# Encoded archival description (EAD) conversion: a methodological proposal

*Eduardo Peis*
*Félix de Moya and*
*J. Carlos Fernández-Molina*

**The authors**

**Eduardo Peis** is Professor, **Félix de Moya** is Dean and **J. Carlos Fernández-Molina** is Professor, all at the Facultad de Biblioteconomía y Documentación, Universidad de Granada, Granada, Spain.

**Keywords**

Archives, Standard Generalized Mark-up Language, Spain

**Abstract**

The eventual adaptation of archives to new technological possibilities could begin with the creation of digital versions of archival finding aids, which would allow the international diffusion of descriptive information. The Standard Generalized Markup Language (SGML), document type definition (DTD) for archival description known as encoded archival description (EAD) is an appropriate tool for this purpose. Presents a methodological strategy that begins with an analysis of EAD and the informational object to be marked up, allowing the semiautomatic creation of a digital version.

**Electronic access**

The current issue and full text archive of this journal is available at
**http://www.emerald-library.com**




## Introduction

Informational services should be re-examined regularly to make sure that all types of materials and electronic formats are used, and that integrated access to informational sources – even external ones – can be offered. It is not a matter of starting from square one, as Croft (1995) would say, when referring to the current lines of research in information retrieval (IR), but rather of reorienting efforts.

This need to adapt is even more crucial in the case of archives, because of their special characteristics, and the scarcity of standards, economic resources and administrative resources.

If archives adjusted to the new technological possibilities, they might be exemplary in making efficient use of resources and increasing the degree of use of the information. One key step on this road to adaptation is the design of structured finding aids that would facilitate Web diffusion of archival information, as well as the re-utilization of this material in the form of varied information objects.

The special features of archival material make it necessary to develop such systems with consideration of the following aims, to:

- present in an extensive and interrelated sense, the descriptive information normally contained in finding aids;
- preserve the hierarchical relations that exist among the levels of description;
- represent the descriptive information that is "inherited" from one level to another;
- "browse" through a hierarchical information architecture; and
- index and retrieve specific elements[1].

These needs can be satisfied by creating a digital version of the finding aids, and encoding the structured descriptive information with a mark-up language that conforms to international standards.

Library systems are beginning to use Standard Generalized Markup Language (SGML) for importing/exporting information in an electronic format, and for creating databases that allow the inclusion of information in multiple formats (Corthouts and Philips, 1996). SGML allows the logical structure of electronic documents to be represented explicitly and rigorously, in an unequivocal and independent form that is internationally recognized, regardless of applications and systems (International Organization for Standardization, 1986).

By means of marks, SGML describes the structural components of an electronic document, identifying each part according to its purpose in the given text or context[2]. Because SGML is a means to describe a language formally, it is known as a meta-language.

The structured information is "analyzed" in the sense that it is divided into component parts, which in turn have components, and so on. If the components can be identified, then they can be processed as well. At present SGML is, for many reasons, the best option for conceptually codifying a document's structural units in the broadest sense. First, it is a generalized public standard, independent of manufacturers and distributors, which ensures its permanence. Second, it does not impose a fixed set of components: the structures can be defined by the user, under the concept of "document type". Therefore, it is remarkably flexible and consistent, adapting to whatever alphabetical system may have been used to write the text. Finally, it breaks down the structure of the presentation, as it is a descriptive rather than procedural codification. For these reasons, more and more applications are using SGML as a basic tool[3].

It is easy to understand why the first attempt to apply a standardized code to instruments of archival description – initiated by the University of California at Berkeley and directed by Daniel Pitti – selected SGML as the ideal technique for carrying out such a codification.

This project designed a document type definition (DTD) to describe a class of documents which consists of an optional title page, the description of a unit of archival material, and appendixes, which are also optional. The title page could include varied elements such as the identification of the repository, or the type of descriptive tool. The descriptive component, in agreement with the DTD, would offer a brief description of the unit (using markable elements analogous to those employed in a MARC cataloging record), a broader narrative description of the unit plus any of its separate parts (including markable elements such as title, dates, scope and contents) and a formatted list of the parts that contain that unit (Gilliland-Swetland, 1996).

We can no longer conceive an "up-to-date" informational service without taking into account the World Wide Web. HTML (or its foreseeable variations) is not adequate for extracting many of the hypertext or hypermedia possibilities of the Web. With SGML these possibilities are greatly increased, but precisely for this reason, their Web implementation is costly to say the very least (and may be simply impossible). Because of this, a set of subsidiary standards is being developed under the auspices of the W3C: Extensible Markup Language, or XML[4]. It has "complementary" specifications in the form of Extensible Style Language (XSL) and Xlink (XML linking language) to enhance SGML potential while allowing the distribution of informational products over the Web.

The current EAD model (version 1.0) offers the option of using XML. It is possible to activate/de-activate the variant sections SGML/XML option using the SGML feature called "marked sections".

Despite the apparent advantages of applying this model, many archival systems can run into formidable difficulties in converting the finding aids from paper to digital form, because of having to adjust the information components of these tools to an EAD structural mode. The most obvious solution would be to key in all the data again, following the DTD EAD, but this may not be feasible in view of the – usually – enormous volume of such finding aids. The objective of our study was to confirm the viability of conversion through the development of a pilot *ad hoc* system, and to show that the process can confer an added value.

SGML conversion normally implies the construction of a bridge between the world of the printed document or word processor (in which the logical structure is visually perceived by the reader) and the "intelligent" document (in which the logical structure is explicitly codified). In effect, SGML encodes the logical structure; yet in most documents this structure is expressed visually, not by means of well defined marks. When converting to SGML, the implicit structure must be made explicit (Severson, 1995).

Library Hi Tech
Volume 18 · Number 4 · 2000 · 360–368

Therefore, in general terms, our proposal consists of an analysis of the SGML model to which conversion is intended, focusing on the DTD EAD and the documents that are going to be converted (finding aids). This approach allows the process to be arranged as a production line. As we shall see, nearly all this line can be automated with an appropriate strategy and proper software tools.

## Methodology

If we bear in mind the variety of situations involving archival informational services, the advantages of designing a modular method become apparent. Our approach involves three separate stages:

(1) previous analysis;
(2) capture/conversion; and
(3) translation/organization.

The first methodological stage may be the most important one. The initial analysis must contemplate both the DTD EAD and the descriptive instrument *per se*. The purpose of these preliminary analyses is threefold, to:

(1) corroborate the adequacy of the SGML model to the descriptive instrument that is to be labeled;
(2) study the possible structural correspondence to find those informational elements that are not presented in the "traditional" finding aid model and will have to be added "manually"; and
(3) discover the physical and logical characteristics of the format of the "source" document that will automatically allow SGML labeling.

These initial assessments will also enable us to make important methodological decisions with respect to the possibilities of application of the model, such as the type of description (in depth, analytical or combined) that best reflects the scope of the information to be converted.

We must be familiar with EAD to be certain that it works for the task at hand. The DTD EAD (Pitti, 1999) was designed to reflect the inherent hierarchy of the collections, in conjunction with the intellectual organization that archivists impose with their descriptive practices. Archival description is regulated by the International Standard Archival Description-General (ISAD-G). The ISAD-G (International Council on Archives, 1994), developed by the International Council on Archives, can be considered the semantic and syntactic intellectual foundation of archival description. EAD, then, would be the standardized communication format, the formal part of that intellectual structure. From this standpoint, EAD is based on ISAD-G.

The EAD model contains two types of elements:

(1) those that encode specific points in the description of component parts of either the finding aids or the original materials (i.e. the descriptive elements); and
(2) those that could encode any characteristic of the document (i.e. the generic elements).

The latter are, however, generally included among the descriptive elements.

EAD uses the term "finding aid" to refer to any hierarchical tool that has been codified using EAD, and that allows either the record-creator or the user to access the materials that are being described.

At a very basic level, a finding aid that is codified using EAD will have three segments:

(1) one that provides information about the finding aid itself, such as its title, the name of its compiler, and date of compilation (<eadheader>);
(2) a second component that includes the preliminary matters necessary for the formal publication of the description (<frontmatter>); and
(3) a third that provides the description of the archival material itself, in addition to related contextual and administrative information (<findaid>).

The <eadheader>, which is based on the "header" element of the SGML Text Encoding Initiative (TEI) model[5], uses four sub-elements – <filedesc>, <profiledesc>, <revisiondesc> and <requiredft> – to capture or contain most of the information that is normally recorded about the creation, publication and use of a "finding aid document". Additional information about this document that does not fit the TEI model can be included in the <frontmatter> element, which in turn includes the <titlepage> and

Library Hi Tech
Volume 18 · Number 4 · 2000 · 360–368

<div> elements and reflects details of an introductory sort (e.g. preface or introduction) that are needed for the formal publication of the finding aid.

In the <findaid> element, two types of information may be presented under two main sub-elements. The  (“archival description”) element, which can be considered as the fundamental component of EAD, presents hierarchically organized information that describes a unit of records (or papers) together with their component parts (or divisions). In addition, <archdesc> has several “high level” descriptive categories that serve as “containers” for more specific descriptive components.

The most important of the high level elements is <did>, or descriptive identification, which will contain those elements needed to identify the title, creator, date of creation, and volume. The element “additional descriptive data”, or <add>, contains supplementary, optional information that does not directly describe the records, yet facilitates their use by researchers (for example, a bibliography). <Archsdesc> also contains an element that provides a detailed analysis of the collection materials: <dsc> or description of subordinate components. The <dsc> element has a repeatable element, <c> (component), which may include all the descriptive elements that appear in <archdesc>. In this way, <c> can be subdivided into other <c>’s until arriving at the simplest elements that make up the record pool.

Whatever attribute or value is used to describe material on one level is automatically “inherited” by the elements at subordinated levels, unless the EAD specifies otherwise.

The attributes reflect defined properties of an element, and can take on different values, depending on the context in which they appear. In order to set up one or more attributes, a codifier should include the name of the attribute (using the same angle-brackets “<>” that are used to define mark-up labels) and its value.

The point of this first phase, in general terms, is that all the information contained in a traditional finding aid can be structured using the elements that constitute the EAD.

For archives with electronic finding aids, the process of converting to EAD is not terribly time-consuming, though it is somewhat complex. The macro functions of text processors or design-conversion scripts may be used to relate the content of the elements pertaining to the database records with the corresponding elements of the SGML model, by inserting the proper labels. Those finding aids that do not exist in an electronic format must be previously converted, either by recreating the finding aid or by using optical character recognition (OCR) technology.

In either digital or printed form the inconsistencies of formatting, or even individual descriptive idiosyncrasies, may complicate the conversion process. For this reason, there is an essential second stage: the formal analysis of the information object to be converted. The application of uniform formatting characteristics (including tabulation, spaces, and regular columns) leads to the automatic identification of the different information elements, thereby permitting them to be labeled.

The next step is the process of capture and conversion. Capture by scanning implies OCR processing. In this scenario, the inconsistencies of formatting we mentioned earlier are further complicated by the omnipresent OCR errors. This calls for double OCR processing: first, for the elimination of errors and proper identification of the formatting characteristics that will discriminate the structural elements, and secondly, for the proper configuration of the OCR software (using, for example, the capacity for manually creating “zones”). In this way, those features of the format that will be crucial for their later conversion can be saved.

Our proposal treats this second methodological phase as an intermediate stage, since it includes the use of a non-proprietary application for SGML conversion, such as Rainbow Maker[6]. This program uses information on a paragraph or character basis, or even on the basis of common attributes that are not style-dependent, in order to insert labels (in this case Rainbow labels) to set off and indicate the relevant chains of text. Although Rainbow is itself an SGML format (a DTD, to be more specific), it is not appropriate for the permanent representation of data, because it

Library Hi Tech
Volume 18 · Number 4 · 2000 · 360–368

only reflects that structure or content identification that is given in the original document.

Engaging in this intermediate step instead of directly using a programming aid for conversion may be more time-consuming in the short term. However, in the mid-term, it provides dual benefits. On the one hand, we now have an intermediary SGML file that can be used, after applying conversion routines, to generate multiple informational objects; and on the other hand, its use of common characteristics will simplify the production of specifications applicable to more than one finding aid, thereby enhancing the automation process. Although using a database with Rainbow files requires more space for memory, it affords the benefit of "pivot" files that allow the direct unfiltered conversion of SGML to XML, or the placing of already processed descriptive information on to an ISAD-G style-sheet. Moreover, indexing, organizational and selection routines may be re-utilized, requiring only slight adaptations.

The final phase consists of designing scripts for translating Rainbow to EAD and, finally, organizing the SGML labels. The end step allows us to connect all the informational components necessary for obtaining the complete archival product.

## Results

We chose a printed inventory to test our proposal: the stock of the Archives of the Infante don Gabriel de Borbon (Mut, 1985), a work of high quality printing, relatively simple description, and very manageable volume.

The process is semi-automatic. It requires the manual introduction of any information needed for the SGML document that is not found in the source document. This involves using two files, one that will be considered an internal entity (the file designated as "plantilla") and another that is an external entity or system entity (the "maestro" file). For example, the explanations of system (content-type, content-ID, and content description), the EAD clarification (in this case, DTD version 1.0), and the entity references are made from the "plantilla" file to the "maestro" system file using the SGML aid "system entity reference". The "plantilla" file should also hold the information contained in the elements <eadheader> and <frontmatter>. Also, before processing, it is necessary to configure the attributes in order to match the EAD declarations.

In applying attributes, we naturally take advantage of the so-called "default" values. Default values are automatically given by the system if the encoder of the finding aid does not specify an alternative value. For example, within the <ead> element, which indicates to the computer that the document is encoded in EAD, there is an attribute called audience. It indicates whether the contents of the encoded document are of free public access, or private and restricted. If the coder does not specify "private", the system automatically interprets the attribute as public. A different example would be the manual manipulation of an attribute such as type, which can affect elements like <archdesc> or <dsc>. For instance, <dsc> may be configured to have a value "analyticover", indicating to the processor that the description of component parts is presented in the form of a summary.

For our trial, notes were defined on the basis of their location. All the elements that could be used in a great number of other elements were defined generically, but their state (that is, "required/not required", "repeatable/not repeatable") was determined by the element in which each was included.

This configuration of information was also included in the "maestro" file. The use of a master file makes processing easier by saving space and increasing the automatic component of the process, as redundancies are avoided. The other file, "plantilla", features an identifier (ID) that was used as a pointer to connect the information with the product of conversion process. This procedure facilitates handling the files.

Part of the content of the "maestro" file of our trial can be seen in Figure 1.

The capture and OCR processing of the paper finding aid comprises a series of files. Each file contains text in Rich Text Format (RTF) and holds information about the existence of fonts (bold and cursive), as well as a precise representation of the page layout

Library Hi Tech
Volume 18 · Number 4 · 2000 · 360–368

**Figure 1** Sample of Maestro file

```
Content-Type: TEXT/PLAIN; charset=ASCII; name="arch.bor.sgml"
Content-ID: Veleta.UNX.284.3.82.03022564/eubd1.ugr.es
Content-Description:
<!-- Fichero ead.dtd: DTD para instrumentos de descripción -->
<!DOCTYPE EAD PUBLIC "-//Society of American Archivists//DTD ead.dtd
(Encoded Archival Description (EAD))//EN" system "maestro".ead.dtd"
[<!ENTITY hdr system "maestro".header.hdr">
]>
<!-- siguen la declaración de tipo de documento y las declaraciones de
entidad -->
```

(including tabs and blank lines). Figure 2 offers an example.

Rainbow Makers uses the information regarding the format and logic of the processed documents contained in these files to insert the Rainbow DTD labels, and the product is then placed in a single file. The part of the Rainbow file corresponding to the above example, then, is shown in Figure 3. Some spaces and tabs were introduced to make the sample easier to read.

Using language aids from the Practical Extraction and Report Language programming (perl)[7], we proceeded to translate the Rainbow labels to EAD labels in "object by object" fashion, principally using character level format (CLF) and location information (WPLOC). After that, these objects were organized according to the document type. The perl aid was used to match chains against a "pattern" document. Once saved, the resulting file is connected to the essential data contained in the "plantilla" file, again producing a single file. The EAD document model corresponding to the proposed example, in an abbreviated form, is given in Figure 4.

**Figure 2** Sample of the printed inventory

**Secretaría**

| | | |
|---|---|---|
| **Leg. 17** | Tenientes de Gobernadores de Alcázar y **Consuegra**; incidencias sobre si ha de haber uno o Dos. | **1738-1817** |
| **Leg. 18** | Alguacil Mayor. Alguaciles ordinarios de la Gobernación. Alcaldes de feria. | **1602-1820** |
| **Leg. 19** | Insaculaciones e instrucciones sobre el modo de Hacerlas. Gobierno de la villa de Alcázar y sus Anejas en jurisdicción. Competencias con los Alcaldes ordinarios y asuntos controvertidos de su juzgado. | **1632-1795** |

This document instance, or any other obtained through the same procedure, can be validated against its DTD using an SGML/XML parser, in our case James Clark's "XP"[8].

## Conclusions

The use of SGML is presented as a reasonable choice for information services, because it allows proper electronic processing, and is not overly complicated. It is well suited to the rigorous logical structure of bibliographic descriptions, in addition to text. It also allows flexibility in the inclusion of both MARC and non-MARC descriptive elements, and permits the interconnection of multiple information objects. Moreover, SGML addresses the problem of special characters through the technique of entity references, and it controls the quality of the input through the features of the parser. On the other hand, the separation of the structure of the document from its particular style of representation makes it possible to use SGML documents for a variety of purposes (e.g. in a Web environment, or in CD-ROM databases), and with different end-user formats (e.g. HTML, private formats, or markup languages specifically oriented to printed copies, such as RTF, LaTeX, or Post-Script).

SGML encoding of archival descriptive elements – accessible either locally or online – simplifies, improves and expands access to archival collections, making it possible to connect catalog records with finding aids. It also allows keyword searches in connected sets of finding aids, thus enabling the retrieval of information that would otherwise remain hidden.

EAD constitutes an ideal DTD SGML for representing the structure of the archival description elements, since it not only describes the physical and intellectual constituents of the documents, but also preserves the hierarchical relationships between these elements. This permits "browsing" over levels of description, without duplicating information. The fact that it fully complies with international standards for artificial description, such as the International

**Figure 3** After Rainbow DTD processing

```
<!DOCTYPE rainbow PUBLIC "-//EBT//DTD Rainbow 2.5//EN"
[]><RAINBOW><FILEINFO ORIGIN="WinWord-RTF1"
DTDVER="2.5"><STYINFO><PARATYPE FONT-FAMILY="Times New Roman" FONT-
SIZE="10" LINE-SPACING="12" CHARSET="ISO-8859" JUSTIFICATION="LeftJust"
FIRST-INDENT="0" LEFT-INDENT="0" RIGHT-INDENT="0" SPACE-BEFORE="0"
SPACE-AFTER="0" FONT-WEIGHT="Medium" FONT-SLANT="Roman"
NAME="Normal"><PARATYPE FONT-FAMILY="Times New Roman" FONT-SIZE="14"
LINE-SPACING="16" CHARSET="ISO-8859" JUSTIFICATION="LeftJust" FIRST-
INDENT="0" LEFT-INDENT="0" RIGHT-INDENT="0" SPACE-BEFORE="0" SPACE-
AFTER="0" KEEP-WITH-NEXT=1 FONT-WEIGHT="Bold" FONT-SLANT="Roman"
NAME="heading 1"><PARATYPE FONT-FAMILY="Times New Roman" FONT-SIZE="12"
LINE-SPACING="14" CHARSET="ISO-8859" JUSTIFICATION="LeftJust" FIRST-
INDENT="0" LEFT-INDENT="0" RIGHT-INDENT="0" SPACE-BEFORE="0" SPACE-
AFTER="0" KEEP-WITH-NEXT=1 FONT-WEIGHT="Bold" FONT-SLANT="Roman"
NAME="heading 2"><PARATYPE FONT-FAMILY="Times New Roman" FONT-SIZE="12"
LINE-SPACING="14" CHARSET="ISO-8859" JUSTIFICATION="LeftJust" FIRST-
INDENT="0" LEFT-INDENT="0" RIGHT-INDENT="0" SPACE-BEFORE="0" SPACE-
AFTER="0" KEEP-WITH-NEXT=1 FONT-WEIGHT="Medium" FONT-SLANT="Roman"
NAME="heading
3"><CLFTYPE NAME="Default Paragraph Font"></STYINFO>
<DOC><HEAD><WPLOC wp-addr="1"><PARA KEEP-WITH-NEXT=1 FONT-WEIGHT="Bold"
FONT-SIZE="14" LINE-SPACING="16" PARATYPE="heading
1"><PARACONT>Secretar&iacute;a</PARACONT></PARA><STRUCLVL>
<WPLOC wp-addr="2"><PARA KEEP-WITH-NEXT=1 FONT-WEIGHT="Bold" FONT-
SIZE="14" LINE-SPACING="16" PARATYPE="heading 1"
><PARACONT></PARACONT></PARA><WPLOC wp-addr="3"><PARA FONT-SIZE="10"
LINE-SPACING="12" PARATYPE="Normal"><PARACONT></PARACONT></PARA>
<WPLOC wp-addr="4">
<PARA KEEP-WITH-NEXT=1 FONT-WEIGHT="Bold" PARATYPE="heading 2"
><PARACONT>Leg. 17<tab></PARACONT></PARA>
<STRUCLVL><WPLOC wp-addr="5">
<PARA KEEP-WITH-NEXT=1 FONT-WEIGHT="Bold" PARATYPE="heading 2"
><PARACONT><CLF FONT-WEIGHT="Medium" >Tenientes de Gobernadores de
Alc&aacute;zar y Consuegra; incidencias sobre si ha de haber uno o
Dos.<tab><CLF FONT-WEIGHT="Bold" >1738-1817</CLF></PARACONT></PARA>
<WPLOC wp-addr="6">
<PARA FONT-SIZE="10" LINE-SPACING="12" PARATYPE="Normal"
><PARACONT><CLF FONT-WEIGHT="Bold" >Leg. 18<tab></CLF>Alguacil Mayor.
Alguaciles ordinarios de la Gobernaci&oacute;n. Alcaldes de
feria.<tab><CLF FONT-WEIGHT="Bold" >1602-1820</CLF></PARACONT></PARA>
<WPLOC wp-addr="7">
<PARA FONT-SIZE="10" LINE-SPACING="12" PARATYPE="Normal" ><PARACONT><CLF FONT-
WEIGHT="Bold" >Leg. 19</CLF><tab>Insaculaciones e instrucciones sobre el modo de
Hacerlas. Gobierno de la villa de Alc&aacute;zar y sus Anejas en jurisdicci&oacute;n.
Competencias con los Alcaldes ordinarios y asuntos controvertidos de su juzgado.<tab><CLF
FONT-WEIGHT="Bold" >1632-
1795</CLF></PARACONT></PARA></STRUCLVL></STRUCLVL></DOC></RAINBOW>
```

Standard Archival Description (ISAD) and other standards and formats, such as MARC, serves as a guarantee of its versatility.

On the basis of our results, we can affirm that it is possible to develop a method that allows a semi-automatic conversion of paper finding aids into EAD, using technology now in wide use. The use of SGML as a key part of the proposed method allows the use of intermediary products of the conversion process, thus maximizing the yield of the given resources. The simplicity of the SGML file and the versatility of the possible indexing routines facilitate the use of advanced retrieval methods.

Even though SGML is highly complex, it may be transformed into XML. One strategy

Library Hi Tech
Volume 18 · Number 4 · 2000 · 360–368

**Figure 4** Corresponding EAD document

```
<ead>
<eadheader> <eadid>011</eadid>
<filedesc><titlestmt><titleproper>Versión EAD del inventario del
Archivo del Infante Don Gabriel de Borbón</titleproper></titlestmt>
</filedesc></eadheader>
<frontmatter><titlepage><author>Antonio Mut Calafell</author><author>
<persname role="colaborador">Jose Luis de la Peña García</persname>
</author><titleproper>Inventario del Archivo del Infante Don Gabriel de
Borbón<lb>Gran Prior de la Orden de San Juan de Jerusalén<lb>en los
Reinos de Castilla y León, y de sus descendientes</titleproper>
<publisher>Archivo del Palacio Real de Madrid</publisher>
<sponsor>Ministerio de Cultura<lb>Dirección General de Bellas Artes y
Archivos</sponsor><date>1985</date></titlepage>
</frontmatter>
<findaid>
     <archdesc audience="public" level="fond">
      <did><unittitle>Archivo del Infante Don Gabriel de
      Borbón<unitdate>1710-1842</unitdate></unittitle><physdesc>
      <extent>3 metros lineales</extent></physdesc>
     </did>
  <admininfo><custodhist><p>Entrega del archivo a la dirección general
de rentas y arbítrios de amortización el 22 de agosto de
1839</p></custodhist><accessrestrict><p>Acceso libre</p>
</accessrestrict></admininfo>

<! -- ... -- >

   <dsc audience="public" type="analyticover">
            <head>Inventario de series</head><thead><row valign="top">
     <entry colname="1">Legajos</entry><entry
     colname="2">Asunto</entry><entry colname="3">Fechas
     extremas</entry></row></thead>

<! -- ... -- >

   <c01 level="series"><did><unititle>Secretaría</unititle></did></c01>
       <c02 level="item"><did><container>Leg.
       17</container><scopecontent>Tenientes de Gobernadores de
       Alc&aacute;zar y Consuegra; incidencias sobre si ha de haber uno o
       Dos.</scopecontent> <unitdate>1738-1817</unitdate></did></c02>
       <c02 level="item"><did><container>Leg.18</container><scopecontent>
       Alguacil Mayor. Alguaciles ordinarios de la Gobernaci&oacute;n.
       Alcaldes de feria.</scopecontent><unitdate>1602-
       1820</unitdate></did></c02>
       <c02 level="item"><did><container>Leg.
       19</container><scopecontent> Insaculaciones e instrucciones sobre
       el modo de Hacerlas. Gobierno de la villa de Alc&aacute;zar y sus
       Anejas en jurisdicci&oacute;n. Competencias con los Alcaldes
       ordinarios y asuntos controvertidos de su
       juzgado.</scopecontent><unitdate>1632-1795</unitdate></did></c02>
</dsc></archdesc></findaid></ead>
```

for making optimal use of archival resources is to have multiple databases with a variety of information objects (e.g. descriptive information, contextual information, or digital images). The information could be almost instantly converted, putting multiple presentation options at our disposal via the XSL specification, including an effective means for the creation of hyperlinks using the XML Link specification.

Library Hi Tech
Volume 18 · Number 4 · 2000 · 360–368

## Notes

1 Development of the Encoded Archival Description Document Type Definition, available: http://www.loc.gov/ead/eadback.html
2 Sperberg-Mc Queen, C.M. and Burnard, L., *A Gentle Introduction to SGML*, available: http://www-tei.uic.edu/orgs/tei/sgml/teip3sg/SG.htm
3 Cover, R., *The SGML/XML Web Page*, available: http://www.oasis-public.com/cover
4 W3Consortium, XML WP, available: http://www.w3.org/xml
5 *TEI Guidelines for Electronic Text Encoding and Interchange (P3)*, available: http://etext.virginia.edu/TEI.html
6 Sklar, D., "Accelerating conversion to SGML via the Rainbow format", *Electronic Book Technologies EBT*, available: http://ftp.sunet.se/pub/text-processing/sgml/Rainbow/Rainbow.why
7 Practical Extraction and Report Language (perl), available: http://www.bme.unc.edu/facilities/software/perl/perlIndex.html
8 XP – James Clark's XML parser in Java. Version 0.4 available: http://www.jclark.com/xml/xp/index.html

## References

Corthouts, J. and Philips, R. (1996), "SGML: a librarian's perception", *The Electronic Library*, Vol. 14 No. 2, pp. 101-10.

Croft, W.B. (1995), "What do people want from information retrieval?: (The top 10 research issues for companies that use and sell IR systems)", *D-Lib Magazine*, Vol. 1 No. 6, available: http://www.dlib.org/dlib/november95/11croft.html

Gilliland-Swetland, A.J. (1996), *Encoded Archival Description Document Type Definition (DTD). Applications Guidelines Technical Document No. 1*, Library of Congress and Society of American Archivist, Washington, DC, available: http://www.loc.gov/ead/

International Council on Archives (1994), *ISAD(G): General International Standard Archival Description*, Adopted by the Ad Hoc Commission on Descriptive Standards, Stockholm, Sweden, 21-23 January 1993, International Council on Archives, Ottawa, Ontario, available: http://www.archives.ca/ica/cds/isad(g)e.wp

International Organization for Standardization (1986), *ISO 8879-1986 (E), Information Processing – Text and Office Systems – Standard Generalized Markup Language (SGML)*, International Organization for Standardization, Geneva.

Mut, A. (1985), *Inventario del archivo del Infante don Gabriel de Borbón*, Dirección General de Bellas Artes y Archivos, Ministerio de Cultura, Madrid, p. 49.

Pitti, D.V. (1999), "Encoded archival description. An introduction and overview", *D-Lib Magazine*, Vol. 5 No. 11, available: http://www.dlib.org/dlib/november99/11pitti.html

Severson, E. (1995), "The art of SGML conversion: eating your vegetables and enjoying dessert", available: http://www.oasis-open.org/cover/seversonVeg.html (mirror copy).