





XAVIER AGENJO BULLÓN, FRANCISCA HERNÁNDEZ CARRASCAL

EAD3, registros de autoridad, datos abiertos y linked open data¹

La evolución de los archivos en la Web

La norma Encoded Archival Description (EAD) tuvo su origen la Universidad de California en Berkeley, en 1993, en un proyecto que tenía como objetivo analizar la viabilidad de una codificación de instrumentos de descripción de archivos, bibliotecas y museos que proporcionara información legible por máquina más compleja y completa que la de los registros MARC. Los criterios que debía cumplir este proyecto de norma eran, entre otros, representar la información de los instrumentos de descripción de archivos, especialmente extensa e interrelacionada; conservar las relaciones jerárquicas existentes entre los niveles de descripción y representar la información descriptiva heredable de un nivel jerárquico a otro.

Con ser estos requisitos fundamentales, nos interesa resaltar el motivo que provocó la necesidad de crear esta norma, que no es otro que el reconocimiento del “creciente papel de las redes de acceso a la información” y la necesidad de preparar las complejas descripciones de archivos para este fin. En 1993 el servidor <http://info.cern.ch>, el primer servicio Web, tenía solo tres años de edad, pero

Xavier Agenjo Bullón (email: xavier.agenjo@larramendi.es)

Fundación Ignacio Larramendi

Francisca Hernández Carrascal (email: francisca.hernandez@digibis.com)

DIGIBÍS, Producciones Digitales

Recibido: 25-04-2016. Aceptado: 28-04-2016

Citación: Agenjo Bullón, Xavier y Hernández Carrascal, Francisca (2016). “EAD3, registros de autoridad, datos abiertos y linked open data”. *Tábula*, n. 19, pp. 331-348

su influencia ya estaba cambiando el mundo de la información. La primera propuesta de EAD² debía responder ya a lo que avecinaba como futuro inmediato.

En este sentido es muy significativa la elección del lenguaje de codificación, SGML³, que permitía el intercambio de información entre sistemas y aplicaciones, y del que HTML, el primer lenguaje de la Web, y XML serían subconjuntos. EAD DTD v1.0, data de 1998⁴, el mismo año en el que se publicó también la primera versión de XML⁵, con la finalidad de servir, recibir y procesar SGML en la Web, del mismo modo en que era posible hacerlo con HTML. La versión 2002, la segunda, ya reflejaba la normalización ejercida por el W3C y la extensión de XML como lenguaje para el intercambio de información en la Web, e incluía tanto la versión *Definición de Tipo de Documento* de SGML, como un esquema XML.

Y ya era perfectamente compatible con ISAD(G). El esfuerzo normalizador emprendido por el Consejo Internacional de Archivos a principios de los años 90 del pasado siglo tuvo su reflejo en la elaboración y consiguiente extensión de la aplicación de la norma ISAD(G) (1994), lo que allanó considerablemente el camino hacia la normalización de las descripciones, proceso necesario, pero no suficiente para garantizar el intercambio. ISAD(G) no prescribe un formato de codificación y su utilización no implica la existencia de un esquema o formato de intercambio. Por ello, la combinación ISAD(G) y EAD ha sido fundamental para incrementar la presencia de los archivos y de sus datos en la Web. En 2016 ya no cabe preguntarse por qué es recomendable utilizar normas para la descripción o para la codificación, si el intercambio tradicional de documentación entre las oficinas y los archivos ya lo exigía, la administración electrónica no puede funcionar sin ellas, pero, sobre todo, hay que considerar la presión que han ejercido los usuarios de la web y los esfuerzos de los archivos y archiveros para responder a estas expectativas.

Muchos de los sistemas de consulta y herramientas hoy disponibles en la Web no habrían sido posibles sin el intercambio de información entre los archivos y el trabajo cooperativo basado en prácticas unificadas. La página de presentación del Portal de Archivos Españoles (2007)⁶ dejaba bien claro este aspecto: *“El Portal de Archivos Españoles elimina las barreras tradicionales en el acceso a los archivos porque, vía Internet, proporciona la consulta libre y gratuita las 24 horas del día y evita los desplazamientos físicos a los distintos centros. Otra de las ventajas es que permite al usuario obtener en una sola operación información repartida por distintos archivos.”*

La adopción de una normativa archivística de descripción y codificación ha posibilitado la creación de sistemas de consulta globales, y el objetivo de crear sistemas de consulta globales ha presionado para aplicar una normativa de descripción y codificación. El desarrollo de este tipo de sistemas Web como PARES⁷, DARA (Documentos y Archivos de Aragón)⁸, el Archivo Dixital de Galicia, y

otros⁹, se ha basado en la creación de redes y sistemas archivísticos que han fortalecido una sólida comunidad de prácticas normalizadas de descripción y codificación. En último término, el servicio a los usuarios en la Web obliga cada vez más a ofrecer puntos de consulta únicos, que se complementan entre sí y que ofrecen conjuntamente todas las descripciones archivísticas disponibles en una determinada red de archivos.

Esta evolución ha sido muy similar a la que ha tenido lugar en bibliotecas, o en museos, desde los OPAC y sitios web individuales, que requerían múltiples consultas para abarcar todo el contenido y el conocimiento de los recursos donde buscar, a sistemas que priman la localización de la información independientemente del proveedor. De hecho, a pocos usuarios les interesa saber con qué norma se ha descrito algo o qué codificación se ha utilizado. Lo que le interesa primordialmente es poder buscar, seleccionar, localizar y acceder al documento digitalizado, ya sea un documento de acceso público o privado, con las correspondientes garantías. La administración electrónica está basada en esto. Pero para que sea así, es necesario que por detrás, en la trastienda, o por debajo del capó, haya una fuerte dedicación a la aplicación de las normas.

Estas son las líneas de trabajo que han conducido a los sistemas de consulta mencionados, pero, sobre todo, a Hispana (2006), Europeana (2008) y al Portal Europeo de Archivos (2013). El éxito de las funcionalidades de estos portales, y otros a nivel nacional como los National Archives of Australia¹⁰, Archives Hub¹¹, NARA¹², etc., nos han proporcionado una visión ubicua, una voluntad de estar presentes donde están los usuarios, en la web, de una forma en la que nuestros datos aporten una visión particular al conjunto. Así, nuestros datos se complementan con otros del conjunto y el conjunto no puede verse de forma completa sin los datos desarrollados en los archivos.

Esto aporta un valor informacional para los usuarios tan sustancial, que la solución de los OPAC, por más que sean sitios web, los cuadros de clasificación o los instrumentos de descripción individuales como las únicas herramientas de información disponibles, se manifiestan ya como algo muy limitado. A veces puede parecer que la presencia en Hispana, Europeana o el Portal Europeo de Archivos no es más que una cuestión de prestigio técnico, pero en el fondo es contribuir a que la información, los datos que estamos obligados a proporcionar a los usuarios, se encuentren de la mejor manera posible, conforme al estado de la Web y de las tecnologías de la información en cada momento. Y es también una cuestión de cumplimiento de obligaciones legales: la *Recomendación de la Comisión Europea de 27 de octubre de 2011 sobre la digitalización y accesibilidad en línea del material cultural y la conservación digital*¹³, en su recomendación 7.b requiere condicionar toda financiación pública para proyectos de digitalización a la accesibilidad del material digitalizado a través de Europeana¹⁴.

EAD3: razones para revisar la norma

El Technical Subcommittee for Encoded Archival Description de la Society of American Archivists (TS-EAD) inició en 2010 el proceso de revisión de la norma, con la finalidad de dotar a los archivos de una estructura de datos común que reflejara los cambios más importantes acontecidos en los últimos años, como han sido la extensión del software para la descripción archivística, el constante aumento de la presencia de los instrumentos de descripción de archivos en la Web, la ampliación de la normativa archivística a los registros de autoridad EAC-CPF¹⁵ y la consolidación de Open Data y Linked Open Data¹⁶.

El proceso de revisión de EAD se centró en lograr una mayor consistencia conceptual y semántica en el uso de EAD considerada en muchas ocasiones demasiado flexible y permisiva y, por lo tanto complicada por su variabilidad para conectar, intercambiar o incorporar otras normas y protocolos en mejorar la funcionalidad de EAD para la representación de descripciones en un entorno internacional y multilingüe; y en minimizar el efecto que la nueva norma pudiera tener en las aplicaciones ya existentes¹⁷.

Definitivamente, la versión 3 de EAD se publicó oficialmente el 28 de agosto de 2015¹⁸. En el proceso han participado representantes de distintos sectores, fundamentalmente archivos, pero también universidades, bibliotecas u otras organizaciones como OCLC. Nos gustaría mencionar entre los miembros del TS-EAD a Daniel Pitti, creador de EAD, y a Kerstin Arnold, de los Bundesarchiv alemanes y coordinadora técnica de APEX¹⁹. Kerstin Arnold²⁰ condensa muy bien los objetivos de EAD3: mejorar la interoperabilidad con otras normas archivísticas, mejorar la usabilidad e intercambio de codificaciones EAD, y facilitar el uso de EAD no solo en el marco de un archivo, sino también en proyectos más amplios, sean archivísticos, temáticos, o interdisciplinarios.

Datos abiertos y la información del sector público

Pero la Web, y las tecnologías de la información y de las comunicaciones, no sólo ha presionado a los archivos, bibliotecas y museos a abrir y a conectar su información públicamente, también ha empujado a los gobiernos a desarrollar la administración electrónica, estableciendo mecanismos de interoperabilidad entre las distintas administraciones, a buscar la transparencia de las acciones gubernamentales y a garantizar la participación ciudadana a nivel educativo, político y administrativo²¹.

Estos procesos han reavivado el papel de los archivos como receptores y creadores de una inmensa cantidad de datos, documentos y procedimientos, lo cual puede verse en la activa participación de los archiveros en el análisis, desarrollo

e implantación de las políticas de gestión de los documentos electrónicos. La combinación de las leyes de transparencia²² y de reutilización de la información del sector público²³ ha dado lugar a la publicación de grandes volúmenes de datos y a algunas visualizaciones verdaderamente espectaculares²⁴.

Aunque el movimiento en pro de la apertura y transparencia de los datos gubernamentales y de su libre conocimiento y reutilización es mucho anterior²⁵, la *Open Government Directive*²⁶ de 8 de diciembre de 2009 del presidente Obama se sitúa habitualmente como punto de inicio de la aplicación del concepto de gobierno abierto y de sus principios de transparencia, participación y colaboración. Inevitablemente, esta política de gobierno abierto debía sustentarse en la publicación de datos en la Web e implicaba el desarrollo de proyectos de tecnologías de la información, siendo el catálogo de los datos del Gobierno de los Estados Unidos, *Data.gov*²⁷ el primer portal de este tipo. A esta iniciativa se sumó inmediatamente, en el mismo 2009, la publicación de datos del gobierno del Reino Unido, *Data.gov.uk*, que nació ya inmersa dentro de las tecnologías de la Web Semántica, para permitir la interrelación entre los diferentes conjuntos de datos.

El número de iniciativas de datos abiertos en todo el mundo ha ido creciendo paulatinamente. La organización Open Knowledge (OKFN) mantiene un directorio de portales de datos abiertos²⁸ (519), y sobre estos datos extrae en *Open Data Websites (Portals and Catalogs)*²⁹ los sitios activos (281 sitios). En 2010, el *Catálogo Nacional de Datos Abiertos, Datos.gob.es*³⁰, nace de la *Iniciativa Aporta* (2009), entidad interministerial para la promoción de la cultura de la apertura de la información en España. *Datos.gob.es* ofrece en la actualidad 14.642 conjuntos de datos distintos, muchos de ellos en múltiples formatos. En mayor o menor grado de evolución técnica, todas las Comunidades Autónomas españolas y en gran medida los ayuntamientos³¹ ofrecen diferentes conjuntos de datos abiertos³². Desde luego, los archivos no están bien representados aún.

La información del sector público y los datos abiertos gubernamentales³³ forman parte no sólo de los mecanismos de transparencia, sino también de la reactivación económica. La reutilización de los datos del sector público es uno de los apartados de la *Agenda Digital para Europa* (2010)³⁴, desarrollada para reactivar una economía digital dificultada por la existencia de sistemas incompatibles, la falta de interoperabilidad y la falta de conectividad. Los términos utilizados en la *Agenda* son muy explícitos en este sentido: “*Por ejemplo, los gobiernos pueden estimular los mercados de contenidos ofreciendo la información del sector público en condiciones de transparencia, eficacia y no discriminación. Se trata de una fuente importante de crecimiento potencial de los servicios en línea innovadores*”. La *Acción Clave número 1* contenía el compromiso de revisar la Directiva sobre la reutilización de la información del sector público, lo que fue efectivo desde su publicación en el *Diario Oficial de la Unión Europea*³⁵ en 2013. El objetivo de la *Agenda Digital para Europa* de desarrollar un mercado digital único se incorporó como uno de los

pilares de la *Europe 2020 Strategy*³⁶. Es necesario añadir que el segundo pilar de esta estrategia es la mejora de la interoperabilidad y de las normas³⁷.

En cumplimiento de esta Directiva, se publicó la Ley 18/2015 sobre reutilización del sector público, ya mencionada, que modificaba la Ley 37/2007, y que abría con esta frase: *La información que generan las Administraciones Públicas y los organismos del sector público constituye un importante recurso para promover la economía del conocimiento*. Igualmente, la transposición de las recomendaciones europeas a España dio lugar a la *Agenda Digital para España*³⁸ (2013), que propone, convenientemente actualizada, una serie de objetivos de Tecnologías de la Información y las Comunicaciones (TIC) y de Administración Electrónica para el cumplimiento de los objetivos de la Agenda Digital para Europa en 2020³⁹. Entre sus planes específicos se encuentra el *Plan de impulso de la economía digital y los contenidos digitales*⁴⁰, que incluye el Eje III: *Programa de reutilización de la información del sector público*⁴¹.

Para completar este panorama, el *Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre*⁴², para el ámbito del sector público estatal, establecía la obligación para todos los organismos de la Administración General del Estado de disponer de un plan de medidas de impulso de la reutilización de la información del sector público, por medios electrónicos, lo que ha dado lugar a los “Planes RISP”⁴³. Hasta la fecha se han publicado 34 de estos planes, siendo el más reciente el de la Biblioteca Nacional de España⁴⁴. Por su parte la *Norma Técnica de Interoperabilidad de Reutilización de recursos de información*⁴⁵ establece las condiciones comunes sobre selección, identificación, descripción, formato, condiciones de uso y puesta a disposición de los documentos y recursos de información elaborados o custodiados por el sector público. El portal *The ePSIplatform*⁴⁶, una iniciativa de la DG CONNECT de la Comisión Europea, elabora un marcador, *The PSI Scoreboard*⁴⁷, sobre el estado de la Información del Sector Público y de los Datos Abiertos en Europa, en el que España figura en segundo lugar.

El Catálogo Nacional de Datos Abiertos (data.gob.es)

Las propias descripciones de archivos, bibliotecas y museos constituyen una de las fuentes más importantes para la difusión de la información del sector público. Según dice la Ley 18/2015 en su preámbulo: “*Se ha ampliado el ámbito de aplicación a las bibliotecas, incluidas las universitarias, los museos y los archivos, dado el importante volumen de recursos de información que poseen y los proyectos de digitalización que vienen llevando a cabo*”.

A fecha de 25 de abril de 2016, el intento de localizar la presencia de los archivos en el *Catálogo Nacional de Datos* revela alguna de las deficiencias de este

catálogo. En nuestra opinión la más importante es la ineficacia de la búsqueda, derivada de la ausencia de vocabularios controlados para la indización de los conjuntos de datos. Por ejemplo, el término más utilizado es ‘Estudio cuantitativo’ con 2.011 ocurrencias; mientras los, aparentemente, sinónimos ‘local’, ‘municipio’ y ‘ayuntamiento’ suponen en conjunto el 5,4 % del total de palabras clave; y, lo más importante, el 22 % de los conjuntos de datos no dispone de ninguna palabra clave para su búsqueda. La ausencia de un vocabulario controlado para la indización de los conjuntos de datos provoca que los resultados de la búsqueda ‘archivo’ sea muy ruidosa, puesto que bajo este término se incluye información propia de los archivos, pero también el término *archivo* en el sentido de salvaguardia y almacenamiento de información de cualquier tipo, o el sinónimo fichero.

En cualquier caso, la principal conclusión es que ni en el *Catálogo* se encuentra toda la información que existe de forma abierta en la Web, ni todos los datos disponibles en formatos reutilizables están catalogados. Por ejemplo, no está incluido el *Fichero de Autoridades del Censo-Guía de Archivos de España e Iberoamérica*⁴⁸, que puede descargarse en el esquema EAC-CPF de forma individual. Igualmente las descripciones de PARES pueden descargarse en EAD, y el Directorio de Archivos del Censo-Guía de Archivos de España y de Iberoamérica puede descargarse, también individualmente, en EAG⁴⁹. Sin embargo, *Catálogo* sólo informa de que está disponible en XHTML, como acceso indirecto, al igual que ocurre con el *Boletín Oficial del Estado*. Esta circunstancia hace pensar que en la construcción del *Catálogo* se ha dado más importancia a una rápida recopilación, lo cual es muy comprensible, de fuentes, incluidas las que ya estaban abiertas en la Web antes de la reglamentación legal sobre datos del sector público, y de que fueran técnicamente reutilizables o no. El 9,2 % de los conjuntos de datos están disponibles en formato HTML o XHTML, lenguajes de codificación de sitios web, lo que corrobora nuestra apreciación de que en muchos casos se trata de una recopilación de lo previamente existente, y no tanto una apertura real de nuevos datos.

Se produce también el hecho contrario, hay conjuntos de datos cuyos formatos y protocolos de acceso a los mismos son interoperables y abiertos, pero no están incluidos en el *Catálogo*. Así las entradas de *DARA*, el *Archivo de la Ciudad de Arganda del Rey* o *Catálogo Colectivo de las Bibliotecas de los Archivos Estatales y del CIDA* no proporcionan información sobre la posibilidad de acceder a los datos a través de sus respectivos repositorios OAI-PMH. En el caso de *DARA*, además es posible acceder a los datos en RDF según el Europeana Data Model. Seguramente el caso más llamativo es el de *Hispana*, que aunque aporta a Europeana casi 2,5 millones de descripciones de objetos digitales, generados con la aportación de más de 200 repositorios españoles, y que por su volumen se sitúa entre los cinco primeros proveedores de Europeana, datos que hablan bien claramente de interoperabilidad y apertura, aparece en el catálogo bajo el epígrafe de

formato en HTML. En la misma situación se encuentran otras importantes fuentes de datos como la Biblioteca Digital Hispánica o la Biblioteca Virtual de Prensa Histórica, que según el Catálogo ofrecen sus datos sólo en HTML.

Tampoco están todos los que son. Se aprecia justamente la falta, por lo menos, de esos más de 207 repositorios OAI-PMH, de contenido patrimonial o científico, que aportan su contenido a Hispana, a Recolecta, o a Europea. Pero lo más significativo es la ausencia de conjuntos de datos, especialmente vocabularios de valores nacidos ya en el entorno de Linked Open, que cumplen con las 5 estrellas LOD⁵⁰ y que no están presentes en el Catálogo. Nos referimos a la *Lista de Encabezamientos para Bibliotecas Públicas en SKOS*⁵¹ y a los *Tesauros del Patrimonio Cultural de España en Linked Open Data*⁵². De hecho, tampoco está en el Catálogo el *Directorio Común de Unidades Orgánicas y Oficinas (DIR3)*⁵³, uno de los elementos de la interoperabilidad de la administración electrónica española, ni sus interesantes catálogos⁵⁴ para la normalización de denominaciones en los archivos como el de Catálogo de Comunidades Autónomas, Provincias, Localidades, Tipos de Entidad Pública, Tipos de Unidad Orgánica, y otros.

Hay que mencionar que estos catálogos se difunden únicamente en ficheros Excel (XSLX), pero esto le ocurre también a muchos de los conjuntos del *Catálogo Nacional de Datos*, ya que los ficheros en formatos tipo hoja de cálculo (CSV, XLS, XLSX, ODS y TSV) suponen el 31 % del total (25.910 ficheros). Para completar el panorama diremos que el 9,3 % de los ficheros tienen un formato de texto plano (ASCII, 'Plain') y que PDF representa el 4,4 % del total. Es decir que el 49,6 % de los ficheros se encuentran entre el rango 1 a 3 de estrellas Linked Open Data, el 52,1 % si tenemos en cuenta el formato PC-AXIS utilizado, entre otros, por el Instituto Nacional de Estadística.

Es de esperar que el Catálogo irá evolucionando y que se ampliará paulatinamente el conjunto de datos disponible, en especial los relacionados con los archivos, bibliotecas y museos. La práctica de contribuir con datos ajustados y bien descritos a este Catálogo cambiaría mucho su utilidad. Como ya es sabido, la utilización de una norma de descripción ayuda, pero siempre hay que considerar todas las buenas prácticas anejas a la descripción e indización. En este caso, el hecho de que exista una recomendación europea y nacional para utilizar DCAT⁵⁵ en la descripción de los conjuntos de datos abiertos, no implica por sí misma ni el interés en figurar en el catálogo, ni el cuidado en que las descripciones sean verdaderamente útiles. Se trata de algo que todos hemos experimentado a lo largo de nuestra vida profesional, el hecho de aplicar una norma hace que las buenas descripciones sean más usables, intercambiables y útiles, pero, al contrario, las normas no convierten las malas descripciones en usables. Justamente la interoperabilidad, el intercambio y la reutilización de datos requiere fundamentalmente que estos sean de calidad.

Estamos seguros de que el *Catálogo* se encuentra en una fase de acumulación y que se irá paulatinamente transformando en un directorio detallado y comprensivo de los conjuntos de datos abiertos disponibles en España. Por supuesto, el papel de los archiveros, bibliotecarios y museólogos, ya acostumbrados a los datos abiertos aunque no se llamaran así, y a describir y codificar correctamente los conjuntos de datos que se abren al uso público, así como el experto uso de las técnicas de descripción, catalogación e indización (términos *vintage* vigentes) contribuya a que el *Catálogo* se convierta en una herramienta útil y usable.

Codificación de recursos y datos en EAD3

Veremos a continuación los distintos usos que los datos abiertos y los datos abiertos vinculados disponibles en España pueden tener en una descripción EAD3. En este apartado analizaremos algunas de las formas de incorporar los datos del sector público al trabajo de descripción archivística, desde el punto de vista de la creación de datos como de su consumo⁵⁶.

Elección de nombres y términos

En esta categoría se encuentran muchos de los conjuntos de datos mencionados, pero especialmente los ficheros o catálogos de registros de autoridad, los nomenclátors, o directorios de entidades, organismos y oficinas, etc.

Muchos recursos están descritos en más de una fuente, especialmente los nombres geográficos, bien según formas idénticas, análogas o diferentes, pero no dejan de ser fuentes ya disponibles que pueden ahorrar mucho esfuerzo en la creación de nombres de personas, familias e instituciones, o de descriptores. Desde luego, es especialmente recomendable revisar los datos disponibles dentro de cada ámbito para utilizar aquellos más convenientes y que tengan una garantía de continuidad en su mantenimiento, para evitar proyectos que publican una única vez sus datos y no vuelven a actualizarse, e incluso desaparecen de la Web.

En el siguiente ejemplo se pueden ver nombres obtenidos del Fichero de Autoridades del Censo-Guía, y del fichero de Autoridades de la Biblioteca Nacional en datos.bne.es. En el primer caso, el registro de autoridad dispone de un número de identificación que por sí mismo no lleva a ninguna descripción en la Web, sin embargo en el caso de la Biblioteca Nacional el identificador es un URI que dirige a una descripción RDF.

Como puede verse este mecanismo permite establecer relaciones con cualquier fuente de datos Linked Open Data como el Virtual International Authority File, Wikidata, etc.

```

<origination>
  <corpname lang="spa" source="Censo-Guía de Archivos de España e Iberoamérica. Fichero de Autoridades" identifier="ES-47186-ARCHV46635">
    <part>Audiencia Territorial de Valladolid</part>
  </corpname>
</origination>

<origination>
  <corpname lang="spa" source="Biblioteca Nacional de España. Autoridades" identifier="http://datos.bne.es/entidad/XX208544" relator="http://datos.bne.es/def/OP6001">
    <part>España. Audiencia Territorial de Valladolid</part>
  </corpname>
</origination>

<corpname source="VIAF" identifier="http://viaf.org/viaf/267246561">
  <part>España. Audiencia Territorial de Valencia</part>
</corpname>

<corpname source="Wikidata" identifier="http://www.wikidata.org/entity/Q17496844">
  <part>Archivo Histórico de Euskadi</part>
</corpname>

```

En EAD 2002 existía el atributo *authfilenumber* que ha desaparecido en EAD3, mientras que en esta versión del esquema se introduce el atributo *Identifier* (... “es un número, código o cadena (p.e., un URI) que identifica el término utilizado en un vocabulario controlado, taxonomía, ontología u otro sistema de organización del conocimiento”).

Términos multilingües

Otra de las posibilidades que ofrece EAD3 es la utilización del atributo *lang* en todos los elementos con contenido (no vacíos), para especificar el idioma en el que está expresado un determinado elemento. Este aspecto es diferente del idioma en el que se realiza una descripción archivística y permite repetir elementos en varios idiomas, lo que confiere a EAD3 un carácter multilingüe.

En EAD 2002 el atributo *langencode* podía utilizarse únicamente en los elementos *abstract*, *language* y *langmaterial*. En el ejemplo que sigue puede apreciarse el uso del atributo *lang* en el elemento *subject*, en el que se han utilizado términos y URIs procedentes de la *Lista de Encabezamientos para Bibliotecas Públicas en SKOS*. Este recurso de la Subdirección General de Coordinación Bibliotecaria tiene como objetivo proporcionar una lista de encabezamientos de mate-

rias en todos los idiomas del Estado Español (el elemento *source* de los ejemplos incluye el nombre de cada una de las fuentes), vinculados además a otros recursos Linked Open Data procedentes de la *Library of Congress Subject Headings*⁵⁷, *RAMEAU (Répertoire d'autorité-matière encyclopédique et alphabétique unifié)*⁵⁸, y del *Gemeinsame Normdatei (GND)*⁵⁹.

El hecho de que cada vez haya más recursos vinculados⁶⁰, como es el caso de las autoridades de la Biblioteca Nacional de España⁶¹, permite reutilizar esos vínculos para trasladarlos a las descripciones archivísticas y a sus registros de autoridad, posibilitando su búsqueda multilingüe.

```
<controlaccess>
  <subject lang="eus" identifier="http://lemav.sgcb.mcu.es/Autoridades/
  LEMAV2015007092/concept" source="Lista de encabezamientos de materia en
  galego (LEMAG)">
    <part>Albaitaritza</part>
  </subject>
  <subject lang="cat" identifier="http://lemac.sgcb.mcu.es/Autoridades/
  LEMAC201235805/concept" source="Llista de encapçalaments de matèria de la
  Biblioteca de Catalunya (LEMAC)">
    <part>Veterinària</part>
  </subject>
  <subject lang="glg" identifier="http://lemag.sgcb.mcu.es/Autoridades/
  LEMAG201304310/concept" source="Lista de encabezamientos de materia en
  galego (LEMAG)">
    <part>Veterinaria</part>
  </subject>
  <subject lang="spa" identifier="http://id.sgcb.mcu.es/Autoridades/
  LEM201013155/concept" source="Lista de Encabezamientos de Materia para las
  Bibliotecas Públicas en SKOS">
    <part>Veterinaria</part>
  </subject>
  <subject lang="eng" identifier="http://id.loc.gov/authorities/subjects/
  sh85143046" source="Library of Congress Subject Headings">
    <part>Veterinary medicine</part>
  </subject>
</controlaccess>
```

Fuentes e identificadores en el entorno de la Administración Electrónica

Al igual que en la categoría anterior algunos recursos, especialmente los publicados por el INE, incluyen códigos de identificación. El siguiente ejemplo está construido sobre el *Listado de Unidades Orgánicas de EELL* (se entiende que son Entidades Locales, pero nos tememos que las máquinas no lo interpreten correctamente) del

Directorio Común de Unidades Orgánicas y Oficinas (DIR3) del Portal de la Administración Electrónica.

Como puede apreciarse, la transformación de estos identificadores en URIs y los datos en formato RDF, lo que no sería otra cosa que su adaptación a Linked Open Data y a la Web Semántica, supondría un gran avance para la interoperabilidad de la administración electrónica, especialmente en aquellos casos en los que estos códigos sirven para elaborar otros identificadores como los códigos de archivos o los identificadores de registros de autoridad.

```
<corpname lang="cat" source="Directorio Común de Unidades Orgánicas y Oficinas (DIR3). Listado de Unidades Orgánicas de EELL" identifier="L05080042">
  <part>Mancomunitat Intermunicipal de Serveis d'Alella, El Masnou i Teià</part>
</corpname>
```

```
<corpname lang="spa" source="Directorio Común de Unidades Orgánicas y Oficinas (DIR3). Listado de Unidades Orgánicas de EELL" identifier="L01080039">
  <part>Ayuntamiento de Alella</part>
</corpname>
```

```
<geogname>
  <part lang="spa" source="Relación de municipios y códigos por provincias a 01-01-2009. Instituto Nacional de Estadística" identifier="08-0044">Alella</part>
</geogname>
```

```
<geogname lang="spa" identifier="2705370" source="Nomenclátor Geográfico Básico (IGN)" locality="Municipio">
  <part>Candeleda</part>
  <geographiccoordinates coordinatesystem="INSPIRE">40.150000 - 5.230000</geographiccoordinates>
</geogname>
```

También se pueden utilizar vínculos y relaciones con algunas de las fuentes Linked Open Data más utilizadas en la Web como GeoNames.

```
<geogname lang="spa" identifier="http://sws.geonames.org/3126622/about.rdf" source="GeoNames" locality="http://www.geonames.org/ontology#P.PPLA3">
  <part>Candeleda</part>
  <geographiccoordinates coordinatesystem="wgs84">40.15521 - 5.24045</geographiccoordinates>
</geogname>
```

Descripción de roles y funciones

En el entorno de los *Metadatos para la Gestión del Documento Electrónico (e-EMG-DE)*⁶² se sugiere que cada entidad genere su propio vocabulario de funciones (y roles), aunque no estaría de más que se intentara a nivel nacional la creación de un vocabulario RDF que recogiera estas funciones, a poder ser multilingüe, para su libre utilización por todos los archivos. Con toda seguridad el vocabulario, aunque amplio, contribuiría a la normalización y relación entre las descripciones de archivos. También podría desarrollarse una terminología de este tipo a partir de los vocabularios controlados de los archivos individuales, pero el soporte de la Administración Electrónica estatal o comunitaria supondría un fuerte respaldo.

```
<origination label="Fotógrafo">
  <persname lang="spa" source="Censo-Guía de Archivos de España e Iberoamerica. Fichero de Autoridades" identifier="http://censoarchivos.mcu.es/CensoGuia/productordetail.htm?id=46415" relator="http://id.loc.gov/vocabulary/relators/pht">
    <part>Capa, Robert</part>
    <part> 1913-1954</part>
  </persname>
</origination>
```

En este ejemplo hemos sustituido el literal del rol ‘fotógrafo’ por su URI en la lista *MARC Code List for Relators*⁶³, disponible en RDF en <http://id.loc.gov>. Igualmente hemos sustituido el código de identificación del registro de autoridad del Censo-Guía por su URL, lo que nos dará una idea de la ventaja que supondría que los registros del Censo-Guía estuviesen identificados por URIs. También se pueden utilizar otras ontologías como Archives Hub⁶⁴, más adecuada al entorno archivístico (véase por ejemplo la propiedad <http://data.archiveshub.ac.uk/def/isOriginationOf>), o la DBpedia, de ámbito mucho más general, pero que es la fuente de recursos Linked Open Data más utilizada de la Web (véase, por ejemplo la propiedad <http://dbpedia.org/ontology/author>).

Datos contextuales

Muchos datos contextuales de las descripciones de archivos pueden encontrarse parcialmente en Linked Open Data, especialmente en recursos generalistas como DBpedia. El aprovechamiento de estos recursos puede tener varios efectos, hacer más denso el entramado de interrelaciones entre los datos existentes en la web y, por tanto facilitar la consulta a los usuarios, pero también reducir el esfuerzo de creación de descripciones. Indudablemente, es un camino de ida y vuelta que muchos archivos, bibliotecas y museos están explotando: reutilizar datos de la Wikipedia y contribuir con sus propios datos a la Wikipedia.

En la versión 2002 de EAD no era posible establecer estas relaciones, justamente por la falta del atributo *href* (localizador de un recurso en línea), que ahora sí está definido.

```
<bioghist localtype="Historia institucional" lang="spa">
  <p>
    <ref href="https://es.wikipedia.org/wiki/Consejo_Superior_de_Investigaciones_Cient%C3%ADficas" linktitle="Consejo Superior de Investigaciones Científicas en Wikipedia">La agencia estatal Consejo Superior de Investigaciones Científicas (CSIC) es la mayor institución pública dedicada a la investigación en España y la tercera de Europa. Adscrita al Ministerio de Economía y Competitividad de España, a través de la Secretaría de Estado de Investigación, su objetivo fundamental es desarrollar y promover investigaciones en beneficio del progreso científico y tecnológico, para lo cual está abierta a la colaboración con entidades españolas y extranjeras. [...]</ref>
  </p>
</bioghist>

<bioghist localtype="Historia institucional">
  <p>
    <ptr href="http://dbpedia.org/resource/Spanish_National_Research_Council" linktitle="CSIC en DBpedia"/></p>
</bioghist>
```

Descripciones archivísticas relacionadas

La introducción de atributos que permitan establecer vínculos que se acaba de mencionar es especialmente interesante para identificar las descripciones archivísticas relacionadas, ya sea dentro del propio archivo o externas al mismo. Esto facilita que las relaciones no sean meramente textuales, sino enlaces activos para los usuarios y relaciones semánticas bien definidas para la Web de los Datos.

```
<relatedmaterial>
  <head>[Otras colecciones sobre Isaac Newton en la University of Cambridge]</head>
  <archref>
    <ptr href="https://janus.lib.cam.ac.uk/db/node.xsp?id=EAD%2FGBR%2F0012%2FMS%20Add.3958-4007" linktitle="The Portsmouth Collection"/>
  </archref>
  <archref>
    <ptr href="https://janus.lib.cam.ac.uk/db/node.xsp?id=EAD%2FGBR%2F0012%2FMS%20Add.9597" linktitle="The Macclesfield Collection"/>
  </archref>
</relatedmaterial>
```

Se pueden construir ejemplos análogos para cualquier tipo de relación entre descripciones archivística, como en los elementos *altformavail*, *originalsloc*, etc.

Otros recursos de bibliotecas y museos

Se pueden utilizar procedimientos similares para establecer relaciones con recursos bibliográficos o museísticos existentes en la Web y/o en Linked Open Data, con los mismos efectos en la usabilidad e interrelación entre recursos mencionada en el ejemplo anterior.

```
<bibliography>
  <bibref>
    <title>
      <part>Castelao e os galeguistas do interior : cartas e documentos,
1943-1954</part>
    </title>
    <ptr href="http://datos.bne.es/edicion/Mimo0001679216.html"
linktitle="Castelao e os galeguistas do interior : cartas e documentos, 1943-
1954"/>
  </bibref>
</bibliography>
```

Conclusiones

La norma EAD3 supone un gran avance para renovar las descripciones archivísticas dentro de la Web, ya que permite la vinculación y reutilización de recursos propios de los archivos, pero también de otras fuentes de bibliotecas y museos o del sector de la información pública. Un uso intensivo de esas capacidades permitirá:

1. mejorar la usabilidad de las aplicaciones, ofreciendo información contextual de contenido generalista y especializado complementario a las propias descripciones archivísticas.
2. facilitar la búsqueda, navegación, e interrelación de datos entre los documentos de un mismo archivo, entre diferentes archivos, y entre sistemas de información de archivos, bibliotecas y museos.
3. reutilizar los datos abiertos producidos en el sector público que se generan en el marco de las políticas de transparencia e interoperabilidad de la administración y en el que los archivos tienen un papel crucial como productores y consumidores de datos abiertos y de datos abiertos vinculados.

4. presionar para que los datos del sector público español se difundan en formatos abiertos y bien documentados, conforme a las recomendaciones del W3C para Linked Open Data.

Notas

¹ Todos los enlaces citados se han comprobado en esta fecha.

² Development of the Encoded Archival Description DTD. 2013. <https://www.loc.gov/ead/eaddev.html>

³ ISO 8879:1986. Information processing — Text and office systems — Standard Generalized Markup Language (SGML)

⁴ La traducción española se publicó en 2000:

EAD: Descripción archivística codificada. Repertorio de etiquetas / elaborado y actualizado por el Encoded Archival Description Working Group de la Society of American Archivist y el Network Development and Marc Standards Office de la Library of Congress; traducción de la Fundación Histórica Tavera. Versión 1.0. Madrid: Fundación Histórica Tavera, 2000.

EAD : Descripción archivística codificada. Directrices de aplicación / elaboradas por el Encoded Archival Description Working Group de la Society of American Archivists ; traducción de la Fundación Histórica Tavera. Versión 1.0. Madrid : Fundación Histórica Tavera, 2000. Véase Anejo C “Preguntas más frecuentes”, pp. 232-236.

⁵ Extensible Markup Language (XML) 1.0. W3C Recommendation 10-February-1998. <https://www.w3.org/TR/1998/REC-xml-19980210>

⁶ <http://www.mcu.es/novedades/novedadesPARES.html>

⁷ <http://pares.mcu.es/>

⁸ <http://www.sipca.es/dara/>

⁹ LÓPEZ FALANTES, Lola. Archivos en acceso abierto. Experiencias en nuestro país (España) = Open Access Initiative in Archives. Experiences in our country (Spain), 2015 [Preprint]. <http://eprints.rclis.org/28805/>

¹⁰ <http://www.naa.gov.au/>

¹¹ <http://archiveshub.ac.uk/>

¹² <http://www.archives.gov/>

¹³ Recomendación de la Comisión Europea de 27 de octubre de 2011 sobre la digitalización y accesibilidad en línea del material cultural y la conservación digital (2011/711/UE). Diario Oficial de la Unión Europea 29.10.2011. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2011:283:0039:0045:ES:PDF>

¹⁴ AGENJO, Xavier y CAMPILLEJO, Maribel. “El observatorio de la digitalización en Europa” En Desafíos y oportunidades de las Ciencias de la Información y la Documentación en la era digital: actas del VII Encuentro Ibérico EDICIC 2015 (Madrid, 16 y 17 de noviembre de 2015). Universidad Complutense de Madrid, Madrid. ISBN 978-84-608-3330-7. http://eprints.sim.ucm.es/34723/1/344-Agenjo_el-observatorio.pdf

¹⁵ EAC-CPF Homepage. <http://eac.staatsbibliothek-berlin.de/>

¹⁶ <https://www.w3.org/standards/semanticweb/data>

¹⁷ Encoded Archival Description: Tag Library. Version EAD3. Prepared and maintained by the Technical Subcommittee for Encoded Archival Description of the Society of American Archivists. Chicago: SAA, 2015. <http://www2.archivists.org/sites/all/files/TagLibrary-VersionEAD3.pdf>.

¹⁸ <https://www.loc.gov/ead/EAD3taglib/index.html>

- ¹⁹ <http://www.apex-project.eu/index.php/en/>
- ²⁰ ARNOLD, Kerstin. EAD3 and the consequences of the new version. <http://www.apex-project.eu/index.php/en/articles/149-ead3-and-the-consequences-of-the-new-version>
- ²¹ Excede del ámbito de esta comunicación, aunque es de suma importancia, detallar la normativa y reglamentación de la administración electrónica española y europea, que en todo caso puede encontrarse en la numerosa bibliografía y especialmente en el PAe (Portal de la Administración Electrónica), http://administracionelectronica.gob.es/pae_Home.html#.VxtSC_I9670
- ²² Ley 19/2013, de 9 de diciembre, de transparencia, acceso a la información pública y buen gobierno. BOE número 295, de 10/12/2013. <https://www.boe.es/buscar/act.php?id=BOE-A-2013-12887>
- ²³ Ley 18/2015, de 9 de julio, por la que se modifica la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público. BOE número 164, de 10 de julio de 2015. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2015-7731
- ²⁴ Data USA. The Most Comprehensive Visualization of U.S. Public Data. <http://datausa.io/>
- ²⁵ TAUBERER, Joshua. Open Government Data: The Book. Second Edition: 2014. <https://opengovdata.io/>
- ²⁶ https://www.whitehouse.gov/sites/default/files/omb/assets/memoranda_2010/m10-06.pdf
- ²⁷ <http://www.data.gov/metrics>
- ²⁸ Data Portals: A Comprehensive List of Open Data Portals from Around the World <http://dataportals.org/>
- ²⁹ <http://data.okfn.org/data/okfn/opendatasites>
- ³⁰ <http://datos.gob.es>
- ³¹ UNE 178301:2015. Ciudades Inteligentes. Datos Abiertos (Open Data). (Versión corregida en fecha 2015-07-29). <http://www.aenor.es/aenor/normas/normas/fichanorma.asp?tipo=N&codigo=N0054318#.Vx5mvP19670>
- ³² Mapa actual de las iniciativas Open Data en España. Actualizado a enero de 2015. <http://datos.fundacionctic.org/2014/03/mapa-actual-de-las-iniciativas-open-data-en-espana/>
- ³³ Véase la diferencia entre ambos conceptos en Open Data Goldbook for Data Managers, p.13. European Data Portal. <http://www.europeandataportal.eu/sites/default/files/goldbook.pdf>. Este manual contiene recomendaciones para la publicación de datos abiertos, e incluye tanto los requisitos técnicos, como su planificación, estrategia y seguimiento.
- ³⁴ <http://eur-lex.europa.eu/legal-content/ES/TXT/?uri=URISERV%3Aasi0016>
- ³⁵ Directiva 2013/37/UE del Parlamento Europeo y del Consejo, de 26 de junio de 2013, por la que se modifica la Directiva 2003/98/CE relativa a la reutilización de la información del sector público. <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2013:175:0001:0008:ES:PDF>
- ³⁶ <https://ec.europa.eu/digital-single-market/en/europe-2020-strategy>
- ³⁷ <https://ec.europa.eu/digital-single-market/en/our-goals/pillar-ii-interoperability-standards#Latest>
- ³⁸ <http://www.agendadigital.gob.es/Paginas/Index.aspx>
- ³⁹ España. Ministerio de Hacienda y Administraciones Públicas. Informe Anual de la Agenda Digital para España. Julio 2015. <http://www.agendadigital.gob.es/Seguimiento/InformesAnuales/Informes/informe-agenda-digital-espana.pdf>
- ⁴⁰ <http://www.agendadigital.gob.es/planes-actuaciones/Paginas/plan-impulso-contenidos-digitales.aspx>
- ⁴¹ http://www.agendadigital.gob.es/planes-actuaciones/Bibliotecacontenidos/Detalle%20del%20Plan/Plan-ADpE-3_Contentidos.pdf
- ⁴² Real Decreto 1495/2011, de 24 de octubre, por el que se desarrolla la Ley 37/2007, de 16 de noviembre, sobre reutilización de la información del sector público, para el ámbito del sector público estatal. BOE número 269, de 8 de noviembre de 2011. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2011-17560

⁴³ Planes RISP de la administración pública. <http://datos.gob.es/content/planes-risp-de-administracion-publica>

⁴⁴ Biblioteca Nacional de España. Plan de medidas de impulso de la Reutilización de la Información. 2016. http://datos.gob.es/sites/default/files/files/plan_risp_bne.pdf.

⁴⁵ Resolución de 19 de febrero de 2013, de la Secretaría de Estado de Administraciones Públicas, por la que se aprueba la Norma Técnica de Interoperabilidad de Reutilización de recursos de la información. BOE número 54, de 4 de marzo de 2013. https://www.boe.es/diario_boe/txt.php?id=BOE-A-2013-2380

⁴⁶ <http://www.epsiplatform.eu/>

⁴⁷ <http://www.epsiplatform.eu/content/european-psi-scoreboard>

⁴⁸ <http://censoarchivos.mcu.es/CensoGuia/completo.htm>

⁴⁹ La norma Encoded Archival Guide creada originalmente por la Subdirección General de los Archivos Estatales se incluyó, tras distintos proyectos europeos, en el repertorio normativo del Portal de los Archivos de Europa, que dirige la recientemente creada Archives Portal Europe Foundation. <http://wiki.archivesportaleuropefoundation.eu/index.php/Category:Standards>. Véase HERNÁNDEZ, Severiano. El Proyecto APENet : portal de archivos europeos en Internet : creación y mantenimiento. Jornada de Difusión de EuropeaLocal, 17-nov-2010. <http://hdl.handle.net/10421/4766>

⁵⁰ The 5 stars of open linked data. En: Inkdroid, June 4, 2010. <http://inkdroid.org/2010/06/04/the-5-stars-of-open-linked-data/>

⁵¹ <http://id.sgcb.mcu.es/>

⁵² <http://tesauros.mecd.es/tesauros/tesauros>

⁵³ El Directorio Común proporciona un Inventario unificado y común a toda la Administración de las unidades orgánicas / organismos públicos, sus oficinas asociadas y unidades de gestión económica - presupuestaria. http://administracionelectronica.gob.es/ctt/dir3#.Vxunx_19671.

⁵⁴ http://administracionelectronica.gob.es/ctt/dir3/descargas#.Vx3me_19670

⁵⁵ Data Catalog Vocabulary (DCAT): W3C Recommendation 16 January 2014. <https://www.w3.org/TR/vocab-dcat/>. El Catálogo Nacional de Datos puede descargarse en RDF (DCAT) <http://datos.gob.es/sites/default/files/catalogo.rdf>, en ATOM y XSL

⁵⁶ AGENJO BULLÓN, Xavier, HERNÁNDEZ CARRASCAL, Francisca. “Cómo y qué consumir en el mundo Linked Open Data; cómo y qué producir en Linked Open Data”. En Cultura Abierta y Conocimiento Compartido: actas de las XIV Jornadas Españolas de Documentación FESABID ‘15: Gijón, 28, 29 y 30 de mayo de 2015. pp. 169-195. http://www.fesabid.org/sites/default/files/repositorio/actas_fesabid_2015.pdf

⁵⁷ <http://id.loc.gov>

⁵⁸ data.bnf.fr

⁵⁹ http://www.dnb.de/EN/Service/DigitaleDienste/LinkedData/linkedata_node.html

⁶⁰ Linked Open Vocabularies (LOV). <http://lov.okfn.org/dataset/lov/>

⁶¹ <http://www.bne.es/es/Inicio/Perfiles/Bibliotecarios/DatosEnlazados/DescargaFicheros/>

⁶² <http://administracionelectronica.gob.es/ctt/eemgde#.Vx423v19670>

⁶³ <http://www.loc.gov/marc/relators/relaterm.html>

⁶⁴ <http://data.archiveshub.ac.uk/def/>

LA GESTIÓN EFICIENTE TIENE UNA FÓRMULA

Innovación + calidad + agilidad



Calidad

Eficacia, optimización de costes y excelencia en la gestión son nuestras señas de identidad. La trayectoria de Servicio Móvil se fundamenta en la constancia y el trabajo de un equipo de profesionales con una clara vocación de servicio al cliente.

La calidad es la clave para ofrecer soluciones globales a través de la gestión de la información y la logística especializada. Nuestra compañía ha desarrollado su know-how durante años de trabajo continuado con grandes instituciones y corporaciones públicas y privadas, así como pyme en general.

**LOGÍSTICA
HOSPITALARIA**

**DIGITALIZACIÓN
DOCUMENTAL**

**OUTSOURCING
HOSPITALARIO**

M Servicio Móvil

www.serviciomovil.com

Atención al cliente:
T. 902 121 181
E. info@serviciomovil.com

Visítenos en el Stand 1-2