

El papel de los metadatos en la preservación digital

Por Jorge Candás Romero

Resumen: *El problema de la preservación de la información digital está muy presente en nuestros días. A medida que el volumen de información en este formato aumente, irá creciendo la preocupación por lograr que ésta se mantenga accesible a largo plazo. Los metadatos, como elementos para la descripción de los documentos digitales o DLO (document like object), han de cubrir un papel fundamental a la hora de aportar información sobre las características lógicas de los DLO a que se refieren, y por tanto para su correcta preservación. Asimismo, al ser los metadatos información digital, su preservación ha de ser también tenida en cuenta. A este respecto se han desarrollado programas y estudios relacionados con la utilización de metadatos para la preservación digital y con la propia preservación de metadatos.*

Palabras clave: *Metadatos, Información digital, Preservación, Preservación digital.*

Title: The role of metadata in digital preservation

Abstract: *The problem of preserving digital information has now taken on great importance. As the amount of digital information continues to grow, the concern about maintaining access to this information accessible will increase. As elements for describing digital documents or DLO (Document Like Object), metadata play an important role in providing information about the logical aspects of the DLO they refer to, and consequently, in all the tasks related to preservation. At the same time, metadata are digital information, so their preservation must be borne in mind. In this respect, several projects and work groups have been devoted to studying the use of metadata for digital preservation and the preservation of metadata themselves.*

Keywords: *Metadata, Digital information, Preservation, Digital preservation.*



Jorge Candás Romero, diplomado en biblioteconomía y documentación (V promoción 2000–2003) y licenciado en documentación (I promoción 2003–2005) por la Universidad de A Coruña. Actualmente realizando el doctorado en documentación en la Universidad Carlos III de Madrid.

Content+Tools+Services



Ovid

a Wolters Kluwer business

Precision search and discovery—it's more than a simple search box

How can I find what I'm looking for?

Where do I start my search?

How do I get to full text fast?

Sound familiar?

You need to present important scholarly information to your patrons—whether they are novice or expert searchers. At Ovid, we offer a complete e-resource management solution that connects them with exactly what they're looking for fast. Ovid SearchSolver™ starts the information discovery process. Precision search tools drill down further, refining results. And Ovid LinkSolver™ takes them directly to full text or other relevant online resources.

Visit www.ovid.com/precisionsearch today to learn more about Ovid's innovative e-resource management tools.

Contact Ovid: +34 91 418 62 75
spain@ovid.com www.ovid.com

1. Preservación digital

PUEDE DEFINIRSE COMO EL CONJUNTO de actuaciones relativas a la preservación de los documentos existentes en formato digital. Así, si la entendemos como un conjunto de actividades, entre las que se incluye la conservación, y que están destinadas a que un objeto perdure el mayor tiempo posible en su estado original, la preservación digital supondría llevar a cabo estas tareas sobre información en formato digital, preocupándonos no sólo por el mantenimiento del objeto, sino también (y principalmente) por su contenido informativo.

Esta idea es expresada por **Rotenberg**¹ al decir, “la preservación de documentos digitales es análoga a la de textos antiguos. Al igual que con los documentos digitales, a veces es necesario actualizar un texto antiguo transcribiéndolo [...]. Un texto antiguo puede ser preservado de dos modos: bien copiándolo en su idioma original o bien traduciendo a cualquier otro idioma contemporáneo al momento de la transcripción”. Por tanto, la planificación y desarrollo teórico de las tareas de preservación será igual tratándose de información digital o sobre cualquier otro objeto de información.

El aumento de la producción de información en formato digital, el crecimiento del número de programas para la digitalización de documentos y el valor de esta información, hacen imprescindible el desarrollo de una política de actuación específica sobre estos materiales en todas las organizaciones. Proyectos como el *Cedars*² muestran esta preocupación por mantener accesibles los recursos digitales más allá del corto plazo.

Las propuestas actuales se centran en su mayoría en la información de carácter administrativo, aunque deberían hacerse extensi-

bles a todas las expresiones culturales, desde imágenes, periódicos, libros, películas e incluso videojuegos, para perpetuar así una gran parte del acervo cultural de los últimos años.

«La planificación y desarrollo teórico de las tareas de preservación es igual cuando se trata de información digital que de cualquier otro objeto de información»

De acuerdo con la *Cornell University Library*³ la preservación digital tiene como objetivo “mantener la capacidad de visualizar, recuperar y utilizar colecciones digitales frente a las infraestructuras y elementos tecnológicos y de organización que cambian con mucha rapidez”. De esta definición cabe señalar dos ideas fundamentales, por un lado la necesidad de prestar atención a los elementos tecnológicos asociados a la información digital, y por otro la rapidez con que se producen los cambios en este entorno.

La dependencia tecnológica, es decir, la necesidad de un entorno tecnológico concreto para acceder a la información, aun siendo un factor exógeno a la información y al propio soporte, es la característica básica y definitoria de la información digital y la que la distingue de la encontrada en soporte tradicional, a la cual el acceso se hace directamente.

De acuerdo con esto se puede decir que la preservación digital ha de basarse en la de los soportes de almacenamiento y en la actuación que evite que el avance tecnológico deje inaccesible la información digital, ya sea mediante migración, emulación, conservación de la tecnología, u otros métodos. Todo ello bajo tres puntos de garantía que aseguren: la originalidad (el docu-

mento primario), la autenticidad (autoría) e integridad (estar completo) del objeto de información o *DLO* (*document like object*) y de la información misma. Estas cuestiones suponen una gran preocupación debido a la fácil manipulación de gran parte de los documentos digitales.

Todas las definiciones consultadas consideran que la preservación digital ha de desarrollarse como una tarea global, continua y compleja, en la que han de tenerse en cuenta los factores físicos y lógicos de la información, así como un tratamiento de descripción formal correcto y adecuado. De ahí se deriva la importancia de imbricar las tareas de preservación digital en el conjunto de actuaciones desarrolladas en la gestión de cualquier centro de información.

La evolución de la tecnología puede darse tanto en un cambio de hardware (componentes físicos) como de software (elementos lógicos). El problema más común en la preservación digital se da en este último caso, en la evolución de la estructura lógica (el formato del archivo y/o los programas de lectura y escritura de documentos), lo que dificulta la lectura de los *DLO* en su forma original. Es también entonces cuando el problema y la actuación son más complejos, ya que las manipulaciones habrán de realizarse, en el caso de optar por la migración, sobre el propio documento y sus contenidos, siendo mucho mayor el riesgo de corromper la originalidad, integridad y autenticidad de la información. Para el tratamiento de estos cambios existen distintas actuaciones:

a. Migración. Se busca evitar la obsolescencia del hardware y/o de los formatos de archivo (estructura lógica, el modo en que se codifica la información) mediante su transferencia a otro hardware y/o su conversión a otros formatos de archivo o nuevas versiones de los mismos.

Todo ello con el fin de garantizar (temporal o definitivamente) su acceso. Por tanto es recomendable realizar la migración hacia formatos estándar (de iure o de facto), procurando que sea un proceso reversible y sin pérdida de información.

Implica la manipulación directa de la cadena de bits (*bit stream*), lo que puede afectar a la autenticidad, integridad y originalidad del documento, y también de los metadatos embebidos (si se opta por esta disposición de los mismos). Por ello los estándares elegidos para la conservación y los procesos de migración habrían de permitir también la de los propios metadatos a otros estándares. En este sentido, la *IFLA* defiende la utilización de formatos estándar en *DLO* y metadatos al decir⁴ “el formato de los archivos para las imágenes máster debe además ser interoperable, como tiff y jpeg; y el esquema de metadatos no debe requerir de un software específico para su lectura, como texto ascii expresado en xml”.

b. Emulación. Es un método de preservación que no requiere la manipulación del fichero original. Supone la implementación de una máquina virtual, o simulación de un entorno informático en una máquina, para permitir la visualización de documentos originalmente ilegibles en esa máquina. Sistemas operativos actuales y algunos paquetes ofimáticos ya incluyen este tipo de opciones por defecto a un nivel muy básico.

c. Preservación de la tecnología. Para la *Cornell University Library*³ significa “preservar el entorno técnico que ejecuta el sistema, incluyendo software y hardware, como por ejemplo: sistemas operativos, software de aplicación original, unidades de medios y similares”. Este método supone la conservación de la información en el soporte original y la de la tecnología (hardware y software) necesaria para su acceso.

2. Los metadatos en la preservación digital

Una correcta identificación y descripción de los documentos provee de la información necesaria para el tratamiento de los mismos. Al hablar de documentos en formato digital o *DLO*, son los metadatos los que cumplen esta función. De acuerdo con gran parte de las propuestas actuales, también deben contener información para el desarrollo adecuado de actuaciones de preservación al aportar información sobre el tipo de formato, modificaciones realizadas sobre el mismo y otras informaciones relevantes.

Méndez⁵ los define como “elementos o estructuras de organización de la información que, asignados a cada objeto de información electrónica, la clasifican, categorizan o describen. Son, en definitiva, datos sobre datos, información estructurada sobre la información distribuida”. Además indica como papales emergentes de los metadatos la actualización de la información y la preservación.

«Deben contener información para el desarrollo correcto de actuaciones de preservación»

Day⁶ se pronuncia en este sentido y mantiene que los metadatos de conservación se han convertido en un campo de trabajo muy popular, que “archiveros y gestores de documentación se han centrado en el desarrollo de *recordkeeping metadata*, mientras que otros grupos se han centrado en la definición de especificaciones de metadatos para necesidades concretas”.

Recordkeeping metadata, o también *record management metadata*, podría traducirse por «metadatos para la gestión de documentos».

El *CIA* (*Consejo Internacional de Archivos*)⁷ igualmente les reconoce una función principal al hablar de los fines de la preservación digital, esto es, conseguir que los documentos sean y permanezcan “auténticos, completos, accesibles y comprensibles, procesables, y potencialmente reusables”, indica que la preservación correcta estará asegurada, entre otras cuestiones si: “los metadatos indican el nombre y la versión del formato de cada archivo y la identificación del software con que ha sido creado y con que se ha realizado la última modificación”. Además plantea tres tipos y funciones con respecto a tareas de preservación:

«a. *Recordkeeping metadata* (metadatos descriptivos) son creados con los propios documentos o en la misma organización que los ha creado. Deben recoger autor, fecha de creación, título y palabras clave.

b. *Archival metadata* (metadatos de archivo), añadidos para facilitar la gestión de los documentos tras su creación, incluyen elementos como la última revisión, fecha en que pasó al archivo, condiciones de uso, etc.

c. *Technical metadata* (metadatos técnicos), necesarios para comprender y procesar los documentos, pueden ser, por ej., formato de archivo y fecha de la última migración de formato».

Siguiendo esta distinción podemos ver que cumplen un doble papel, como elementos de descripción y como herramientas de gestión administrativa y archivística. Es decir, no sólo cumplen una función de representación del documento (*surrogate*), sino también ofrecen información sobre cómo gestionarlo y su trayectoria documental. En este sentido, otra aplicación de los metadatos habrá de ser la documentación de los cambios y procesos a que se sometán los documentos en

las migraciones de formato (*technical metadata*).

Premis (Preservation metadata: implementation strategies) define⁸ *preservation metadata* como “la información que un repositorio usa para apoyar el proceso de preservación digital. Concretamente, el conjunto de metadatos encargado de gestionar las funciones de mantenimiento de la viabilidad, legibilidad y accesibilidad, autenticidad e identificación en el contexto de preservación”. Asimismo indica que esta categoría recogerá información antes asignada a otras, como metadatos administrativos (incluyendo derechos y permisos) y técnicos y estructurales. También considera de gran importancia la documentación de la procedencia digital (la historia del objeto) y la documentación de las relaciones existentes entre los diferentes objetos y de éstos con el repositorio.

Para la disposición de los metadatos y su tratamiento, se plantean dos posibilidades: embeber esa información en el propio documento a que se refieren, o bien crear un documento independiente para el almacenamiento de los metadatos vinculado al referido. Esta problemática es de suma importancia pues, los metadatos, como información digital que son, están sujetos a la misma fragilidad e inestabilidad que los *DLO* a los que se refieren, por lo que su tratamiento ha de tenerse muy en cuenta, y también han de cumplir los mismos requisitos técnicos que los *DLO* para garantizar así su uso a largo plazo.

Al almacenarlos embebidos en el mismo fichero que el documento a que se refieren, cualquier modificación de ellos modificaría la cadena de bits del documento (*bit stream*¹), y este cambio podría cuestionar la integridad y autenticidad de éste. Mientras, la separación y crea-

ción de ficheros independientes puede dificultar la gestión de los fondos y perderse la vinculación entre ambas informaciones.

La *Koninklijke Bibliotheek (Biblioteca Nacional de Holanda)*⁹, en su proyecto de creación de depósitos digitales, opta por la separación con el objetivo de no acceder ni modificar el documento original cuando se trabaja con los metadatos, y dice “la separación del contenido informativo y la información de descripción y preservación está basada en razones de coherencia funcional y flexibilidad. Separando el objeto digital de los metadatos técnicos, se permite al *DIAS (digital information archiving system)* actualizar y gestionar los metadatos técnicos sin modificar el objeto digital archivado”.

En la misma línea de actuación se pronuncian **Gail Hodge** y **Bonnie C. Carroll**¹⁰, aunque también definen una serie de casos en los que, para un mejor funcionamiento y utilización de los *DLO*, los metadatos se encontrarán embebidos en el propio documento: “en general, los archivos de metadatos se almacenan separados de los propios archivos [...]. De cualquier modo, en algunas circunstancias, como diarios electrónicos con cabeceras sgml, la información debe almacenarse en el propio archivo y extraerse para el catálogo. En el caso de archivos distribuidos, los metadatos deben almacenarse centralmente con los recursos electrónicos distribuidos. Para algunas herramientas de búsqueda utilizadas los metadatos deberán almacenarse como etiquetas embebidas en el recurso en línea”.

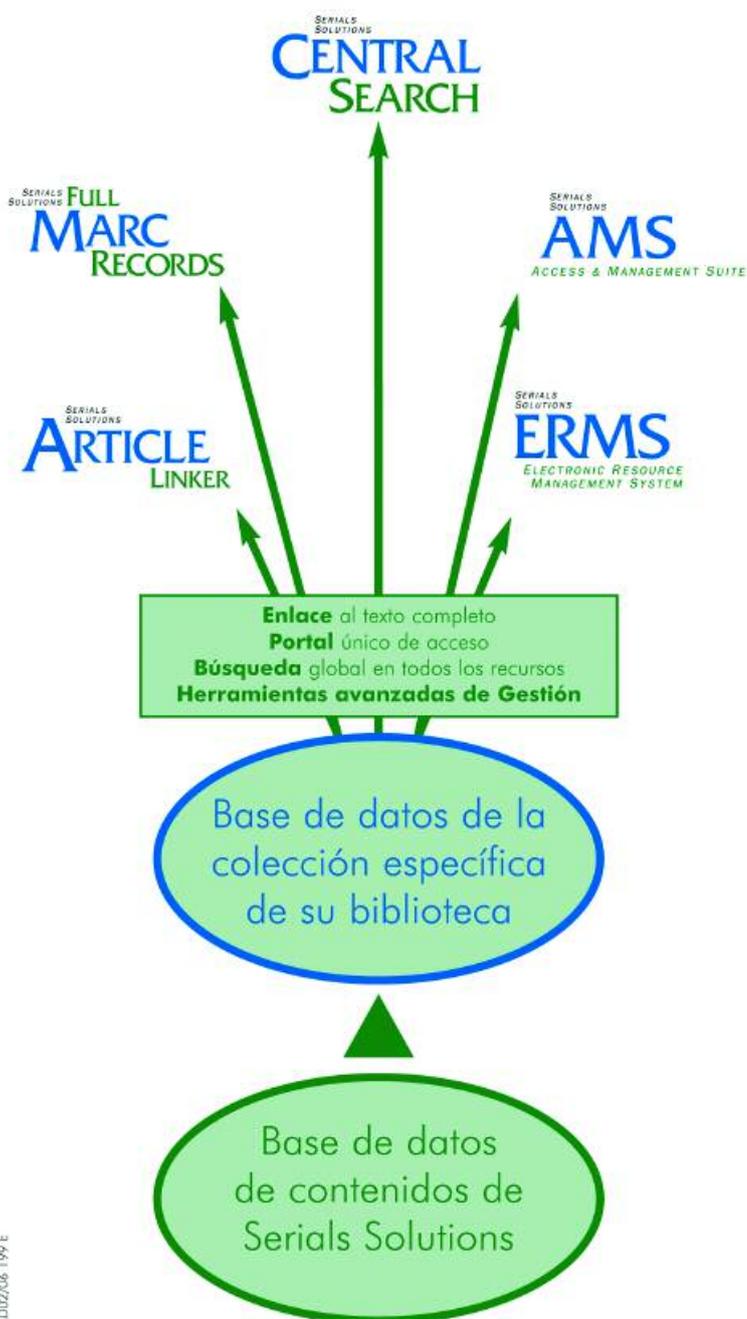
Igualmente, estos autores atribuyen a los metadatos un papel fundamental para facilitar la interoperabilidad de los sistemas de información mediante su estandarización, y por tanto la posibilidad de compartir información a gran escala.

<i>Premis</i>	<i>Cedars</i>	<i>LoC</i>
"format" (identificación del formato informático del <i>DLO</i>)		"associated_file_type" (formato del archivo informático)
"creatingApplication" (aplicación con la que se creó el <i>DLO</i>)		
"preservationLevel" (identificación de las tareas de preservación que han de aplicarse)	"preservation description information" (la información necesaria para una correcta preservación),	"reformatting_guidelines" (políticas e instrucciones para el cambio de formato)
"objectIdentifier" (la designación que identifica unívocamente al <i>DLO</i>)	"reference information" (los identificadores del recurso)	"archive_ID" (identificador de la localización del <i>DLO</i>)
"relationship" (relaciones existentes entre el <i>DLO</i> y otro u otros <i>DLO</i>)	"context information" (relación del <i>DLO</i> con su contexto)	
"fixity" (información sobre si el <i>DLO</i> ha sido alterado de forma no autorizada o no registrada)	"provenance information" (la historia del <i>DLO</i>),	"archive_history" (historia de hechos referentes al archivo del <i>DLO</i>)
"objectCharacteristics" (características técnicas del <i>DLO</i>)	"custody history" (datos de las organizaciones que han sido depositarias)	"reformatted_original_information" (información sobre las condiciones del documento original)
	"management history" (modificaciones realizadas en el <i>DLO</i> para su preservación)	"archive_date_time" (fecha y hora de creación, copia de seguridad, etc.)
	"rights management" (información sobre los derechos de propiedad intelectual relativos al <i>DLO</i>)	

Comparativa de los principales elementos de los esquemas de metadatos



Las soluciones más
actualizadas y completas
para sus colecciones
electrónicas



www.serialssolutions.com

ProQuest Information and Learning
C/ Castello nº 95 - 6º D
28006 - Madrid
Spain
Tel: +34 91 575 5597
Fax: +34 91 575 9885
E-mail: editor@iberica.proquest.com

SerialsSolutions

Serials Solutions es una empresa afiliada a ProQuest Information and Learning

3. Proyectos actuales

La necesidad de incluir en los metadatos información sobre el tratamiento que han de recibir los *DLO* para garantizar su correcta preservación queda patente en el hecho de que gran parte de los proyectos sobre este asunto y de los modelos de esquemas de metadatos recogen esta opción. A continuación se analizan brevemente algunos de los más destacados.

3.1. Dublin Core Metadata Initiative Preservation Working Group

Establecido en 2003, funciona como un grupo de discusión y debate para personas y organizaciones interesadas o que trabajen sobre metadatos para la preservación. Sus objetivos¹¹ son:

- a. Recopilar información y hacer un estudio de los esquemas de metadatos existentes.
- b. Investigar la necesidad de un campo específico para el esquema de metadatos de preservación (p. ej. archivos, bibliotecas, información científica, etc.).
- c. Ponerse en contacto con otros proyectos globales que tengan el mismo propósito como *Premis*, si es necesario.

Los datos recogidos en el modelo de Dublin Core se agrupan en tres tipos⁵:

- a. Principalmente relacionados con el contenido del *DLO*.
- b. Relacionados con la propiedad intelectual del recurso.
- c. Vinculados a la temporalidad del documento: fecha, tipo, formato e identificador de recurso.

Este último grupo de elementos es el que nos ocupa, el que debe contener las instrucciones necesarias para el tratamiento correcto del *DLO* (programa/s utilizado/s para su creación, para su lectura, para su modificación, etc.) y los que con-

tingan el historial de los posibles cambios sufridos (de formato, de soporte, etc.).

3.2. Premis Working Group

Patrocinado por *OCLC* y *RLG* (*Research Libraries Group*) está compuesto por expertos internacionales, incluyendo bibliotecas, museos, archivos, administración pública y sector privado.

«Otra función de los metadatos habrá de ser la documentación de los cambios y procesos a que se sometan los documentos en las migraciones de formato»

Sus objetivos son dos¹²:

- a. Desarrollar el núcleo para el conjunto de metadatos de conservación, apoyado por un diccionario de datos con una amplia aplicabilidad por la comunidad interesada en este tema.
- b. Identificar y evaluar estrategias para la codificación, almacenamiento y gestión de metadatos en sistemas de preservación digital.

Desde mayo de 2005 están disponibles las conclusiones y resultados de su trabajo en el *Data dictionary for preservation metadata: final report of the Premis working group*¹³. Este documento “define y describe un núcleo de metadatos de preservación con amplia aplicabilidad en repositorios, contextualiza el *Data dictionary*, explica las suposiciones menos evidentes y el modelo de datos, y también aclara el significado y uso de las unidades semánticas definidas en el *Data Dictionary*”, las cuales describen las propiedades de los documentos o, como se denominan en la propuesta, entidades.

El informe presenta y analiza un esquema de metadatos, sus dife-

rentes elementos, así como indicaciones de uso y ejemplos de aplicación a distintos tipos de *DLO* (documentos de *MS Word*, páginas web, fotografías, etc.). Trata de presentarse como una base sobre la cual desarrollar futuros esquemas.

De cada una de las unidades semánticas se definen los siguientes componentes:

—Nombre (*name*): usado para identificar los componentes de manera única, pero no utilizados para el funcionamiento interno.

—Componentes semánticos (*semantic components*): cada una de las partes de las unidades semánticas que no tienen valor por sí mismas.

—Definición (*definition*): significado de la unidad semántica.

—Fundamento (*rationale*): por qué es necesaria.

—Límite de los datos (*data constraint*): cómo debe codificarse.

—Categoría del objeto (*object category*): indica si la unidad semántica se refiere a una representación, documento o una cadena de bits (*object bitstream*).

—Aplicabilidad (*applicability*): indicación de a qué categoría de objetos se aplica.

—Ejemplos (*examples*): casos ilustrativos de uso.

—Repetición (*repeatability*): indicación de si la unidad semántica puede tener múltiples valores.

—Obligatoriedad (*obligation*) a la hora de incluir un valor en dicha unidad semántica.

—Notas de creación/mantenimiento (*creation/maintenance notes*): cómo debe elaborarse y/o actualizarse.

—Notas de uso (*usage notes*): información sobre el uso de la unidad semántica o aclaraciones sobre su definición.

Algunas de las unidades semánticas definidas son: *objectIdentifier* (designación unívoca del *DLO*), *preservationLevel* (tareas de preservación que han de aplicarse), *objectCharacteristics* (características técnicas del *DLO*), *fixity* (sobre si ha sido alterado de forma no autorizada o no registrada), *format* (formato informático), *inhibitors* (características técnicas del objeto que limitan el uso, acceso o migración), *creatingApplication* (aplicación con la que se creó), *software* (el necesario para acceder), *hardware* (el necesario para acceder), *relationship* (relaciones existentes entre el *DLO* y otro u otros *DLO*), etc.

3.3. The Cedars Project². Metadata for digital preservation

Fue creado en la tercera fase del programa de bibliotecas digitales *eLib* en abril de 1998, con una duración prevista de tres años. Su principal objetivo era “dirigir cuestiones estratégicas, metodológicas y experimentales y aconsejar sobre las mejores prácticas para la preservación digital”. Un segundo objetivo que se extendió durante un año (posterior a los tres iniciales) se centraba en “consolidar y compartir información sobre el trabajo realizado los tres años anteriores”.

El trabajo práctico del proyecto se centró en la creación de un prototipo de archivo distribuido (*distributed archive prototype*) basado en el sistema *OAIS* (*Open archival information system*) (del que se hablará más adelante) empleando una especificación de metadatos (*metadata specification*).

En la presentación de esta propuesta¹⁴ de esquema de metadatos se justifica la utilidad y necesidad del mismo al afirmar que “el acceso significativo a objetos digitales durante décadas o incluso siglos precisará de un sofisticado sistema de metadatos que describa detalladamente procesos técnicos, de gestión y de actividades administrativas,

así como de control de acceso y copyright”.

El texto de la propuesta, titulada *Metadata for digital preservation. The Cedars project outline specification. Draft for public consultation*¹⁵, también incide en estos aspectos en varias ocasiones al decir “estos metadatos son necesarios para dar acceso significativo al contenido digital archivado, e incluye información descriptiva, administrativa, técnica y legal” y “la segunda razón para desarrollar esta especificación provisional es contribuir, en un nivel estratégico, al desarrollo en colaboración e internacionalmente de una especificación estándar de metadatos para la preservación digital”.

«No sólo se preocupa de la utilización de los metadatos, sino también de la correcta manipulación y preservación de éstos»

Como problema de la aplicación de metadatos se menciona el nivel de detalle de la descripción (*granularity*), problema intrínseco a cualquier proceso de análisis documental ya sea formal o de contenido, tanto en fondos bibliográficos como archivísticos. En este sentido se indica: “[los metadatos] deben asignarse a nivel de colección digital, de objeto digital o incluso a nivel de archivo individual. La granularidad de los metadatos la determinará el propio objeto digital y el nivel de descripción necesario para asegurar la preservación, pero también dependerá de la política en la gestión de la colección”. Frente a ello, el *Cedars* optó por desarrollar un esquema que busca permitir la descripción a cualquier nivel dejando la decisión de especificar el nivel al centro de información.

Lo que sí se determinan en la propuesta del *Cedars* son tres nive-

les de metadatos de acuerdo con la significación o importancia que tienen los elementos de descripción. Así, se especifican elementos:

- a. Muy significativos: muy útiles para la presentación en un amplio abanico de objetos digitales y, por tanto, necesarios si están disponibles.
- b. Significativos: útiles para la mayor parte de los objetos digitales.
- c. Poco significativos: se consideran útiles pero no estrictamente necesarios.

Entre los primeros estarían: *preservation description information* (información necesaria para una correcta preservación), *reference information* (identificadores del recurso), *context information* (relación del *DLO* con su contexto), *provenance information* (historia del *DLO*), *custody history* (datos de las organizaciones que han sido depositarias), *management history* (modificaciones realizadas en el *DLO* para su preservación), *rights management* (derechos de propiedad intelectual relativos al *DLO*), etc.

Y entre los «significativos» estarían: *existing metadata* (identificación de otros metadatos referentes al *DLO*), *reason for creation* (motivos para la creación del *DLO*, sea digitalización para conservación u otros), *original technical environments* (información del hardware y software utilizado en la creación), *prerequisites* (hardware y software necesarios originalmente para el acceso al documento), *reason for preservation* (motivos para su preservación y archivo), etc.

Para la estructura y los apartados, la propuesta del *Cedars*¹⁶ sigue en gran medida el modelo *OAIS*, al que se refiere al indicar el tipo de información que recoge cada elemento de descripción. Estos elementos son:

a. Información de referencia (*reference information*). Identifica y, si es necesario, describe uno o más mecanismos usados para asignar identificadores tanto para la *content information* como para permitir a los sistemas externos la referencia unívoca a una *content information* concreta.

b. Información de contexto (*context information*). Documenta las relaciones de la *content information* con su entorno y con otras *content information* existentes en otros lugares.

c. Información de procedencia (*provenance information*). Documenta la historia de la *content information*.

d. Información de estabilidad (*fixity information*). Recoge los mecanismos de autenticación y muestra las claves de autenticación usadas para asegurar que un objeto concreto de *content information* no ha sido alterado de un modo no documentado.

e. Información de representación (*representation information*). Es necesaria para hacer comprensible el *data object*.

f. Derechos de propiedad intelectual y uso (*handling intellectual property rights*). Son considerados parte de la *provenance information*. El trabajo en este campo supone un aumento significativo del modelo de *OAIS* y está pensado para animar a los propietarios del *copyright* de contenidos digitales que no están convencidos de si archivar o no sus recursos.

En este caso son los elementos correspondientes a *fixity information* y a *representation information* los que recogen la información necesaria para la preservación y tratamiento de los *DLO*.

3.4. National Library of New Zealand

Ha desarrollado el *Metadata Standards Framework–Preservation Metadata*¹⁷.

Entre los objetivos de este esquema están “detallar los elementos necesarios para respaldar la preservación de objetos digitales y será la base para el diseño de un repositorio y sistemas de entrada de datos para recoger y almacenar metadatos de preservación”. Así pues, no sólo se preocupa de la utilización de los metadatos, sino también de su correcta manipulación y preservación.

3.5. Library of Congress

Ha creado un esquema propio¹⁸ para la descripción de repositorios digitales agrupados en tres tipos: descriptivos, administrativos y estructurales. Siendo los administrativos los usados para gestionar y preservar los *DLO*.

Este esquema está basado en el *Structural metadata dictionary for LC digitized material, version 1.03*¹⁹, la *Metadata table for the Colidge–consumerism experiment* y la experiencia del *Thomson Editorial Asset Management System (Teams) Repository*.

Entre las funciones que deben cumplir los metadatos definidos se incluyen la información para preservación digital y aquella para el cambio de formato como medida de preservación de los *DLO*. Éstos se denominan *preservation (D)* y *preservation (R)* respectivamente²⁰.

Entre los primeros están: *archive_date_time* (fecha y hora de creación, copia de seguridad, etc.), *archive_history* (historia de hechos referentes al archivo del *DLO*), *archive_ID* (identificador de la localización del *DLO*), *archive_next_date_time* (fecha y hora de creación de la última copia de seguridad), *associated_file_type* (formato del archivo informático), etc.

Y entre los segundos: *reformatted_original_information* (condiciones del documento original), *re-*

formatting_guidelines (políticas e instrucciones para el cambio de formato), etc.

4. Conclusiones

La preservación de la información digital es, y seguirá siendo, un tema de preocupación y de investigación, debido al creciente aumento de la información en este formato. Como hemos visto, son los metadatos, y más concretamente los de preservación (*recordkeeping metadata* o *preservation metadata*), los elementos que aportan la información necesaria para un adecuado tratamiento de los *DLO* que garantice su preservación.

En este ámbito se ha de trabajar conjuntamente desde la informática y la biblioteconomía y la documentación para coordinar así los aspectos técnicos con una correcta descripción y tratamiento de los *DLO* como objetos de información, así, esta colaboración habrá de desarrollarse tanto en los aspectos teóricos como posteriormente en el trabajo de asignación de metadatos. Es de esperar que en los múltiples lugares comunes existentes entre ambas disciplinas se empiece a dar en España una mayor colaboración y comunicación científica. Para ello, los aspectos relativos a metadatos son un buen punto de partida.

El reto de los metadatos está ahora mismo en la definición de un esquema de descripción aceptado y aplicado por el mayor número de partes posibles. La diversidad de materiales digitales dificultará la creación de un modelo estándar para la descripción de los *DLO*, si bien, con los metadatos de preservación, por tratarse de información de carácter técnico, sí resultaría más plausible la determinación de unos elementos comunes.

Las propuestas realizadas por el *Cedars*, el esquema *Premis* o el utilizado por la *Library of Congress* son un buen comienzo. Sin embargo, habrá de ser una entidad de ca-





baratz

En primera línea en
Sistemas de Información y Gestión del Conocimiento



Raimundo Fernández Villaverde, 28 28003 Madrid (España) Teléfono +34 91 456 03 60 - Fax +34 91 533 09 58 www.baratz.es - E-mail: informa@baratz.es

- © Soluciones para bibliotecas:
Absys, **absysNET**, **Absys express**
- © Gestión documental y del conocimiento:
BKM, **Baratz Windows**, **BRSCGI**
- © Soluciones para archivos: **Albalá**
- © Servicios de Catalogación Retrospectiva
- © Edición en CD-Rom



UNE-EN ISO 9001:2000

rácter supranacional de prestigio quien determine el estándar de iure a seguir, para facilitar así el intercambio de información. La viabilidad del desarrollo del estándar se ve en el hecho de que, como ya se ha comentado, los esquemas actualmente utilizados tienen un gran número de aspectos comunes.

Referencias bibliográficas

1. **Rothenberg, Jeff.** *Ensuring the longevity of digital information.* Consultado en: 09-04-05. <http://www.clir.org/pubs/archives/ensuring.pdf>
2. **Cedars.** *Project overview.* Consultado en: 31-07-05. <http://www.leeds.ac.uk/cedars/overview/overview.html>
3. **Biblioteca de la Universidad de Cornell.** *Departamento de Investigación. Tutorial de digitalización de imágenes—Preservación digital.* Consultado en: 06-04-05. <http://www.library.cornell.edu/preservation/tutorial-spanish/preservation/preservation-01.html>
4. **IFLA.** *Guidelines for digitization projects for collections and holdings in the public domain, particularly those held by libraries and archives.* Consultado en: 10-04-05. <http://www.ifla.org/VII/s19/pubs/digit-guide.pdf>
5. **Méndez Rodríguez, Eva M^a.** *Metadatos y recuperación de información: estándares, problemas y aplicabilidad en bibliotecas digitales.* Gijón: Trea, 2002, Isbn 84-9704-055-4.
6. **Day, Michael.** *Metadata for digital preservation: an update.* Consultado en: 28-04-05. <http://www.ariadne.ac.uk/issue22/metadata/>
7. **International Council on Archives. Committee on Current Records in an Electronic Environment.** *Electronic records: a workbook for archivists (ICA Study 16).* Consultado en: 22-11-05. http://www.ica.org/biblio/Study16ENG_5_2.pdf
8. **OCLC; RLG; Preservation metadata: implementation strategies (Premis).** *Data dictionary for preservation metadata: final report of the Premis working group.* Consultado en: 12-08-05. <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>
9. **Van Diessen, Raymond J.** *Preservation requirements in a deposit system.* Consultado en: 09-04-05. http://www.kb.nl/hrd/dd/dd_onderzoek/reports/3-preservation.pdf
10. **Hodge, Gail; Carroll, Bonnie C.** *Digital electronic archiving: the state of the art and the state of the practice.* Consultado en: 28-04-05. http://www.icsti.org/Dig_Archiving_Report_1999.pdf
11. **Dublin Core Metadata Initiative. DCMI Preservation Working Group.** Consultado en: 26-04-05. <http://dublincore.org/groups/preservation/>
12. **OCLC.** *Background [OCLC-Premis: Preservation Metadata Implementation Strategies].* Consultado en: 12-08-05. <http://www.oclc.org/research/projects/pmwg/background.htm>
13. **OCLC; RLG; Preservation Metadata: Implementation Strategies (Premis).** *Data dictionary for preservation metadata: final report of the PREMIS working group.* Consultado en: 12-08-05. <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>
14. **Cedars.** *Metadata for digital preservation.* Consultado en: 27-04-05. <http://www.leeds.ac.uk/cedars/metadata.html>
15. **Cedars.** *Metadata for digital preservation. The Cedars project outline specification. Draft for public consultation.* Consultado en: 27-04-05. <http://www.leeds.ac.uk/cedars/documents/Metadata/cedars.html>
16. **The Cedars project team; Ukoln.** *Metadata for digital preservation: the Cedars project outline specification.* Consultado en: 27-04-05. <http://www.leeds.ac.uk/cedars/MD-STR-5.pdf>
17. **National Library of New Zealand.** *Metadata standards framework: preservation metadata.* Consultado en: 21-04-05. http://www.natlib.govt.nz/files/4initiatives_metastchema.pdf
18. **Library of Congress.** *Introduction to metadata elements: Library of Congress.* Consultado en: 29-04-05. <http://www.loc.gov/standards/metadata.html>
19. **Library of Congress.** *Structural metadata dictionary for LC digitized material.* Consultado en: 29-04-05. <http://memory.loc.gov/ammem/techdocs/repository/structmeta.html>
20. **Library of Congress.** *Table of core metadata elements for Library of Congress Digital Repository Development.* Consultado en: 12-09-05. <http://www.loc.gov/standards/metatable.html>

Jorge Candás Romero.
jorgecandas@gmail.com