headings.

followed optionally by one or more of the following:

- **<nameadds>** name-additions contains subelements for distinguishing additions to the base heading. While additions can be made directly within the <*grp> and <*head> elements when they are used uniquely within the <*grp> or <*head> to qualify names, they can, when shared by all of the headings, be contained here and shared in indexing, sorting, and display of the headings

- **<didentifier>** digital-identifier contains machine-readable reference to internet accessible digital portrait or other non-textual digital identifiers of the described entity

The <pershead>, <corphead>, and <famhead> elements each contain the same subelements:

- **<part>** contains a part of the name. A TYPE attribute may be used to provide a precise designation of the name component, "forename," "surname," "parent body," and so on. Repeatable.

followed by:

- **<date>** contains the life dates of individuals, or the active dates of families and organizations. It should not be confused with the <usedate>, which contains the date or date range when the name was used by or for the entity. Optional.

- **<place>** contains the name of a place associated with the heading. A TYPE attribute may be used to provide a precise designation of the role of the place name in relation to the heading, for example, "Birthplace." Optional.

- **<nameadd>** contains additions made to the base name to distinguish it from the same name used for another entity, or to enhance the base name's intelligibility. A TYPE attribute may be used to specify a precise designation for the addition, for example, "expansion" for expansion of initials used in the name. Repeatable.

- **<usedate>** contains the date or date range when the name was used by or for the entity. It should not be confused with <date>, which contains the life or active dates of the entity. Optional.

The <persgrp>, <famgrp>, and <corpgrp> elements each contain two or more <*head> elements of the same entity type as the parent element. After the <*head> elements, the following may be used: <nameadds>, <sourceref> or <sourcerefs>, or <note> or <notes>.

The following elements are available directly inside <identity>, the <*grp> elements, and the <*head> elements.

- **<sourcerefs> or <sourceref>** <sourcerefs> contains two or more <sourceref>; <sourceref> contains both a reference to the resource used in composing a heading or headings that is declared in the <eadheader> using the <sourcedecl>. The <sourceref> also contains a <sourceinfo> subelement for containing a transcription of the source information used in composing one or more headings

- **<notes> or <note>** <notes> contains two or more <note>; <note> contains editor's description of judgments and decisions not otherwise documented in the declarations made in the <eacheader>, evaluations of the evidence when there are contradictions or suspected or known inaccuracies, and so on.

Because <nameadds> is a grouping element for <nameadd> that facilitates economic reuse of name additions, it is directly available in <identity> and the <*grp> elements. The <nameadd> element is available directly inside the <*head> elements because its use there intended to be specific to the heading.

Similarly, the <sourcerefs>, <sourceref>, <notes> and <note> elements available directly inside a <*head> element are intended to apply only to the specific heading, while those available inside <identity> and the <*grp> elements are intended to apply to all sibling <*head> elements.

The <identity> element is intended to facilitate control of the names used by and for an entity. Unlike traditional authority control, the notion of "authorized heading" and "variant headings" is not explicitly privileged in the naming of the elements. Instead, there is an AUTHORIZED attribute. To privilege one heading over the others for indexing, sorting, or display, users enter their owner or repository code in AUTHORIZED attribute. Where more than one heading is authorized within the context of a descriptive system, or different headings are authorized for different contexts, there are additional attributes available: RULE, the descriptive rules used in the composition of the heading; LANGUAGECODE, the language of the heading; and SCRIPTCODE, for the script of the heading. For example, within the context of the Archive of Ontario, parallel French and English headings can be designated by placing the repository code of the Archive in two parallel <corphead> elements using the AUTHORIZED attribute, with the two different headings being distinguished by the values in the LANGUAGECODE.

Relations

As a component of archival description, creator description must be brought into relation with the other descriptive components. Creator descriptions must be dynamically related to the record descriptions for which they provide context, and the functions and activities in which the creators engage and that the records document. With the exception of unique relations, it is the nature of relations that they take place between entities and not within them. Creators are related to other creators, to activities and functions, and to records. Similarly, activities and functions are related to other functions and activities, creators, and records; and records are related to other records, to creators, and to functions and activities. Each creator, record, or function/activity description can thus function as a node in a set of relations.

Because relations are *between* the descriptive nodes, they are most efficiently created and maintained outside of each node.[6] A person, for example, can be related to one or more persons, organizations or families, to one or more archival records, books, journals, and museum objects; and to various occupations and

---

[6] This is a cardinal principle of both relational database and hypermedia theory.

11

activities. Each of the related entities can be related to one or more other entities. To record all of these relations in the description of each node is inefficient, as correction of an error would require updating two or more descriptions.[7]

While maintaining relations independent of descriptions is efficient, when communicating descriptions between systems or to users, it will be necessary to assemble or gather and represent the related descriptions using surrogates. Each surrogate for a related description will optimally include both human- and machine-readable information. The human-readable information provides succinct description of the related entity, creator, records, or function/activity sufficient to enable identification and a relevancy judgment. The machine-readable information supports a traversable link to the related description.

There are three elements for describing EAC relations with other descriptive entities: <eacrels>-EAC to EAC relations, <resourcerels>-EAC to resource relations, and <funactrels>-EAC to function and activity relations. The <eacrels> contains one or more <eacrel> as well as <sourcerefs>, <sourceref>, <notes> or <notes>, for documenting the source or sources of information documenting the relation, and descriptive notes. The <resourcerels> and <funactrels> are similarly structured, though the principal contained element is <resourcerel> and <funactrel> respectively. The <*rel> elements have attributes to facilitate traversable links to resources described in contained descriptive surrogates.

Each <*rel> has a RELTYPE-relationship type attribute, with a closed list of primitive values.

<eacrel> has the following RELTYPE values available:

- superordinate
- subordinate
- earlier
- later
- associative
- coordinate
- identity
- other

<resourcerel> has the following RELTYPE values available:

- origination
- destruction
- control
- causa

---

[7] In order to assist in the design of EAC, a provisional, related DTD, called Encoded Archival Relations (EAR), was developed to document relations.

- other

Since there has been no attempt at this time to design and implement function and activity description, preliminary decisions concerning the typing of EAC to function and activity relations were deferred.

There are two principal rationales behind the primitive or basic typing of relations. First, there is general interest in being able not only to enable coherent expression and navigation of relations, but also the ability to create graphic displays of organizational charts, family trees, and timelines. The primitives are an experimental attempt to provide the data necessary to construct such displays. At this point, there has been no attempt to test the utility of the structures with graphic displays. Second, basic information about the nature of relations is necessary to display to users in order to make the relation intelligible. Given cultural and institutional differences, the number of possible relation types is, in principle, unlimited. EAC designers decided, though, that to achieve a minimum level of functionality that there needed to be consensus on a set of basic or primitive relation types.

Each <*rel> element has one or more subelements available for representing surrogate description of a related entity. Because <eacrel> is self-referential, the surrogate description presents no major semantic and structural difficulties: <eacrel> simply contains <persname>, <famname>, and <corpname>, which accommodate the heading subelements in <pershead>, <famhead>, and <corphead>. The designers chose to provide a minimally structured element, <funact>, to accommodate surrogate representation of function and activity description, pending development of a descriptive structural standard for function and activity description.[8] In essence, the <funact> element is merely a placeholder.

The representation of surrogate information for records presents difficult technical challenges because any EAC entity can in principle be related to records, broadly defined, outside the control of archivists, and therefore outside the scope of archival standards. For example, EAC documents may be related to archival records as well as books and journals, for which librarians have responsibility, and museum artifacts and collected natural objects, for which museum catalogers have responsibility.

EAC to EAD relations can be addressed by negotiations within the archival community to reconcile EAC and EAD semantics and structure. Bibliographic descriptions and museum descriptions must be accommodated using a different

---

[8] EAC designers assume that the international archival community will, in the near future, undertake an attempt to develop a standard for function and activity description. The <funact> element and related elements would necessarily need to be modified for compatibility if and when a standard emerges. A similar strategy was adopted in the development of EAD with respect to elements now being more rigorously defined in EAC.

strategy. Since the primary function of the information is to make surrogate intellectual description to provide context for the presence of a traversal link to a related resource, the elements need only accommodate a minimal semantics and structure. The surrogate need not support detailed, sophisticated searching, as this is addressed in the referenced description. It need only facilitate coherent display of the description of the resource that will be sufficient for the user to decide whether or not to pursue the link. An alternative approach would use XML Namespace, which supports incorporating different semantic and structural XML standards into one document.[9]

Assuming and pending more robust implementation of XML Namespace, the designers of EAC have chosen to provide minimally defined surrogate elements inside <resourcerel> for bibliographic and museum description: <bib>-bibliographic description and <mus>-museum description. The <arch>-archival description elements contain all of the EAD <did>-descriptive identification subelements, though at this stage of development it is not entirely compatible with the EAD <did>. The <bib> contains a minimal set of elements to support a basic bibliographic citation. Assuming and pending the emergence of one or more museum description encoding standards, <mus> element has the same structure and semantics as <bib>.

Description

The <desc>-description element accommodates a variety of both controlled and prose description of creators. Three elements are available for grouping controlled or element specific description of each entity type: <persdesc>, <famdesc>, and <corpdesc>.

<persdesc> contains the following:

- **<legalstatus>** legal status

- **<sex>**

- **<location>**

- **<descentry>** descriptive entry

followed optionally by any of the following

- **<funactdesc>** function or activity description

- **<character>** personal characteristics

- **<env>** environment

---

[9] For more information on Namespace, see http://www.w3.org/TR/REC-xml-names/

- **<ocd>** other context description

<corpdesc> contains the following elements:

- **<corptype>** corporate body or organization type

- **<legalstatus>**

- **<location>**

- **<descentry>** descriptive entry

followed by:

- **<causa>** mandates and warrants

- **<funactdesc>** function or activity description

- **<assetstruct>** assets and administrative structure

- **<env>** environment

- **<ocd>** other context description

<famdesc> contains the following elements:

- **<legalstatus>** legal status

- **<location>**

- **<descentry>** descriptive entry

followed by:

- **<funactdesc>** function or activity description

- **<assetstruct>** assets and administrative structure

- **<env>** environment

- **<ocd>** other context description

There are two structurally distinct classes of <*desc> subelements. The first class, represented by the elements in the above lists that precede the phrase "followed by," are intended to provide controlled vocabulary description of

important characteristics of the described entity. Each of these elements represents a particular descriptive category and has the same subelements: <value>-value, optionally followed by <date, <place>, <note>, and <sourceref>. The <descentry> element is used as a repeatable means of expanding the descriptive categories, and thus has a TYPE attribute for specifying the category.

The second class of <desc> elements is similar to the first, except that it enables using one or more category-value pairs for representing aspects of the parent descriptive category, followed optionally by prose description. The <descentry> is used for representing the category-value pairs. <ocd>-other context description is a generic element used when no other element of this structural type is appropriate. Like <descentry>, <ocd> has a TYPE attribute for designating the descriptive category.

The <bioghist>-biography/history element, borrowed from EAD, can optionally be used for prose description of any entity type. It enables simple or complex, brief or lengthy biographies and organization histories. Particularly noteworthy among its subelements is the <chronlist>-chronological list, which enables a succession of two or three part entries <date>, <event>, or <date>, <place>, <event>.

TYPE Attribute: One and the Many

As an international effort, the designers of EAC are attempting to agree on as much as possible, while accommodating cultural and institutional differences. The semantics and structure described above represents the current semantic and structural consensus.

In addition to the elements <descentry> and <ocd>, described above, several descriptive elements also have TYPE attribute that accommodate arbitrary textual content to facilitate national, regional, and local extensions to the EAC semantics. <date> and <place> are widely available and thus may be used to qualify other descriptive information in a variety of ways. Other elements, in particular elements representing an abstraction of several suggested and more specific alternatives, also bear the TYPE attribute.

It is widely recognized that such extensions can undermine communication and collaboration objectives. To ameliorate this danger, each TYPE attribute is accompanied by two related attributes, TYPEAUTH, and TYPEKEY. The TYPEAUTH provides a means to reference an authority declared in the <eacheader> using <auth>, and through <auth>, to reference the authority when it is Internet accessible. The TYPEKEY attribute provides the unique identifier for the particular term or phrase in the authority. A similar set of attributes is used to specify the values used in the <value> element: VALUEAUTH and VALUEKEY.

Conclusion

The EAC initiative is just underway. In addition to the two meetings held in New Haven and Toronto, a third meeting was held at the University of Virginia this past June. Two small groups met. The first group worked on elaborating and formalizing the semantic and structural design started in Toronto, and one to write an XML Document Type Definition (DTD) based on the design; and another to develop a strategy for ongoing development.

An Alpha version of the EAC DTD is currently under development. The first draft is complete and preliminary testing by the participants is underway. Several activities will follow the successful testing and revision of the DTD. An Alpha tag library will be written, the DTD will be made available for wider testing, and one or more training sessions will be organized to ensuring that those testing the DTD understand its semantics and structures. The EU funded Linking and Exploring Authority Files project will play a major role in the testing and evaluation, using data from a wide number of European repositories. The EAC listserv, hosted by Yale University, will also be made public, to facilitate wider participation and discussion. A conference will also be organized following the testing to evaluate the results of the testing and, if the evaluation is favorable, to plan the steps necessary to complete development.

Given the preliminary status of EAC, it is far too early to predict its success. There are a great many intellectual, technical, and political challenges to be met. Nevertheless, all participants agree that the initial effort has been encouraging.

Appendix A

Linking and Exploring Authority Files (LEAF, a EU funded project

Tone Merete Bruvik (University of Bergen; EU LEAF)

Adrian Cunningham (National Archives of Australia)

Wendy Duff (University of Toronto)

Joanne Evans (Australian Science and Technology Centre, University of Melbourne)

Margaret Hedstrom (University of Michigan)

Hans Hofman (Information Policy Department, Ministry of the Interior, The Netherlands)

Gunnar Karlsen (University of Bergen; EU LEAF)

Bob Krawczyk (Archives of Ontario)

Michelle Light (Yale University)

Gavan McCarthy (Australian Science and Technology Centre, University of Melbourne)

Per-Gunnar Ottosson (Riksarkivet, Sweden; EU LEAF)

Daniel Pitti (Institute for Advanced Technology in the Humanities, University of Virginia)

Kathleen Roe (New York State Archives)

Dick Sargent (Historical Manuscripts Commission, U.K.)

Richard Szary (Yale University)

Anne Van Camp (RLG)

Stefano Vitali (Archivio di Stato di Firenze)

Stephen Yearl (Yale University)

Toronto Tenets
Principles and Criteria for a Model for Archival Context Information

March 2001

This document defines principles and criteria for designing, developing, and maintaining a representational scheme and communication structure for archival context information.

A description of archival records sufficient to support the accurate interpretation of the records must include a description of the circumstances that surrounded their creation and use. Primary among these circumstances is a recording of information about the creative responsibility for the records, usually vested in an organization or person(s). With this information, users can understand the records more completely since they will know the context within which the organization or person operated and created records.

This model primarily addresses the description of creating entities, a central component to the description of archival records, and clearly an archival responsibility. It recognizes the existence of other information, such as functions and business processes, geographic places, events, concepts, and topics, that are crucial to archival description, which are also important, but which may be defined more fully by other agencies and not included in this model.

While traditional heading control functions may be accommodated by this model, its primary purpose is to standardize descriptions about records creators so that they can be discovered and displayed in an electronic environment, linked to each other to show/discover the relationships amongst record-creating entities, and linked to descriptions of records.

Definitions and Uses

1. Archival context information consists of information describing the circumstances under which records (defined broadly here to include personal papers and records of organizations) have been created and used. This context includes the identification and characteristics of the persons, organizations, and families who have been the creators, users, or subjects of records, as well as the relationships amongst them.

2. Context information is not metadata that describes other information resources, but information that describes entities that are part of the environment in which information resources (i.e., records) have existed.

3. The recording of context information in archival information systems directly supports a more complete description and understanding of records as well as the provenance approach to retrieval of these records across time and domains.

4. Context information also can have value as an independent information resource, separate from its use in supporting the description, retrieval, and interpretation of records.

5. This model is also intended to support the exchange and sharing of context information, especially in those instances where repositories have holdings or interests that have context information in common, especially about creators or subjects of records.

Structure and Content

6. Context information has traditionally been embedded in catalog records, finding aids, and other archival descriptive tools. This model can be used either as a component of existing descriptive approaches that fully integrate contextual information into descriptive products or as an independent system that is linked to descriptive systems and products.

7. Each instance of context information describes a single entity.

8. The model provides a framework within which the full range and depth of context information can be recorded and suggests a minimum set of elements for describing an entity, but defers recommendations for appropriate use of other elements to application guidelines developed for specific implementations.

9. The model defines a universe of elements used to describe entities and the structure of interrelationships amongst those elements. These elements and structure support the discovery, navigation, and presentation of context information and the linking of that information to descriptions of records, especially those encoded according to EAD, MARC, and similar standards.

10. The model supports the linking of descriptions of contextual entities to digital or other surrogate representations of those entities.

Technical Issues

11. The model is expressed as an XML-compliant document type definition to encourage platform independence and portability of information. The model may also be implemented using other approaches.

Components, Relationship to ISAAR(CPF), and Ownership

12. Two parts: dtd and guidelines.

13. The model was designed as an implementation of the International Standard for Archival Authority Record for Corporate Bodies, Persons, and Families -ISAAR(CPF). ISAAR(CPF) was under review at the time the model was being developed and the model may incorporate different approaches than that defined in the original ISAAR(CPF) standard. Principles and approaches adopted for the model will be submitted to the International Council on Archives Committee on Descriptive Standards to inform their review of ISAAR(CPF). It is expected that the model will fully conform to the revised ISAAR(CPF).

14. Responsibility for control and maintenance will be carried out by Yale for some period of time and the original working group will continue to develop the model until it is appropriate to be opened to a wider community for further discussion and verification and testing.

Appendix C
EAC Example

Entity Description

Identity Section

Used:

Rostovzeff, Michael I. (Michael Ivanovitch), 1870-1952.

Not used:

Rostovzeff, Michael Ivanovitch
Rostovtzeff, Michael I. (Michael Ivanovitch)
Rostovtzeff, Michael (Michael)
Rostovtzeff, Michael Ivanovitch
Rostowzew, M.(Michael)
Rostowzew, Michael
Rostovtsev, Mikhail Ivanovich

Resources

Archival Records:

Creator:	Rostovzeff, Michael I. (Michael Ivanovitch), 1870-1952.
Title:	Papers of Michael Ivanovitch Rostovzeff, (1897-1968)
Extent:	Linear ft. of shelf space occupied: 4.5; Number of items: ca. 2,500
Repository:	Duke University. Special Collections Library.
Abstract:	The Michael I. Rostovzeff Papers span the years 1897 to 1968 with the bulk dated1926 to 1954. The collection chiefly consists of the correspondence of Michael Rostovzeff and C. Bradford Welles, a colleague of Rostovzeff's at Yale University, with other scholars in the fields of ancient history, archaeology, and philology. Other materials include autobiographical writings by Rostovzeff, photographs, financial papers, and clippings. The papers primarily reflect Michael Rostovzeff's tenure as a faculty member of the Classics Department at Yale University.

Description

Biographical Note

Date	Event
1870, Nov. 10	Born, Zhitomir (the Ukraine), Russia
1888	Graduated from the First Classical Gymnasium, Kiev, Russia
1892	B.A., University of St. Petersburg
1899	Master of Latin Literature, University of St. Petersburg
1901	Married Sophie M. Kulezycki
1903	Doctor of Latin Literature, University of St. Petersburg
1905-1918	Member, Constitutional Democratic Party
1916-1919	Member, Russian Academy of Sciences
1918	Emigrated to Great Britain
1918-1920	Lecturer, Queen's College,
1920-1925	Professor of Ancient History, University of Wisconsin, Madison, Wis.
1925-1939	Sterling Professor of Ancient History, Yale University
1926-1927	Published *A History of the Ancient World*
1926	Published *The Social and Economic History of the Roman Empire*
1928-1937	Director of the Yale University Expedition at Dura-Europos
1938	Published *Dura-Europos and Its Art*
1939	Appointed Director of Archaeological Studies, Yale University
1941	Published *The Social and Economic History of the Hellenistic World*
1944	Appointed as the Sterling Professor of Ancient History and Classical Archaeology, Emeritus
1952, Oct. 20	Died, New Haven, Conn.

Professor Rostovzeff received honorary degrees from the University of Leipzig (1909), Oxford University (1919), University of Wisconsin (1924), Cambridge University (1934), Harvard University (1936), Athens University (1937), and the University of Chicago (1941). He was also a member of

numerous national academies and learned societies, both in the United States and Europe. Included among these are Phi Beta Kappa, the American Academy of Arts and Sciences, the American Philological Society, La Pontificia Accademia Romana di Archeologia, Academie des Inscriptions et Belles Lettres, and the Polish Academy of Science.

Record Control Information

Record type: personal name

Editorial status: draft

Language encoding standard: ISO 639-2B

Script encoding standard: ISO 15924

Date encoding standard: ISO 8601

Country encoding standard: ISO 3166-1 a2

Owner encoding standard: ISO 11551

Record identifier: US::VaU::Example06

Maintenance history:

Name	Date	Event
Daniel Pitti	3 September 2001	Record created using <bioghist> element in rostov.xml and LCNAF record.

Language/Script of description: English in Latin Script.

Descriptive rules: Anglo-American Cataloging Rules, Second Edition.

Sources:
Guide to the Papers of Michael Ivanovitch Rostovzeff
    Michael Ivanovitch Rostovzeff, 1870-1952

Library of Congress Name Authority File, record id:
    His *Istoriia gosudarstvennago otkupa*, 1899.
    nuc89-43423: His *Iranians & Greeks in south Russia* [MI] 1922 (hdg. on MH rept.: Rostovtsev, Mikhail Ivanovich, 1870-1952; usage: M. Rostovtzeff)
    Römische Bleitesserae, 1979: t.p. (Michael Rostowzew)