

***Exegi monumentum aere perennius* (Hor. C. 3.30): preservación a largo plazo de materiales bibliográficos digitalizados**

Xavier Agenjo Bullón

Fundación Ignacio Larramendi

xavier.agenjo@larramendi.es

Francisca Hernández Carrascal

Digibis, Producciones digitales

francisca.hernandez@digibis.com

Resumen

Se analiza la preservación digital de los materiales bibliográficos desde el punto de vista de los importantes proyectos de digitalización que se han llevado a cabo en los últimos años en España y se deduce de ello la necesidad de convertir una preocupación de carácter teórico en una necesidad práctica, ya verdaderamente crítica dado el volumen de material digitalizado. Se examinan las dos principales normas mediante las cuales la comunidad internacional está desarrollando políticas de preservación digital a largo plazo, la normativa *ISO 14.721*, *OAIS*, y el *Diccionario de Datos* elaborado por la *Estrategia de Implementación de Metadatos de Preservación (PREMIS)*, que han alcanzado ya un importante reconocimiento internacional. Se citan las recomendaciones y buenas prácticas que deben utilizarse para la preservación digital, desde el encapsulamiento de los datos de PREMIS en el esquema de Metadatos METS hasta la ingesta de los mismos en un repositorio de preservación digital OAIS. Y por último se hace una recomendación para la incorporación de metadatos de preservación a los objetos digitales desde el mismo momento de su creación, incluso en un entorno en el que aún no estén definidos repositorios de preservación.

1. Introducción

El objetivo de esta comunicación es realizar un acercamiento a la preservación digital desde la particular perspectiva y experiencia en el desarrollo de proyectos de digitalización. Así, aunque las recomendaciones y prácticas que se mencionarán son pautas unificadas para cualquier procedimiento, nos centraremos en las fases de digitalización y construcción de bibliotecas digitales o virtuales.

Desde una perspectiva nítidamente española, los materiales bibliográficos han comenzado a digitalizarse muy recientemente. Si bien pueden citarse algunos proyectos notables desde los años 90¹, como los del Archivo General de Indias, Biblioteca Virtual Miguel de Cervantes, etc., la mayor parte se inician en el presente siglo. La mejor herramienta de la que se dispone para tener una idea exacta de la digitalización de materiales bibliográficos en España es el *Directorio y Recolector de Recursos Digitales*² del Ministerio de Cultura. Quizá no sea completamente casual que la presentación³ de este *Directorio y Recolector* se realizara en el marco de las *Primeras Jornadas de Preservación Digital: Conceptos Básicos y Principales Iniciativas*⁴ celebradas los días 14 a 16 de marzo de 2006 en el salón de actos del Ministerio de Cultura. Este directorio está basado en tres aplicaciones distintas: un recolector de recursos digitales, un directorio, basado en el *MARC 21 Format for Community Information*, actualizable vía web y una aplicación de búsqueda y recuperación de la información, tanto de recursos digitales como de descripciones de colecciones digitales.

En el momento de su inauguración el *Directorio* describía 45 proyectos de digitalización o de creación de bibliotecas digitales, virtuales, etc., mientras que en la actualidad ofrece 194 colecciones. A fecha de 31 de diciembre de 2009, el *Recolector* permite recuperar cerca de 550.000 documentos digitales⁵ procedentes de 84 repositorios (18 de los cuales se corresponden con proyectos de digitalización⁶, 28.192 documentos, y 66 son repositorios institucionales).

¹ Hay una relación completa en AGENJO, Xavier. La cuarta salida del monje digital y sus problemas hemerográficos : una recapitulación. Boletín de la ANABAD, 54 (40), pp. 119-138.

² <http://roai.mcu.es/es/inicio/inicio.cmd>

³ Corrió a cargo de María Antonia Carrato Mena, Subdirectora General de Coordinación Bibliotecaria. http://www.mcu.es/bibliotecas/CE/jppd/ficheros/mariaantonia_carrato_jpd.ppt

⁴ <http://www.mcu.es/bibliotecas/CE/jppd/index.htm>

⁵ En la misma fecha Recolecta : Recolector de Ciencia Abierta tiene recopilados 415.186 documentos procedentes de 79 repositorios (institucionales, de revistas, artículos, tesis, etc.).

⁶ ARCA (Arxiu de Revistes Catalanes Antiques), Biblioteca Digital Hispánica, Biblioteca Digital de Madrid, Biblioteca Digital de Castilla-La Mancha, Biblioteca Navarra Digital, Biblioteca Valenciana Digital, Biblioteca Virtual de Andalucía, Biblioteca Virtual de Aragón, Biblioteca Virtual de Derecho Aragonés, Biblioteca Virtual de La Rioja, Biblioteca Virtual Lastanosa, Biblioteca Virtual de Prensa Histórica, Biblioteca Virtual de Patrimonio Bibliográfico, Biblioteca Virtual del Principado de Asturias, Biblioteca Virtual Sierra Pambley, Galiciana Digital, Memoria Digital Vasca, Iuris Digital de la Real Academia de Jurisprudencia y Legislación.

En los dos últimos años el Ministerio de Cultura ha promovido, mediante su financiación total o parcial, la creación de nuevas bibliotecas digitales basadas en el protocolo de recolección de metadatos OAI-PMH, por lo que estas cifras crecerán considerablemente

2. Creación de recursos digitales

Es claro que para proceder a la creación de recursos digitales que permitan la máxima visibilidad y accesibilidad en la red, así como, el intercambio de información entre distintas bibliotecas digitales o virtuales, o la recolección de metadatos en diferentes repositorios, es necesario el uso de esquemas de metadatos de diferente tipo cada vez más amplios y ricos. Así, para crear los metadatos descriptivos mínimos del protocolo OAI-PMH sólo es necesario utilizar Dublin Core sin cualificar bajo el esquema 'dc_oai', pero si se pretende transmitir la relación completa de los componentes de un objeto digital y su estructura será necesario contar con el esquema METS⁷. Por otra parte, una correcta descripción bibliográfica en toda su complejidad nos llevaría a utilizar el esquema MARC XML⁸, tanto en lo que se refiere a registros bibliográficos como a registros de fondos y localizaciones, lo cual es interesante cuando se trata, tanto de materiales seriados (para la descripción completa de los niveles de numeración y cronología), como del fondo antiguo, el cual presenta, así mismo, peculiaridades importantes que han de quedar descritas a la hora de señalar las características específicas de un ejemplar concreto (una encuadernación valiosa, anotaciones manuscritas, estado de integridad del ejemplar, etc.).

Asimismo, será necesario incluir todo un conjunto de datos administrativos y técnicos a los que haremos referencia más adelante, definidos en los correspondientes esquemas de metadatos y que se ajustan a diferentes funcionalidades, entre ellas las propias de la preservación digital a largo plazo. Una guía de buenas prácticas para la construcción de colecciones digitales de calidad y para la aplicación de esquemas de metadatos a estas colecciones es

⁷ Metadata Encoding & Transmission Standard. <http://www.loc.gov/standards/mets/>

⁸ <http://www.loc.gov/standards/marcxml/>

En esa misma página puede verse una útil presentación de Rebecca Guenther y Jackie Radebaugh relacionada con los principales temas que se discuten en esta comunicación, <http://www.loc.gov/standards/mods/mods-mets-ala/mods-mets-ala.html>

*A Framework of Guidance for Building Good Digital Collections*⁹ asumidas por la Digital Library Federation¹⁰.

En esta publicación se definen una serie de principios a los que deben ajustarse los metadatos ('good metadata') utilizados en una colección digital.

- Deben seguir la normativa propia de una comunidad y ser apropiados para el tratamiento de los materiales, para los usuarios y para el uso actual y futuro de la colección.
- Deben ser interoperables.
- Deben disponer de control de autoridades y de normas específicas para la elaboración de contenido tanto de las descripciones, como de los objetos y sus relaciones.
- Deben incluir una declaración sobre las condiciones de uso de los objetos digitales.
- Deben soportar la preservación a largo plazo de todos los objetos incluidos en una colección.
- Deben, en tanto que objetos en sí mismos, tener las características de autoridad, autenticidad, persistencia, y estar identificados unívocamente.

3. Preservación digital

La preservación digital es el conjunto de estrategias, procesos y técnicas que dan respuesta a los problemas que plantea la conservación de los materiales digitales y de los medios (hardware y software) que se emplean para su almacenamiento y consulta, y que están derivados fundamentalmente de la obsolescencia provocada por la rápida renovación tecnológica y por la inestabilidad de los soportes. Estas técnicas son muy variadas y responden a diferentes situaciones y líneas estratégicas (copias de seguridad, copia de datos en soportes durables, migración, replicación, emulación, etc.), aunque en general están destinadas a mantener los objetos digitales y sus características de acceso a largo plazo. Según las *Directrices para la preservación del Patrimonio Digital de la UNESCO*¹¹,

⁹ *A Framework of Guidance for Building Good Digital Collections : A NISO recommended practice*. Prepared by the NISO Framework Working Group with support from the Institute of Museum and Library Services. 3rd Edition. December 2007. <http://www.niso.org/publications/rp/framework3.pdf>

¹⁰ <http://www.diglib.org/>

¹¹ *Directrices para la preservación del Patrimonio Digital*. Preparado por la Biblioteca Nacional de Australia. División de la Sociedad de la Información. Organización de las Naciones Unidas para la Educación, la Ciencia y la Cultura. Marzo de 2003. <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>

la preservación digital consiste en los procesos destinados a garantizar la accesibilidad permanente de los objetos digitales. Para ello, es necesario encontrar las maneras de representar lo que se había presentado originalmente a los usuarios mediante un conjunto de equipos y programas informáticos que permiten procesar los datos. Para lograrlo, es necesario que la comprensión y la gestión de los objetos digitales se realice considerándolos desde cuatro puntos de vista: como fenómenos físicos, como codificaciones lógicas, como objetos conceptuales comprensibles para el ser humano y como conjuntos de elementos esenciales que deben ser preservados para ofrecer a los futuros usuarios lo esencial del objeto.

La preocupación internacional por la preservación digital se concretó en 2002 con la publicación del *Reference Model for an Open Archival Information System (OAIS)*¹². Se trata de una recomendación para el desarrollo de un amplio consenso sobre los requisitos que debe cumplir un archivo¹³ para preservar a largo plazo la información digital y definir la norma ISO Reference Model for an Open Archival Information System (ISO 14721:2003). En esta norma un sistema OAIS es un archivo, que consiste en una organización, de equipos humanos y sistemas, que tienen la responsabilidad de preservar información y hacerla disponible para una comunidad específica ('designated community'). Un sistema será conforme a OAIS si soporta el modelo de información descrito en la norma que, por otra parte, no especifica ningún método de implantación. No es el objetivo de esta comunicación detallar el contenido de esta norma, pero sí es importante tener una noción básica del modelo informativo, que nos hará entender con mayor precisión los metadatos necesarios para la preservación.

En el modelo OAIS es básico el concepto Paquete de Información (*Information Package*) que son contenedores conceptuales que agrupan dos tipos de información, la Información de Contenido (*Content Information*) y la Información de Preservación (*Preservation Description Information*). La Información de Contenido es el objeto propio de la preservación (y sus procesos) y está formada por los objetos digitales (bits) y la información de representación asociada (p.e., juego de caracteres) y necesaria para que un objeto digital sea comprensible para una comunidad específica (hardware, programas, componentes, etc.).

¹² *Reference Model for an Open Archival Information System (OAIS)* : CCSDS 650.0-B-1 : Blue Book : Recommendation for Space Data System Standards. Consultative Committee for Space Data Systems. January 2002. <http://public.ccsds.org/publications/archive/650x0b1.pdf>

¹³ Dado el variado uso del término 'archive' en inglés ha de tenerse en cuenta que para esta norma un archivo es la organización que asume la responsabilidad de preservar una o varias colecciones digitales a largo plazo. Igualmente en esta norma el término Abierto ('Open') se refiere a que tanto la norma misma como otras normas o recomendaciones relacionadas que pudieran elaborarse se desarrollarán en foros abiertos. Por tanto, no implica ninguna connotación de restricciones o ausencia de ellas en el acceso a un archivo.

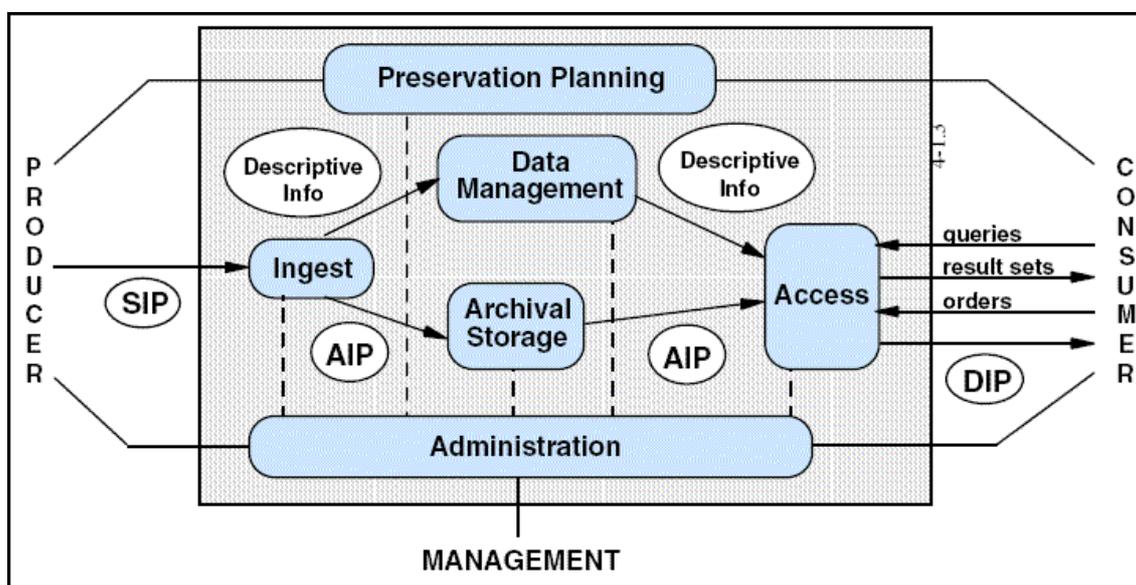
La Información de Preservación es la información necesaria para preservar la Información de Contenido y debe proporcionar suficientes datos sobre su *procedencia, contexto, referencia y fijeza*¹⁴. La procedencia describe tanto la fuente de la Información de Contenido que la ha custodiado desde su origen, como su historia, incluyendo la historia de los procesos realizados. El contexto describe cómo la Información de Contenido se relaciona con otra información externa al Paquete de Información o con otras Informaciones de Contenido. La Referencia proporciona uno o más identificadores o sistemas de identificadores unívocos, de modo que una determinada Información de Contenido sea perfectamente distinguible de otras. Por último, la fijeza proporciona un envoltorio o escudo protector que protege la Información de Contenido de cualquier alteración o que permita su detección (p.e., mecanismos de suma de verificación).

Un sistema OAIIS trabaja con paquetes de información que se distinguen según los procesos en los que participan. Así *un Submission Information Package (SIP)* es un paquete de información que se envía a un sistema OAIIS por un productor. Su forma y contenido (información de contenido e información de preservación) deben ser establecidos entre el productor y el sistema OAIIS. Una vez que este paquete de información haya sido incorporado al sistema se complementará con las informaciones específicas propias para la gestión de preservación con lo que se transformará en un *Archival Information Package (AIP)*. A su vez un sistema OAIIS devolverá a los usuarios que lo soliciten los datos necesarios en forma de un *Dissemination Information Package (DIP)*, de forma que el usuario pueda distinguir claramente la información solicitada.

Indudablemente un sistema OAIIS debe proporcionar una serie de funciones para ser considerado como tal, como son las funciones de integración (*Ingest*), de gestión de datos, de archivo de preservación, de gestión de preservación, de acceso y de administración. Igualmente, forma parte del sistema una definición específica de roles y entidades en la que el Productor son las personas o clientes del sistema que generan o proporcionan la información que debe ser preservada, el Archivo es la entidad que realiza las funciones y servicios necesarios para el control de las operaciones diarias de un sistema OAIIS, y, por último, el Consumidor (usuario) que son las personas o clientes que interactúan con los servicios de un sistema OAIIS para acceder a la información preservada.

¹⁴ En inglés 'fixity'. Según el Diccionario de la Lengua Española la segunda acepción de 'fijeza' es persistencia, continuidad.

El modelo de referencia OAIS fue el adoptado por el estudio realizado en 2002 por el grupo de trabajo *Preservation Metadata Framework*¹⁵, organizado conjuntamente por la OCLC y la RLG para definir los tipos de información que se deberían asociar a un objeto digital a efectos de preservación y que es el origen del *PREMIS Data Dictionary for Preservation Metadata*. La versión 1.0 del *PREMIS Data Dictionary* fue publicada en 2005¹⁶, y la actual 2.0 es de marzo de 2008¹⁷. Este diccionario de datos es una traslación del modelo OAIS a unidades semánticas implementables, bajo la forma de un esquema de metadatos específicos para preservación, y sustentado en encuestas sobre sistemas reales¹⁸ de repositorios de preservación.



Tomado de Reference Model for an Open Archival Information System (OAIS) : CCSDS 650.0-B-1 : Blue Book , Figure 4-1

¹⁵ *A Metadata Framework to Support the Preservation of Digital Objects (the Framework)*. http://www.oclc.org/research/projects/pmwg/pm_framework.pdf.

¹⁶ *Data Dictionary for Preservation Metadata. Final Report of the PREMIS Working Group*, May 2005. Dublin (Ohio) : OCLC ; Mountain View (California) : RLG, 2005. <http://www.oclc.org/research/projects/pmwg/premis-final.pdf>

¹⁷ <http://www.loc.gov/standards/premis/v2/premis-2-0.pdf>

¹⁸ *Implementing Preservation Repositories for Digital Materials : Current Practice and Emerging Trends in The Cultural Heritage Community*. A Report by the PREMIS Working Group, September 2004. <http://www.oclc.org/research/projects/pmwg/surveyreport.pdf>

Se trata también de un producto del consenso internacional¹⁹ basado, sin duda, en la gran experiencia de la RLG en el desarrollo y sostenimiento de proyectos como el *National Register Of Microform Masters (NRMM) Retrospective Conversion Project*²⁰, lo que le habrá proporcionado una gran familiaridad con la amplísima casuística con la que es posible encontrarse a la hora de describir las características, tanto de un material bibliográfico reproducido, como el de la propia copia fotográfica en este caso.

El modelo de datos de PREMIS define cinco entidades especialmente importantes para las actividades de preservación digital, las entidades intelectuales, los objetos digitales, los eventos, los agentes y los derechos, y sobre ellos construye las unidades semánticas apropiadas. Las entidades intelectuales (monografías, publicaciones periódicas, partes componentes, mapas, fotografías, etc.) que son las obras²¹ que se consideran a efectos de descripción y gestión, o en términos de FRBR las manifestaciones de una obra, que pueden tener relaciones con otras obras y con una o más representaciones digitales (p.e., las imágenes TIFF de un artículo, un fichero PDF formado por imágenes y texto oculto, o uno o varios ficheros con el texto obtenido de un proceso de OCR). Los objetos digitales son las unidades de información a las que se refieren los procesos de preservación digital y pueden ser ficheros, cadenas de bits o representaciones. Los eventos son las acciones que se realizan sobre un objeto, o por un agente, en un repositorio de preservación. Los agentes son las personas, organizaciones, sistemas, etc. relacionados con un evento en la vida de un objeto o con los derechos particulares sobre un objeto. Los derechos son declaraciones sobre permisos de preservación asociados a un objeto o agente.

Al igual que el modelo OAIS, PREMIS no define ninguna estrategia o método de implementación, ni prescribe una tipología de acciones, ni el modo en el que deben recogerse, almacenarse o gestionarse los datos referidos a estas acciones.

¹⁹ Las instituciones participantes están citadas en la página de agradecimientos del PREMIS Data Dictionary. En su comité editorial participan la Library of Congress, el General Register Office for Scotland, la Koninklijke Bibliotheek, la National Library of Australia, la British Library, la Universitätsbibliothek Göttingen, la OCLC y los Library and Archives Canada

²⁰ http://www.arl.org/preserv/presresources/Microform_masters~print.shtml

²¹ La definición de PREMIS de 'entidad intelectual' parece corresponderse con la de 'manifestación' en FRBR más que con la de 'obra' o 'expresión', a pesar de la mención al contenido intelectual. El concepto PREMIS de 'representación digital' puede corresponderse tanto con una 'manifestación' (p.e., en el caso de una obra nacida digitalmente) o con un 'item' (p.e., en el caso de la digitalización).

La conformidad con PREMIS requiere que un repositorio de preservación siga su modelo de datos, con una cierta flexibilidad para la definición de unidades semánticas no contempladas en el diccionario de datos. Sí define como obligatorio el mínimo de información necesaria para soportar actividades de preservación que debe acompañar a un objeto digital en los procesos de transferencia de un repositorio de preservación a otro.

¿En qué medida serían por tanto necesarios los metadatos de preservación en el entorno de la digitalización de materiales bibliográficos ya que ninguna de las actuales colecciones digitales españolas puede definirse como repositorio de preservación, ni en términos de OAIS ni de PREMIS?. Ambos modelos establecen la necesidad de registrar una serie de datos mínimos que permitan seguir el ciclo de vida de los objetos digitales para garantizar su accesibilidad futura, su correcta interpretación, su autenticidad, su integridad, etc., cualquiera que sea el sistema o metodología que se utilice para ello. En ambos modelos, el punto clave son los datos asociados a la información digital en el momento de ingreso en un sistema, procedentes de un proveedor de contenidos digitales (p.e., proceso de digitalización) o de un intercambio de información entre repositorios, en la versión más amplia de este término. En consecuencia y como fruto de un proceso de digitalización es recomendable que se registren esa serie de datos mínimos independientemente de que formalmente el ingreso de información se realice en un repositorio de preservación o no. Ha de tenerse en cuenta además que cualquier sistema del tipo de las bibliotecas digitales dispone de procedimientos de preservación básicos (p.e, copias de seguridad, actualización de versiones entre formatos, transformación de un formato a otro, comprobación de alteraciones en ficheros, etc.) que se realizan regularmente, aunque propiamente no puedan denominarse repositorios de preservación.

Así pues, en el momento de ingreso ('ingest') de un conjunto de informaciones en cualquier repositorio, sea o no sea este de preservación, que procedan de un proceso de digitalización (un *Submission Information Package*, según OAIS) se debe disponer de los metadatos propios de preservación (según PREMIS) para poder realizar el seguimiento actual o futuro del ciclo de vida de los objetos digitales. Esto implica registrar los metadatos propios de la descripción de las unidades intelectuales (descripción bibliográfica); los metadatos técnicos que reflejan las características de los ficheros (su tipología, su formato, la versión del formato, el software y hardware con el que ha sido creado, el agente que lo ha creado, los niveles de resolución, la profundidad de bits, el formato de compresión, el índice de reducción, el tamaño del fichero, la fecha de creación, los algoritmos de verificación, etc.); los metadatos administrativos propios de los derechos de acceso, uso y reproducción; la relación de ficheros que puede componer una representación digital; y la estructura lógica y física de sus ficheros.

El registro de estos metadatos puede realizarse de diferentes maneras, o bien usando PREMIS como esquema básico o bien utilizando PREMIS como una extensión²² de otros esquemas como METS²³, para lo que pueden utilizarse las recomendaciones *Guidelines for using PREMIS with METS for exchange*²⁴. Desde un primer momento se previó la posibilidad de incluir los metadatos PREMIS dentro de un fichero METS, que se transformaría de este modo en un *Submission Information Package (SIP)* de entrada en cualquier repositorio de preservación, o en cualquier repositorio según hemos recomendado. La versión 1.8 de METS hace una recomendación concreta para se haga de esta manera. Incluso existen otros esquemas de metadatos que pueden coincidir total o parcialmente con lo establecido en PREMIS (p.e., la unidad semántica *ObjectCharacteristics* y el esquema *Metadata for Still Images, MIX*²⁵ o el esquema *Technical Metadata for Text*²⁶), lo que supone que su registro puede realizarse por medio de diferentes esquemas o combinándolos. Evidentemente, cualquier sistema que reciba un paquete de información procedente de un proceso de digitalización, debe establecer sus requisitos mínimos y obligatorios, o bien haciendo referencia a alguno de los perfiles existentes²⁷ o bien desarrollando sus propios perfiles. Es incluso posible que los datos que hemos mencionado como obligatorios a efectos de preservación digital puedan incluirse en un fichero METS sin hacer uso de PREMIS como extensión, pero es imprescindible registrarlos.

En cualquier caso, parece muy lógico utilizar el esquema PREMIS, dada la riqueza y precisión creciente de las distintas versiones de su diccionario de datos. En resumen, debe quedar claro que la no existencia de un repositorio de preservación normalizado no exime de la necesidad de utilizar metadatos de preservación y en concreto PREMIS.

²² Así viene solicitándose de una manera creciente en los pliegos de prescripciones técnicas para la prestación de servicios de digitalización.

Using Premis with METS refleja una serie de perfiles que especifican el uso de los metadatos PREMIS envueltos en un esquema METS y la relación con otras extensiones de metadatos que pueden coincidir total o parcialmente con METS. <http://www.loc.gov/standards/premis/premis-mets.html>.

Resulta de gran interés el perfil ECHO Dep Generic METS Profile for Preservation and Digital Repository Interoperability <http://www.loc.gov/standards/mets/profiles/00000015.html>

²⁴ <http://www.loc.gov/standards/premis/guidelines-premismets.pdf>

²⁵ <http://www.loc.gov/standards/mix/>

²⁶ <http://www.loc.gov/standards/textMD/>

²⁷ Estos perfiles incluyen la definición de aplicación de una institución (p.e, National Library of Australia) o de un sistema concreto (p.e., Dspace).

Así, es muy razonable optar por la solución de incluir encapsulados en un fichero METS aquellos elementos de PREMIS que hacen referencia al proceso de digitalización (o de OCR o de generación de versiones diferentes de una copia master), detallando todo el proceso (hardware y software), sus agentes y sus resultados (ficheros y características técnicas de los mismos). El registro de estos metadatos hará posible que en cualquier momento puedan volver a ser regenerados, detectando los cambios que hayan podido afectar a los objetos digitales, sobre todo aquellos que puedan verse afectados por una falta de normalización, en cualquiera de las características básicas de un objeto digital.

Dada la creciente implantación tanto de METS como de PREMIS se están desarrollando diferentes herramientas para la extracción de metadatos²⁸ y, lo que es muy importante, para la validación de los esquemas utilizados. Entre ellos convendría destacar a JHOVE (*Harvard Object Validation Environment*), entre otras razones porque facilita la creación de un SIP para su ingreso en un repositorio OAIS y además porque en la nueva extensión de METS para los metadatos técnicos de ficheros de texto *Textmd*²⁹ la incluye como herramienta recomendada.

Todas las normas, recomendaciones y perfiles mencionados, empezando por el modelo OAIS, están sustentados en el beneficio de la normalización. Probablemente, sería muy difícil intentar llevar a cabo una preservación digital a escala global, y mucho menos cooperativa o coordinada, si no se utilizarán procedimientos normalizados que, por un lado permitieran crear y diseñar un software específico, y por otro lado establecer acuerdos de colaboración. Cuando se piensa, por ejemplo, que la Library of Congress, probablemente la institución más potente en el ámbito bibliotecario, ha llevado a cabo acuerdos estratégicos con terceras instituciones para proceder a la preservación del patrimonio digital norteamericano, no cabe duda que esto no sería posible si no se contara con esos instrumentos de normalización.

²⁸ METS Tools & Utilities. <http://www.loc.gov/standards/mets/mets-tools.html>. Y Tools for preservation metadata implementation. <http://www.loc.gov/standards/premis/tools.html>

²⁹ No es objeto de esta comunicación profundizar en *Textmd*, pero no puede dejarse de mencionar la importancia estratégica que habrá de tener a la hora de normalizar los objetos digitales a texto completo resultantes de los procesos de OCR, incluyendo el reconocimiento de segmentos, y el predominio del texto en el siglo XXI sobre las imágenes como representación digital de los materiales bibliográficos.

De aquí, evidentemente, se extrae la lección de que por muy voluntariosos que sean determinados proyectos que puedan llevarse a cabo en España para la preservación del patrimonio digital, si no participan, ni cuentan con la normalización que suponen las normas descritas, es muy posible que su resultado sea baldío. No cabe duda que OAIS y PREMIS serán los modelos que sustenten cualquier iniciativa en este sentido.

En este punto ha de tenerse en cuenta todas las recomendaciones elaboradas por la Comisión Europea³⁰ que no sólo van dirigidas a la creación de bibliotecas digitales, sino también promover una política de preservación digital, sin duda basada en los estándares citados. Posteriormente a esa resolución se añadió una segunda que pretendía no sólo la preservación y difusión del patrimonio cultural europeo sino también lo que podríamos denominar del bagaje científico, es decir los recursos basados en repositorios institucionales y académicos, fundamentalmente vinculados con los denominados documentos electrónicos de investigación.

Últimamente, con el objetivo de estructurar estos repositorios conforme a una normativa, ha surgido en Europa, y también se está aplicando en España tanto por REBIUN como por el Ministerio de Cultura, la utilización de la normativa DRIVER que tiene como objetivo certificar determinadas características de los repositorios OAI. De esta manera y mediante tres acrónimos que guardan una gran similitud OA (Open Access), OAI (Open Archive Initiative) y OAIS (Open Archival Information System), se puede intentar lograr red de repositorios, candidatos a futuros repositorios de preservación, el acceso abierto a los mismos y la preservación digital a largo plazo de todos estos recursos. Por último, conviene hacer notar que desde el punto de vista bibliotecario resultan fundamentales las experiencias de la Koninklijke Bibliotheek³¹ o de la National Library of Australia³², o del National Digital Information Infrastructure and Preservation Program (NDIIPP)³³ de la Library of Congress.

30 *Report on Digital Preservation, Orphan Works, and Out-of-Print Works* : Selected Implementation Issues : (adopted by the High Level Expert Group at its third meeting on 18.4.2007). i2010: Digital Libraries High Level Expert Group – Copyright Subgroup.

http://www.edlproject.eu/downloads/report_HLEG_preserv_orphan_works.pdf

³¹ Digital Preservation Strategies at the KB
http://www.mcu.es/bibliotecas/CE/jppd/ficheros/barbara_siermann_jpd.pdf

³² <http://www.nla.gov.au/policy/digpres.html>

³³ <http://www.digitalpreservation.gov/>

Parece que en la situación española, de carácter fuertemente distribuido, un proyecto como el de los Estados Unidos (quizá sustentado en repositorios OAI y OAIS) podría significar un modelo real por el carácter participativo y colaborador entre diferentes instituciones y administraciones.

Finalmente, se insistirá en la recomendación del poeta latino cuando recomendaba crear monumentos más duraderos que el bronce. Así nosotros deberemos crear objetos digitales más duraderos, que estén accesibles y consultables a lo largo del tiempo, luchando contra el reloj del peligro de desaparición a causa de la acidez de los materiales bibliográficos y contra el peligro de la rapidísima evolución tecnológica.