

XML y registros electrónicos: principales estándares en la descripción archivística

Rogério Müller Fernandes

Doutorando em ciência da computação na Universidad de Alcalá,
Espanha.
E-mail: romuller@uel.br

Resumen

Describe los principales estándares XML en la descripción archivística concomitantemente con la definición y presentación de las herramientas de XML y conceptos del campo archivístico.

Palabras clave

XML. Registros electrónicos. Estándares. Descripción archivística.

XML and electronic records: main standards in the archival description

Summary

It describes the main standards XML in the archival description concomitantly with the definition and presentation of the tools of XML and concepts of the archival field.

Keywords

XML. Electronic records. Standards. Archival description.

INTRODUCCIÓN

Desde cuando los hombres comenzaron a emplear símbolos para registrar los sonidos de las palabras, más que para representar objetos o pensamientos, el conocimiento viene siendo registrado y acumulado, utilizándose la tecnología de producción existente, en una evolución que va desde las placas de arcilla en la Mesopotámica, hasta los registros electrónicos de hoy.

A pesar de las muchas dudas aún existentes relacionadas a aspectos legales, preservación digital e inherentes migraciones, entre otras, cada vez más informaciones y documentos son creados y/o convertidos para la realidad digital.

Para describir los diferentes tipos de contenido de forma electrónica, para que se pueda los almacenar, procesar y transmitir, es necesario disponer de formatos que puedan representar los datos que circularán por la red [1], tarea y definición del XML (eXtensible Markup Language), que se hizo el metalenguaje padrón de las aplicaciones informáticas.

En el área archivística los instrumentos de descripción son las principales herramientas de apoyo al efectivo uso de los archivos, y para que puedan ser utilizados de manera integrada y compartida, posibilitando la construcción de sistemas globales unificados de información en la web es necesario que sean utilizados de acuerdo con normas o estándares ampliamente aceptos.

En el transcurrir del texto se pretende contextualizar el uso del XML en la descripción archivística, trazando un panorama de los principales estándares en la gestión de los registros electrónicos. Para ello, además de esa primera sección introductoria, el contenido está estructurado en cinco partes más: en la segunda son presentadas las definiciones y demás informaciones relevantes sobre XML, en la tercera son detallados algunos de los principales conceptos archivísticos y sus relaciones con el XML, en la cuarta sección son descritos los principales estándares, en la quinta las iniciativas de nuevas herramientas en el campo archivístico y en la última la conclusión.

XML

XML o Lenguaje Extensible de Marcado es un metalenguaje (lenguaje para definir otros lenguajes), desarrollado y mantenido pelo World Wide Web Consortium (W3C), – consorcio internacional fundado para desarrollar protocolos y líneas directivas visando la evolución de la web – que describe una clase de objetos de datos llamada documentos XML y parcialmente describe el comportamiento de los programas que los procesan [1]. Su desarrollo empezó en 1996 y es una recomendación W3C desde febrero de 1998.

Es un formato de texto simple y muy flexible, diseñado para hacer frente a los desafíos de la publicación electrónica de gran escala y que está también desempeñando un papel cada vez más importante en el intercambio de una gran variedad de datos en la web y en otras partes.

XML es también una forma restringida de SGML (Standard Generalized Markup Language), por eso, por construcción, documentos XML son conformes con los documentos SGML. Ha sido diseñado para facilitar la implementación y interoperabilidad con SGML y HTML, siendo su objetivo posibilitar un SGML genérico que pueda ser enviado, recibido y procesado en la web de la manera de como ahora es posible con HTML, pero sin la complejidad del SGML y limitaciones del HTML.

SGML o Lenguaje Estandarizado de Marcado General, a vez, es un metalenguaje usado para definir lenguajes de marcado específicos para codificar tipos particulares de textos electrónicos. Es reconocido como un estándar internacional (ISO 8879, 1986) y suministra códigos no propietarios que son independientes de cualquier software específico o plataforma de hardware. [1]

Fue definido en el inicio de los años 80 por iniciativa de las editoriales de los EE.UU que querían un *lenguaje en que se pudiesen separar las dos funciones principales de relevancia en el mundo editorial, que son los contenidos y la forma de presentar esos contenidos, en este caso los libros, las publicaciones* [2]. Cuando se desarrollo era un lenguaje independiente de las redes y de Internet, sin embargo, fue la base en la definición del HTML (HyperText Markup Language).

Guiados por la experiencia acumulada con HTML, los proyectistas de XML sencillamente tomaran lo mejor de SGML y producirán algo que es tan potente como SGML pero más regular e sencillo de usar [3].

El éxito de XML, sin embargo, no proviene del objetivo que esperaban sus diseñadores; no viene porque se haya separado forma y contenido, sino por una razón en la que no habían pensado, y es que los datos tienen que circular por la red. Para ello la mejor forma es estructurar esos datos de forma jerárquica y para representar información jerárquica, XML es un formalismo perfectamente adaptado [2].

Estructura del XML

Un lenguaje de marcado es un lenguaje que permite añadir marcas o etiquetas que describen el contenido del documento para dar semántica o indicar la presentación del contenido, que habitualmente, es textual [2]. En otras palabras, los documentos escritos en un lenguaje de marcado contienen básicamente dos cosas: datos y metadatos, siendo los metadatos la información extra que añade un contexto o significado a los datos [4]. El XML, diferente del HTML que tiene marcado predefinido, posibilita al autor del documento proyectar su propio marcado.

Además de los caracteres especiales (“, “, <, >, /) que típicamente establecen los límites del marcado, separando datos de estructura, un documento XML es formado por dos componentes básicos, elementos y atributos:

- *Elementos*: son la base de la estructura de los documentos, denomina al conjunto formado por un cierto marcado y el contenido que delimita.

```
<elemento>contenido</elemento>
```

- *Atributos*: son componentes que se insertan en el interior de los elementos para añadir información adicional a los mismos.

```
<elemento atributo="valor">
  contenido</elemento>
```

La Figura 1 presenta un texto y su correspondiente en XML, y como puede observarse, el XML representa los datos de forma “inteligente” preocupándose con el significado, diferente del HTML, por ejemplo, que se preocupa con el aspecto de los datos.

DTD y XML Schemas

Para que se pueda decir que un documento es XML lo mismo debe cumplir con las normas sintácticas generales de XML (ser bien formado) y para que pueda ser utilizado para el intercambio de información entre sistemas debe cumplir con la declaración de elementos y atributos realizada en una DTD (ser válido).

Una DTD (Declaración de Tipo de Documentos) es el conjunto de definiciones estructurales creadas para validar un documento XML contra ellas. En dicha DTD se realiza la declaración formal de la gramática en la que se definen los elementos, atributos, entidades y reglas que se deberán utilizar en el documento XML que aspire ser conforme a ella [6].

La DTD es un archivo que contiene una descripción del lenguaje del XML que describe lo que es posible en la su lenguaje XML y lo que no lo es. Así, cuando se habla del “lenguaje XML”, se está hablando en realidad, de una DTD específica. *Una última gran característica del XML, es la habilidad de utilizar varias DTDs al mismo tiempo, en el mismo documento. Esto significa que puede utilizar varios tipos de datos diferentes, al mismo tiempo, en el mismo documento* [4].

DTDs son frecuentemente utilizadas por comunidades especializadas que comparten intereses comunes más allá de la frontera de organizaciones particulares o localidades, por eso, los creadores de herramientas de descripción no tienen que crear sus propias DTDs porque pueden utilizar una que ya existe y es reconocida por los otros miembros de la comunidad archivística.

Entre otras cosas, una DTD define [7]:

- qué elementos pueden aparecer en un documento
- qué elementos son obligatorios y cuáles son opcionales
- qué elementos pueden aparecer más de una vez en el mismo documento
- qué atributos pueden usarse con cada elemento
- qué atributos son obligatorios y cuáles son opcionales
- donde puede aparecer cada elemento en el documento.

Como una alternativa al uso de una DTD, los elementos y atributos pueden ser definidos usando un XML Schema. Schemas son potencialmente más poderosos y

proporcionan un mayor control sobre la información do que las DTDs, pero tienen una sintaxis más compleja.

Tecnologías XML

El XML tiene aún diversas tecnologías en su entorno que fueran desarrolladas para complementar o extender las posibilidades del estándar. La Figura 2 presenta un resumen de algunas de las principales [6]:

FIGURA 1
Ejemplo de texto y su correspondiente en XML (basado en [5])

<pre><?xml versión="1.0"/> <catálogo de dirección> <entrada> <nombre>Pablo García</nombre> <dirección> <calle>Mayor</calle> <ciudad>Alcalá de Henares</ciudad> <provincia>Madrid</provincia> <código postal>28805</código postal> <país>España</país> <dirección> <télefono principal="true">9111728234</télefono> <télefono>9164332214</télefono> <email>pablogarcia@net.es</email> </dirección> </entrada> <entrada> <nombre>Juan Sanches</nombre> <email>juansan@net.es</email> </entrada> </catálogo de dirección></pre>	<p>Catálogo de dirección Pablo García Calle Mayor, 135 Alcalá de Henares, Madrid 28805 España 9111728234 (principal) 9164332214 pablogarcia@net.es Juan Sanches juansan@net.es</p>
--	--

FIGURA 2
Tecnologías XML

Tecnología	Descripción/Función
XHTML	Permite trabajar con documentos HTML en aplicaciones que hacen uso del estándar XML
XPath	Es un lenguaje que se utiliza para acceder al contenido de un documento XML
Xlink	Permite introducir enlaces en documentos XML
XPointer	Derivada de XPath, permite cargar o seleccionar únicamente aquella parte del documento que le interese al usuario
XSL	Permite dar una apariencia en el navegador al contenido de un documento XML
DOM y SAX	Permiten realizar programas en diversos lenguajes de programación que sean capaces de utilizar XML.

CONCEPTOS ARCHIVÍSTICOS

Un documento de archivo es la información registrada, independientemente de forma o soporte, producida o recibida y mantenida por una institución o persona en el ejercicio de sus actividades públicas o privadas. El conjunto de esos registros que fueron identificados formalmente como teniendo valor a largo plazo para su creador, para una organización o para la sociedad en general, forma el archivo. [8]

Para apoyar el uso efectivo de archivos son necesarios instrumentos de descripción (finding-aids) que indiquen que los archivos existen y como fueron creados, como los archivos son organizados y como ellos pueden ser accedidos. Además de describir la naturaleza, contenido y origen de los materiales archivados, un típico instrumento de búsqueda provee información sobre la fecha de su creación o acumulación, su extensión, cantidad o tamaño, y en que condiciones pueden ser usados [7].

De ahí que un instrumento de descripción, que en la definición de la ISAD (International Standard Archival Description) es el término más amplio que comprende cualquier descripción o medio de referencia elaborado o recibido por un servicio de archivo, con vistas al control administrativo o intelectual del acervo archivístico, necesita llevar considerablemente más información contextual que, por ejemplo, un catálogo de biblioteca [8].

Los instrumentos de descripción también proveen información sobre la estructura de los archivos que ellos describen. Información estructural es necesaria porque los archivos pueden ser vistos bajo diferentes niveles. A nivel de ítem, hay normalmente numerosos documentos individuales (por ejemplo, cartas, memorandos, facturas). Cada ítem tiene una estructura interna, mientras ítems relacionados son frecuentemente agregados en archivos o carpetas. Estos por su parte, son normalmente agregados para formar series de archivos. Una serie es compuesta de archivos y/o ítems que tienen algo en común, normalmente porque ellos fueron creados en el curso de actividades relacionadas. Cuando los varios ítems, archivos y series creados o acumulados por una única organización, familia o individuo son vistos como un todo, la agregación resultante de los archivos es conocida por los archiveros como un fondo. [7]

Ese modelo jerárquico (ver Figura 3) en forma de pirámide [8], donde un único fondo contiene varias series, y cada

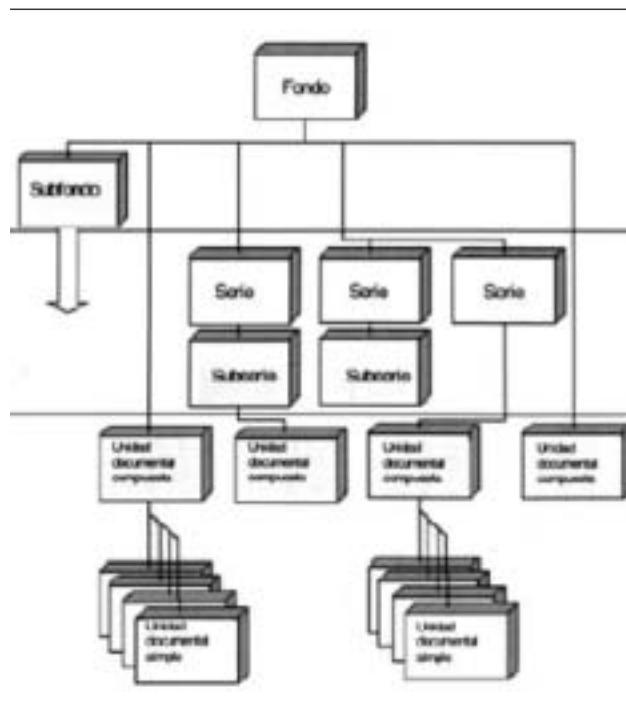
serie por su parte puede contener muchos archivos o ítems, es utilizado por la mayoría de los instrumentos de búsqueda de archivos y es normalizado por la Descripción Bibliográfica Internacional Normalizada General - ISAD(G), estándar internacional publicada por el Consejo Internacional en Archivos.

Como el XML fue proyectado para apoyar jerárquicamente la información estructurada puede representar fácilmente el tipo de construcción empleada en la mayoría de los instrumentos de descripción, que cuando codificados empleando XML, normalmente utilizan EAD (Encoded Archival Description) para propiciar un padrón común.

En cuanto la ISAD(G) constituye una norma de estructura de datos muy general y básica para las descripciones de documentos de archivos la EAD proporciona una estructura de datos precisa para los instrumentos de descripciones, que incluyen representaciones ricas e detalladas de las unidades de descripción.

Además de ISAD(G) y EAD la comunidad archivística tiene aún un otro estándar internacional con la respectiva representación también codificada en XML, la International Standard Archival Authority Record for

FIGURA 3
Modelo de los niveles de organización de un fondo



Corporate Bodies, Persons, and familias - ISAAR(CPF)2 y el EAC (Encoded Archival Context).

El principal objetivo de la ISAAR(CPF) es proporcionar reglas generales para la normalización de las representaciones de los productores y el contexto de creación de los documentos de archivo, pero puede ser utilizada también para controlar la forma de cualquier nombre de institución, persona o familia que constituya un punto de acceso a una descripción. [9]

La Figura 4 presenta el ámbito de aplicación de los actuales estándares internacionales de descripción archivística [10]. Donde las diferentes y esenciales representaciones interrelacionadas de un sistema de información archivística (descripciones de documentos de archivos, autoridades de contexto, autoridades de contenido y entidades externas) son representadas gráficamente en función de las normas y estándares internacionales de descripción archivísticas. Las clases A, B, y C y se refieren a los diferentes niveles de descripciones establecidas pelo Comité de Normas de Descripción del CIA (Internacional Council on Archives), a saber: A (al nivel solamente de fondos y subfondos), B (descripción de lo material de archivo de todos los niveles abajo al nivel del archivo), C1 (descripciones de los ítems tratados como el último nivel de descripción de un fondo) y C2 (descripciones de los ítems tratados como simple ítems sin el contexto jerárquico de que son parte).

En un enfoque más general en relación a la comprensión del papel y naturaleza de las normas descriptivas, se puede diferenciar dos tipos particulares de normas: normas de estructura de datos y normas de contenidos de datos. Las normas de estructura identifican los elementos de información requeridos para la descripción archivística. Por ejemplo, un elemento del Título que nombra la entidad, un elemento de la Extensión que declara cuánto hay, y un elemento de la Fecha que indica cuando el material fue creado. Normas de estructura para la descripción archivística como MARC (Machine-Readable Cataloging) y EAD, son bien desarrolladas y ampliamente usadas, sin embargo, no son las únicas. A pesar de facilitaren el intercambio de información, no

FIGURA 4

Ámbito de Aplicación de las Actuales Normas Internacionales de Descripción Archivística



proporcionan ninguna guía sobre como exactamente entrar con la información de algún elemento específico. Sin las correspondientes normas de estructura la información se presenta incoherentemente.

Las normas de contenido, por su vez, proporcionan las reglas que declaran exactamente como entrar la información en cada elemento de la estructura de datos. Por ejemplo, para el elemento de Extensión, la norma de contenido de datos informará como expresar la extensión del material archivístico en términos de número de ítems, indicando si el espacio del estante estará en pies lineales, metros o alguna otra medida [11].

ESTÁNDARES XML EN LA DESCRIPCIÓN ARCHIVÍSTICA

Actualmente los dos principales estándares reconocidos y adoptados por diversas instituciones en todo el mundo para la descripción archivística utilizando el XML son: EAD y EAC. A través de sus específicas DTDs u otras herramientas auxiliares, como las bibliotecas de etiquetas, posibilitan el intercambio de datos e incrementan la cantidad de informaciones disponibles en el campo archivístico.

A pesar de no tener una normalización formal o de *iure*, como otros esquemas de metadatos como Dublín Core, por ejemplo, tanto EAD (ya más consolidado) cuanto EAC (mas reciente y aguardando aún su versión final) constituyen normas de facto, respetadas y utilizadas,

desarrolladas en conformidad con las normas internacionales existentes y aunque la aprobación por parte de una organización como ISO conllevaría algunos aspectos positivos, las propias políticas, procedimientos y intereses de ese tipo de organización (como cobrar sumas desorbitadas por las publicaciones para financiar sus operaciones) indican el riesgo de pérdida del control de las normas y llevan muchos a considerar que estándares archivísticos deben ser gestionados por archiveros [12].

EAD y EAC no incluyen también reglas relativas a cómo deben formalizarse los datos dentro de los elementos, ni tampoco proporcionan las formas de los nombres/términos o los códigos que están autorizados para su inclusión en un determinado elemento. Para tanto son necesarias normas de contenido y de valores de datos, nacionales o internacionales, para facilitar la interoperabilidad entre distintos sistemas. En EE.UU y Canadá existen normas de contenido archivísticas muy utilizadas con formatos MARC, respectivamente: Archives, Personal Papers and Manuscripts (APPM) y Canadian Rules for Archival Description (RAD) [13].

EAD

EAD o Descripción Archivística Codificada es un estándar para codificar instrumentos de descripción archivística. Está escrita en XML en la forma de una Definición de Tipo de Documento (DTD) y tiene el objetivo de describir, controlar y suministrar acceso para los materiales archivísticos producidos por archivos, bibliotecas, asociaciones, agencias y otras organizaciones.

La EAD fue desarrollada originalmente en SGML (posteriormente convertida para XML), entre 1993 a 1995, a partir de un proyecto de la Universidad de California en Berkeley, EE.UU, pero en el transcurso de los años se transformó en un esfuerzo cooperativo patrocinado por varias instituciones y asociaciones profesionales.

La primera versión de EAD (EAD DTD Version 1.0) es de 1998 y la segunda y actual, Version 2002, está disponible desde diciembre del referido año. Actualmente la Society of American Archivists (SAA) a través del EAD Working Group es la responsable por su contenido intelectual y por lo monitoreo y asistencia de su corriente desenvolvimiento, en cuanto la Library of Congress, EE.UU, a través del *Network Development and MARC Standards Office* por su mantenimiento online [14].

Asociados a EAD hay tres documentos técnicos [15]:

- La DTD: EAD DTD, *EAD Technical Document No. 1*, con material complementario como archivos de entidades opcionales.
- La biblioteca de etiquetas: *EAD Tag Library Version 2002*, *EAD Technical Document No. 2*, documento de referencia sobre cada uno de los elementos y atributos definidos en EAD, con apéndices útiles como pasarelas (crosswalks) a otros estándares (ISAD(G) 2000 y MARC21) y ejemplos de uso.
- Las directrices de aplicación: *EAD Version 2002 Application Guidelines*, *EAD Technical Document No. 3*.

EAC

EAC o Contexto Archivístico Codificado es un estándar para estructurar e intercambiar información sobre los creadores de materiales archivísticos. Esta en desarrollo desde 2001 cuando un grupo internacional de archivistas se reunió en un *Meeting* en Toronto, Canadá y definió los *principios y criterios para diseñar, desenvolver y mantener un esquema figurativo y una estructura de comunicación para la información de contexto archivístico* [16].

Los registros que pueden ser creados utilizando EAC trabajan de una manera similar a los archivos de autoridad de las bibliotecas pero con énfasis en la provisión de historias biográficas o administrativas para personas u organizaciones interesadas [7].

La denominación EAC fue utilizada con el objetivo de enfatizar las relaciones con EAD, y *técnicamente constituye una DTD de XML y SGML que, en este caso, regula la codificación de la información sobre el contexto archivístico, proporcionando un formato electrónico para los registros de autoridad ajustados a la ISAAR(CPF)* [10].

Para una interpretación exacta de los registros de un archivo es necesaria la descripción de las circunstancias que rodearan su creación y uso, sendo la primera el registro de la información sobre la responsabilidad de creación, normalmente centrada en una organización o persona(s). Con esta información, los usuarios pueden entender los archivos más completamente, una vez que sabrán el contexto en lo cual la organización o persona operó y creó los archivos [17].

Actualmente el Ad Hoc EAC Working Group es responsable por su desarrollo, y aunque oficialmente no posa ser considerado normativo, ya que la versión sendo

utilizada aún es Beta, EAC ha sido ampliamente utilizado de forma experimental. Las herramientas existentes asociadas al EAC tuvieron la última revisión en agosto de 2004 y son las siguientes:

- EAC Beta DTD = FTP EAC Beta DTD y EAC Beta Tag Library
- EAC Beta XML Schema = EAC Schema Diagram y EAC Beta XML Schema
- EAC Relax NG = EAC Beta Relax NG Schema

Otras Normas y Estándares utilizados en la Archivística

Además de EAD y EAC, específicos estándares archivísticos, otras normas y estándares desarrollados en diferentes campos son también utilizados en la descripción de los archivos para informaciones adicionales. La Figura 5 presenta algunos de los principales con una breve explicación.

INICIATIVAS DE NUEVAS HERRAMIENTAS EN EL CAMPO ARCHIVÍSTICO

Como estándares para codificación, EAD y EAC apoyan una efectiva estructuración y distribución de información sobre archivos para los usuarios. Sin embargo, como herramientas únicas, no pueden dar a los usuarios acceso para el contenido actual de los materiales archivados, por eso, existe un gran potencial para el desarrollo de recursos que faciliten la integración de EAD y EAC con otras herramientas que permitan el uso remoto de contenidos de colecciones archivistas.

Para ejemplificar algunas de las diferentes iniciativas que utilizan y/o integran estándares XML en el campo archivístico, son presentados en la secuencia, tres de los principales proyectos desarrollados con sus respectivos resúmenes:

- *LEADERS – Linking EAD to Electronically Retrievable Sources* [7]

Proyecto financiado por la Arts and Humanities Research Board (AHRB) y que tiene como base la School of Library, Archive and Information Studies de la University College London. Su objetivo es incrementar el acceso de los usuarios remotos a los archivos suministrando los medios para presentar las fuentes archivistas dentro de su contexto. Esto es obtenido uniendo las codificaciones de las transcripciones y de las imágenes digitalizadas de papel basadas en los materiales

FIGURA 5

Otras Normas y Estándares utilizadas en la Archivística

Normas/ Estándares	Descripción
TEI (Text Encoding Initiative) [18]	Internacional e interdisciplinar estándar que facilita a las bibliotecas, museos, editores, y estudiosos individuos representar una variedad de textos lingüísticos y literarios para pesquisa online, enseñanza y preservación. Presentada en forma de una DTD XML/SGML.
MARC (Machine Readable Cataloging Records) [19]	Formato base de las iniciativas de automatización en bibliotecas, posibilita el intercambio y compartimiento de datos. Tiene DTD en XML/SGML y una herramienta llamada MARCXML para conversión de las informaciones entre los formatos.
DCMES (Dublin Core Metadata Element Set) [20]	Usado para complementar métodos existentes de pesquisa e indexación de metadatos web, es compuesto básicamente por un conjunto de 15 elementos de información que pueden ser usados para describir una gran variedad de recursos de información en la Internet.

de archivos (contenido) a las codificaciones de los instrumentos de descripción y de registros de autoridades (contexto). El proyecto pretende desarrollar un conjunto común de instrucciones a partir de tres diferentes codificaciones XML: EAD para instrumentos de descripción, EAC para registros de autoridad y TEI (Text Encoding Initiative) para transcritos. Hasta el presente, EAD y EAC son frecuentemente usados independientemente de TEI e no existe un ambiente generalizado para unir los tres esquemas de codificación. El TEI posibilita que textos electrónicos de todos os tipos sean pesquisados y presentados al usuarios en una variedad de diferentes maneras. Para LEADERS cuando EAD, EAC y TEI son reunidos al lado de las imágenes digitalizadas de archivos el beneficio potencial para los usuarios es amplio.

- *COVAX – Contemporary Culture Virtual Archive in XML* [21]

Proyecto financiado por la Information Society Technology, es un consorcio que reúne organizaciones de diferentes países europeos que pretende analizar y

diseñar las soluciones técnicas para permitir el acceso, a través de Internet, a descripciones de documentos de archivos, bibliotecas y museos. La viabilidad del proyecto se comprobará mediante un prototipo que incluirá una muestra significativa de las distintas clases de documentos existentes, sobre los que se construirá un sistema global de búsqueda y recuperación. El proyecto se basa en el supuesto de que una enorme cantidad de descripciones creadas por bibliotecas, archivos y museos pueden hacerse disponibles en Internet convirtiendo los registros existentes, o creando otros nuevos, a las DTD apropiadas de SGML/XML. Se ha considerado como núcleo operativo del proyecto la utilización de XML y de las distintas DTD ya en uso en bibliotecas (MARC DTD), archivos (EAD) y museos (AMICO), así como la DTD definida por la TEI para el marcado de las versiones electrónicas de textos culturales. El empleo de SGML/XML permitirá la normalización, interoperabilidad e interconexión entre bibliotecas, archivos y museos en los procesos de almacenamiento, consulta, búsqueda y recuperación de todo tipo de descripciones y documentos. La aplicación de XML mejorará la distribución normalizada de fuentes documentales primarias vía Internet y facilitará su acceso a un mayor número de usuarios del mundo entero.

- *CUSTARD project (Canadian-U.S. Task Force on Archival Description)* [11]

Proyecto financiado por el National Endowment for the Humanities (NEH) que intenta reconciliar la norma americana de descripción APPM, la canadiense RAD, y la ISAD(G) para crear un conjunto de reglas descriptivas que posan ser usadas con EAD y MARC21. El desarrollo de normas de contenidos ha tenido lugar al largo de las líneas nacionales (APPM en el EE.UU, RAD en Canadá, *Manual of Archival Description (MAD)* en Bretaña, y *Australian Common Practice Manual*), con resultados muy diferentes. Como los medios de intercambio de información internacional se pusieran disponibles más prontamente, los beneficios de estar de acuerdo con las normas internacionales se pusieran claros. Las normas internacionales de alto nivel ISAD(G) e ISAAR(CPF) fueran desarrolladas por el ICA, pero ellas todavía requieren reglas más detalladas para hacerlas útiles. El objetivo de CUSTARD es crear una norma robusta de contenido para la descripción de materiales archivísticos basada en una armonización de APPM y RAD dentro del *framework* de ISAD(G), y totalmente compatible con las normas de estructura de datos existentes.

CONCLUSIÓN

La integración de las informaciones archivísticas en redes, en niveles nacionales e internacionales, visando el desarrollo de sistemas integrados para el compartimiento de los datos, es hoy una realidad cada vez mas presente. Sin embargo, la consecución de ese objetivo está directamente relacionada con el desarrollo y amplia utilización de normas o estándares que posibiliten descripciones comunes entre los diferentes archivos.

Basados en la tecnología XML, EAD y EAC constituyen los principales estándares para descripción e intercambio de datos archivísticos, pero en función de no posibilitaren una efectiva descripción de los contenidos, son utilizados en diferentes proyectos en conjunto con otras normas o estándares.

El XML con su facilidad de uso y cada vez mayor aceptación ha revolucionado el mundo de la informática y por extensión impactado mucho fuertemente en todas las otras áreas. Específicamente en la archivística, posibilitó a través de EAD y EAC una descripción mas detallada en relación a las normas existentes y una interoperabilidad en los niveles básicos fundamental para el acceso global de los registros electrónicos.

Con la amplia utilización de los estándares XML evaluaciones e intercambios de experiencias posibilitarán el perfeccionamiento y desarrollo de nuevas herramientas para el control e intercambio archivístico.

Artigo submetido em 29/11/2005 e aceito em 07/03/2007.

REFERENCIAS

- [1] W3 CONSORTIUM. *Extensible Markup Language (XML)*. Disponível em: <<http://www.w3c.org/XML/>>. Acesso em: 25 abr. 2005.
- [2] SANCHÉS FERNANDÉZ, Luis; DELGADO KLOOS, Carlos. XML: el ASCII del siglo XXI. *Novática/Upgrade*, n. 158, p. 8-12, jul./ago. 2002.
- [3] W3 CONSORTIUM. *XML in 10 points*. Disponível em: <<http://www.w3.org/XML/1999/XML-in-10-points.html>>. Acesso em: 25 abr. 2005.
- [4] LAMBRECHTS, Floris. Aprendendo XML. *LinuxFocus*, artículo n. 242. Disponível em: <http://main.linuxfocus.org/Portugues/Archives/lf-2002_05-0242.pdf>. Acesso em: 25 abr. 2005.

- [5] ALMEIDA, Mauricio Barcillos. Uma introdução ao XML, sua utilização na Internet e alguns conceitos complementares. *Ciência da Informação*, Brasília, v. 31, n. 2, p. 5-13, maio/ago. 2002. Disponível em: <<http://www.ibict.br/cienciadainformacao/viewarticle.php?id=173&layout=abstract>>. Acesso em: 02 maio 2005.
- [6] GUTIÉRREZ MARTÍNEZ, José María; PALACIOS ESCRIBANO, Fernando; GUTIÉRREZ DE MESA, José Antonio. *Estándar XML y tecnologías asociadas*. Madrid: Danysoft, 2003.
- [7] LEADERS: linking EAD to electronically retrievable sources. Disponível em: <<http://www.ucl.ac.uk/leaders-project/index.htm>>. Acesso em: 25 abr. 2005.
- [8] INTERNATIONAL COUNCIL ON ARCHIVES. ISAD (G): General International Standard Archival Description. 2nd edition. 2000. Disponível em: <http://www.ica.org/biblio/cds/isad_g_2e.pdf>. Acesso em: 25 abr. 2005.
- [9] INTERNATIONAL COUNCIL ON ARCHIVES. ISAAR (CPF): International Standard Archival Authority Record for Corporate Bodies, Persons and Families. 2nd edition. 2004. Disponível em: <<http://www.ica.org/biblio/ISAAR2EN.pdf>>. Acesso em: 25 abr. 2005.
- [10] SANTAMARÍA GALLO, Abelardo. Nuestras normas nacionales de descripción archivística en el marco internacional. *Archivamos*, n. 47-48, p. 5-14, 2003.
- [11] SOCIETY OF AMERICAN ARCHIVISTS. *Statement of principles for the CUSTARD project*. Disponível em: <<http://www.archivists.org/news/custardproject.asp>>. Acesso: 02 maio 2005.
- [12] SANTAMARÍA GALLO, Abelardo. Charlando de normalización con... *Archivamos*, n. 47-48, p. 26-29, 2003. Entrevista realizada por Abelardo Santamaría.
- [13] INTERNATIONAL COUNCIL ON ARCHIVES; COMMITTEE ON DESCRIPTIVE STANDARDS. *Report of the sub-committee on finding Aids guidelines for the preparation and presentation of finding Aids*. 2001. Disponível em: <<http://www.icacds.org.uk/eng/findingaids.htm>>. Acesso em: 10 maio 2005.
- [14] ENCODED ARCHIVAL DESCRIPTION - EAD. *Official EAD version 2002 web site*. Disponível em: <<http://lcweb.loc.gov/ead/>>. Acesso em: 25 abr. 2005.
- [15] NOGÁLES FLORES, J. Tomás. *EAD (Encoged Archival Descripton) version 2002*. Disponível em: <<http://rayuela.uc3m.es/~nogales/cursos/ead.html>>. Acesso em: 02 maio 2005.
- [16] ENCODING ARCHIVAL CONTEXTS - EAC. *Encoding Archival Context: EAC*. Disponível em: <<http://www.library.yale.edu/eac/>>. Acesso em: 25 abr. 2005.
- [17] ENCODING ARCHIVAL CONTEXT - EAC. *Report from Toronto Archival Context Meeting*. Disponível em: <<http://www.library.yale.edu/eac/torontotenets.htm>>. Acesso em: 15 maio 2005.
- [18] TEXT text encoding for interchange consortium. Disponível em: <<http://www.tei-c.org/>>. Acesso em: 10 maio 2005.
- [19] MARC: Machine Readable Cataloging Records. Disponível em: <<http://www.loc.gov/marc/>>. Acesso em: 10 maio 2005.
- [20] DUBLIN Core Metadata Initiative. Disponível em: <<http://dublincore.org/about/>>. Acesso em: 10 maio 2005.
- [21] COVAX: contemporary culture virtual archive in XML. Disponível em: <http://www.covax.org/covax_e/project/project.htm>. Acesso em: 15 maio 2005.