

## LOS ARCHIVOS Y LAS BIBLIOTECAS ANTE LA PRESERVACIÓN DIGITAL: ¿UN SOLO ENFOQUE?

### ARCHIVES AND LIBRARIES BEFORE DIGITAL PRESERVATION: A SINGLE APPROACH?

**Térmens Graells, Miquel.** Universidad de Barcelona. Departamento de Biblioteconomía y Documentación. Melcior de Palau, 140. 08014 Barcelona. termens@ub.edu

**Resumen:** La investigación y el desarrollo en preservación digital se están efectuando de forma indistinta en archivos y bibliotecas de distintos países. Los dos tipos de instituciones están usando aparentemente los mismos estándares y las mismas herramientas para llegar a soluciones parecidas, pero un análisis más pormenorizado muestra que existen amplias diferencias. ¿Son estas diferencias herencia de los desencuentros históricos entre archivos y bibliotecas o son el resultado de nuevas diferencias objetivas? Esta comunicación pretende dar respuesta a esta incógnita.

A partir de una revisión de la literatura en preservación digital y de los proyectos en curso a nivel internacional se detectarán que puntos en común y que diferencias presenta el reto de la preservación digital en archivos y bibliotecas. A nivel indicativo, avanzamos que se observan grandes diferencias en la percepción del concepto de preservación “a largo plazo”, que la urgencia de la preservación digital es vista de forma muy distinta, y que se dan pesos desiguales a los conceptos de acceso, autenticidad, recuperación y descripción, entre otros. Las diferencias continúan en temas centrales actualmente a debate internacional como la función de los metadatos, que esquemas usar, el control del proceso de vida de los documentos y las características significativas a conservar. A nivel organizativo, las instituciones actualmente pioneras en implantar soluciones se articulan de forma diferente en los dos ámbitos de actuación y también varían en su capacidad de relacionarse con otras instituciones o departamentos. A nivel de volumen de trabajo y de requerimientos informáticos las situaciones pueden ser dispares y no siempre en el mismo sentido.

Si nos centramos en los profesionales, archiveros y bibliotecarios identifican peligros y oportunidades distintas en la gestión de los documentos digitales y en su preservación. Conocer todas estas diferencias y también coincidencias ha de permitir que cada colectivo respete mejor al otro y vea los beneficios de colaborar con él en objetivos comunes.

**Palabras clave:** preservación digital; bibliotecas digitales; archivos electrónicos; administración electrónica; digitalización.

**Keywords:** digital preservation; digital libraries; digital archives; electronic administration; digital imaging; digital curation.

## 1. Antecedentes

Desde mediados de la década de 1990 distintas instituciones y empresas están estudiando la problemática de la preservación futura de los documentos digitales. La expansión de este tipo de documentos ha llevado a que en determinados ámbitos su presencia sea mayoritaria y por tanto se convierta en ineludible su conservación por razones patrimoniales, culturales, jurídicas o de otro tipo.

Actualmente aún están en desarrollo las soluciones técnicas que han de hacer posible la preservación de objetos digitales a largo plazo, pero sí existe un consenso sobre los parámetros básicos de estas soluciones que, de una forma muy sintética, se pueden relacionar en: el seguimiento del modelo teórico OAIS (Reference, 2002); el control de los ficheros en el momento de la ingestión en el sistema; la recopilación de metadatos descriptivos, técnicos y de gestión; y la aplicación de técnicas de migración de formatos y de emulación de aplicaciones y plataformas informáticas. En la práctica el consenso no es tan evidente; baste recordar que el modelo OAIS fue desarrollado fuera de las comunidades bibliotecarias y archivísticas, que se ha reclamado la necesidad de un modelo de preservación menos complicado y más aplicable a casuísticas simples (Allinson, 2006), y también el escepticismo o entusiasmo, según los casos, con el que es recibida la opción de emulación.

Se plantea la incógnita de hasta qué punto es posible establecer un modelo y unas soluciones universales de preservación. Si ello es cuestionable a nivel de tipologías documentales –son evidentes las diferencias entre una colección finita de fotografías (objetos digitales únicos) y una biblioteca de libros digitales (objetos compuestos con múltiples formatos de salida)–, también lo es a nivel de centros en los que se implementa –una empresa multinacional, una biblioteca nacional, un archivo municipal–, como en la tipología de éstos –archivo, biblioteca, museo...–. El presente artículo se centra en estas últimas diferencias, más concretamente en las que se pueden establecer entre las necesidades de preservación de las bibliotecas y los archivos. La exploración de las necesidades de estos dos ámbitos solo pretende poner en evidencia sus respectivos requerimientos y prioridades y contribuir a una reflexión sobre las mismas. Solo a partir del reconocimiento de las diferencias se podrán establecer puentes sólidos de colaboración.

## 2. Archivos y bibliotecas como gestores de preservación digital

En el futuro inmediato sin duda surgirán instituciones especializadas en la preservación digital, como única función, sea como misión propia o como soporte a otras entidades; de hecho esto ya está ocurriendo, véanse los ejemplos de Physikalisch-Technische Bundesanstalt (PTB) (Siegfried, 2005), Portico (Fenton, 2006) y iArxiu (Alamillo, 2007). Sin embargo la situación más general será aquella en la que la preservación digital se añade como una función más a las que ya realizan las instituciones, veamos cómo ello se produce en bibliotecas y archivos.

En la mayoría de bibliotecas la preservación no es una finalidad en si misma sino simplemente una actividad necesaria que ha de permitir dar cumplimiento a su función

primordial que es dar acceso a la información y la documentación. La preservación es, por tanto, una actividad complementaria, no esencial y a la que se presta una atención limitada, actitud que luego repercute en los recursos dedicados. Como consecuencia, los procedimientos y las aplicaciones informáticas de preservación se diseñan como apéndices a otros procedimientos o aplicaciones, o como simples módulos o rutinas de aquellos; un ejemplo es el complicado encaje de la preservación con software de repositorios tan usados como DSpace y Fedora.

En los archivos la preservación es la razón de ser de los mismos, aunque sin perder de vista que el objetivo no es simplemente conservar sino estar preparados para dar acceso a los contenidos en el momento que sea preciso. Por ello los procedimientos y las aplicaciones informáticas de preservación son parte del núcleo del archivo y han de estar perfectamente en conexión con los restantes procedimientos y aplicaciones. En este sentido se considera vital la comunicación y la integración con los sistemas de gestión documental.

Actualmente las bibliotecas que están encarando el reto de la preservación son aquellas que han acumulado grandes volúmenes de datos sea por digitalización de materiales impresos –es el caso, entre otras, de las bibliotecas nacionales–, por la creación de repositorios instituciones –caso de las bibliotecas universitarias– o, más raramente, por el afán de conservar materiales adquiridos, como las revistas electrónicas. Se trata de un problema claramente delimitado en un sistema informático concreto y con un gran volumen de datos (terabytes, petabytes) que a medio plazo dificulta fiar su conservación exclusivamente a los sistemas informáticos habituales de copias de respaldo. La previsión de estas bibliotecas es de un crecimiento muy rápido de sus necesidades de preservación digital. A nivel de formatos se puede dar una gran variabilidad, dada la tendencia de las bibliotecas a crear colecciones digitales de cualquier material que esté a su alcance.

Los archivos a la larga acumularán más materiales pero ésta no es su preocupación actual, que está centrada en sobrevivir a –o sea, en gestionar correctamente– las remesas que ya les están llegando y en preparar la infraestructura para dar servicio a las necesidades de la administración electrónica. Así pues, el reto digital depara una paradoja relevante: si tradicionalmente se ha entendido que los archivos están preocupados por el largo plazo, con la preservación digital lo que les preocupa es el corto plazo, como alguno de ellos diría: si conseguimos sobrevivir a los próximos cinco años seguro que también tendremos éxito en los diez siguientes. De forma significativa hoy en día es más fácil escuchar elucubraciones sobre la preservación a largo plazo en reuniones de bibliotecas que en las de archivos; no olvidemos que las bibliotecas están sometidas a presiones de globalización mucho más fuertes, véase sino el previsible impacto de iniciativas como la de Google Books o el poder de proveedores comerciales como Elsevier o Thompson Reuters.

### **3. Los documentos custodiados**

En los archivos los documentos están muy pautados, siguen tipologías documentales conocidas y habitualmente se organizan en series. Un problema especial de los archivos administrativos, en especial de la administración pública, es la organización en expedientes,

que agrupan los documentos por su relación con un trámite administrativo con independencia de su tipología documental. Los expedientes presentan un grave problema de gestión en el ámbito digital pues obligan a mantener la relación de ficheros dispares, por ejemplo un registro de una base de datos con diversos documentos de procesador de textos, posiblemente organizados y almacenados de forma distinta a nivel informático.

En las bibliotecas el número de tipologías documentales puede ser menor (libros, revistas, artículos, *working papers*...) pero éstas pueden llegar a complicarse mucho en el caso de bibliotecas digitales que recopilan materiales diversos: mapas, fotografías, manuscritos, etc.

Entre los dos tipos de centros no solo es distinta la naturaleza de los documentos gestionados sino también las condiciones en las que éstos son preservados. Los archivos custodian documentos únicos, por eso en cualquier sistema archivístico es primordial asegurar la autenticidad de los documentos; autenticidad que no se confiere por sí, por el mismo archivo, sino que ha de ser propia de cada documento. La fiabilidad proviene del origen del documento, de que es auténtico y que se garantiza y se puede probar que no ha sido alterado desde su creación. En documentos electrónicos este principio de autenticidad viene asegurado por técnicas como la firma electrónica y el sellado de tiempo, entre otras, unas medidas de seguridad informática que se aplican como instrumento para fijar el valor jurídico de los documentos.

La situación en las bibliotecas es bien distinta: dan acceso a documentos no necesariamente únicos, al contrario, en muchas ocasiones son meramente copias de un producto industrial llamado libro o revista. Incluso en los casos de colecciones propias de documentos únicos, como pueden ser colecciones fotográficas, se considera que la autenticidad viene garantizada por la institución: ¡las bibliotecas no nos engañan! En las bibliotecas digitales no existe una gran preocupación por demostrar la autenticidad de los documentos, ni siquiera por garantizar su integridad a un nivel elevado, pues son propiedades que se transmiten desde el origen. Aunque la posibilidad de deterioro tecnológico o incluso de manipulación no se descarta, en la práctica no se espera que las bibliotecas respalden jurídicamente la veracidad de los documentos que ofrecen.

La protección de los datos personales también obliga a los archivos a prever mayores niveles de control en el acceso a los documentos custodiados, preocupación que en la mayoría de casos no se presenta en las bibliotecas dada la naturaleza de documentos ya publicados o bajo licencia de la mayoría de objetos que están bajo su cobertura.

El distinto peso que bibliotecas y archivos dan a las propiedades de los documentos tiene después su correspondencia en el acceso que se da a los mismos. Destaca la preocupación que se observa en las bibliotecas por las características de las interfaces de acceso y por mantener el *look and feel* original de los documentos, que no son otra cosa que la traslación de las preocupaciones de sus clientes habituales. Se prima por tanto mantener la apariencia del presente, incluso siguiendo a formatos impresos sin demasiada viabilidad futura, frente a soluciones a largo plazo. Sin duda son visiones que en los archivos digitales resultan minoritarias.

#### 4. Los modelos de cooperación interinstitucional

La cooperación entre instituciones es cada vez más necesaria para conseguir dar un servicio adecuado. En las bibliotecas la cooperación normalmente se traduce en políticas conjuntas de adquisiciones y en el acceso y la agregación de contenidos externos; en los archivos la forma principal de cooperación es la transferencia de fondos entre organismos relacionados. A continuación observaremos el comportamiento de estos modelos en la preservación digital.

Las bibliotecas no transfieren sus fondos, pues a nivel patrimonial básicamente acumulan y, en su caso, aplican expurgos. Por otro lado ya hemos indicado que para ellas dar el máximo acceso a los materiales y al mayor número de los mismo es su máximo objetivo. Ante esta disyuntiva –un patrimonio cerrado frente un acceso en expansión– el medio digital ha ofrecido la solución de la recolección virtual ante el usuario de fondos dispersos en distintas bibliotecas; ello se lleva a cabo principalmente mediante metabuscadores y recolectores de metadatos –no de los documentos en sí–, entre los que las iniciativas Europea (<http://www.europeana.eu/>) y Driver (<http://search.driver.research-infrastructures.eu/>) son ejemplos destacados. En este escenario lo más importante es asegurar la intercomunicación de las aplicaciones en tiempo real o casi real, función en la que actualmente está teniendo un papel central el protocolo OAI-PMH (<http://www.openarchives.org/OAI/openarchivesprotocol.html>) y el esquema de metadatos Dublin Core (<http://dublincore.org/>) y sus adaptaciones, como Driver (<http://www.driver-repository.eu/>). En otros casos la recuperación de la información es posible gracias al uso de esquemas de descripción unificados, básicamente MARC21 (<http://www.loc.gov/marc/>), sea en formato nativo o bajo pasarelas XMLMARC (<http://www.loc.gov/standards/marcxml/>) o similares. Conseguida la intercomunicación, no se considera primordial, como mínimo en el estadio actual, llegar a otros niveles de homogeneización. Esta vocación de intercomunicación casi universal –si duda uno de los *leiv motiv* de las bibliotecas– no impide que se estén dando pasos importantes hacia la constitución de unos pocos grandes hosts en determinadas áreas, como las revistas y los libros electrónicos (Hoorens, 2007; Kenney, 2006).

Por su parte, los archivos hasta ahora no han facilitado la consulta agregada de sus fondos, ni siquiera a nivel de inventario o de catálogos, posiblemente debido a un bajo nivel de descripción de sus fondos. La transferencia de documentos sí que forma parte de sus prácticas profesionales tradicionales y bien pronto se aplicará a los fondos digitales. Como resultado de esta situación, los archivos no están excesivamente preocupados por la interconexión de sistemas, que en todo caso no tienen porqué ser en tiempo real, y sí en normalizar sus instrumentos de descripción. En este punto destaca la iniciativa de la Comisión Europea en la promoción del estándar Moreq2 (Moreq, 2008).

Paradójicamente algunas bibliotecas sí que han avanzado bastante en el análisis de la problemática inherente a la transferencia de colecciones digitales, como es el caso del Archiving, Ingest and Handling Test (AIHT) ejecutado el 2004-2005 por la Library of Congress y otras 4 bibliotecas universitarias (Shirky, 2005).

Llama la atención que en los Estados Unidos, uno de los países que más está avanzando a nivel práctico en preservación digital, tanto su biblioteca nacional como sus archivos

federales están avanzando rápidamente hacia sistemas de preservación en red, a nivel institucional, funcional y tecnológico. El año 2000 la US National Digital Information Infrastructure and Preservation Program (NDIIPP), consagró la visión de la preservación en red bajo la coordinación de la Library of Congress (Friedlander, 2002). En paralelo la National Archives and Records Administration (NARA) en el programa Electronic Records Archives (ERA) ha planeado y está implementando un sistema de archivos federales en red (Thibodeau, 2007).

## **5. La visión de los profesionales**

Las actitudes de los profesionales bibliotecarios y archiveros ante el reto de la preservación presentan coincidencias y diferencias a la vez.

Si nos fijamos en el espacio que a la preservación digital se está dedicando en los respectivos congresos y seminarios, no cabe duda de que los archiveros como mínimo siguen con más atención el tema y un número importante de ellos intenta formarse en las nuevas habilidades y conocimientos que se demandan. Para los archiveros la preservación digital representa un doble oportunidad: por un lado conseguir una visibilidad social de la que los archivos nunca han gozado (de aquí algunos proyectos ligados al acceso abierto de los fondos digitalizados); de otro, gracias al impulso dado por la administración electrónica, conseguir, en el caso de los archivos de la administración, la integración total en las respectivas instituciones y alcanzar suficiente poder para influir en una visión amplia de la gestión documental. Sin duda las presiones, las oportunidades y las posibles ganancias futuras son mayores en la administración electrónica que en la preservación digital pura, aunque ambas van ligadas. Si las oportunidades son grandes, los peligros también: más de un responsable de archivos ha dicho que se encuentran ante la última oportunidad para que los archiveros y los archivos sean algo más que simples depósitos a ojos de sus respectivas instituciones; si no lo logran significará que buena parte de las funciones del archivo digital habrán sido directamente asumidas por las áreas de informática.

A ojos de muchos bibliotecarios la situación es distinta: su función y la de las bibliotecas está siendo cuestionada por las crecientes facilidades de acceso a la información digital. Sus problemas profesionales son el resultado del número creciente de servicios digitales que han de crear y gestionar, y que demandan una mayor productividad y una mayor capacitación técnica; en esta perspectiva, la preservación digital no es otra cosa que un servicio más a mantener. Los bibliotecarios no se enfrentan a la preservación de forma fría como los archiveros –pues para ellos la conservación es una función inherente– sino como un problema cercano, que explota a partir de la creación de bibliotecas digitales y de repositorios institucionales con un gran volumen de información almacenada. Para los bibliotecarios la preservación es básicamente un problema de sostenibilidad tecnológica de sus servicios digitales, y en esto se encuentran más cercanas a la visión pragmática de los informáticos, una visión que puede tener la tentación de primar las soluciones a corto y medio plazo ante previsiones más sostenibles y costosas a largo plazo.

## 6. Conclusiones

Archivos y bibliotecas se enfrentan al mismo reto de adaptar sus funciones tradicionales a las exigencias de los nuevos documentos digitales y asegurar su pervivencia para los usuarios futuros. El catálogo de soluciones será común pero su implementación estará adaptada a cada ámbito. Hoy en día, cuando las acciones de preservación aún sólo son viables en unas pocas grandes instituciones, ya se observa que archivos y bibliotecas cuentan con necesidades distintas; destacan en este sentido las diferencias en el volumen de los datos a gestionar (a corto y largo plazo), los requerimientos funcionales presentes y futuros de los documentos tratados, y el distinto grado de integración con las instituciones fuente de los documentos y con otros archivos y bibliotecas.

Es deseable que las dos comunidades sigan cooperando más allá de sus diferentes necesidades pues las soluciones que una quizás esté aplicando ahora serán necesarias para la otra más adelante. Pueden ser indicadoras de este camino inverso los almacenamientos a gran escala que actualmente planean algunas bibliotecas –sin duda, el futuro de muchos archivos– y la robustez que en el aseguramiento de la autenticidad de los documentos desean alcanzar la mayoría de archivos –una mejora necesaria en según qué colecciones bibliotecarias–.

### Referencias bibliográficas

- Alamillo, Ignacio; Cruellas, Marta.** “El projecte iArxiu: custòdia segura i preservació a llarg termini de documents electrònics”. *Lligall*, 2007, n. 26, pp. 263-300.
- Allinson, Julie.** *An introduction to OAIIS and reference models for repositories*. University of Bath, 2006. Consultado en: 30-10-2009. <http://www.ukoln.ac.uk/ukoln/staff/j.allinson/oais-ja-progmeet-2006-10-25.pdf>
- Fenton, Eileen Gifford.** “An Overview of Portico: An Electronic Archiving Service”. En: *Serials Review*, 2006, v. 32, n. 2, pp. 81-86.
- Friedlander, Amy.** “The National Digital Information Infrastructure Preservation Program. Expectations, Realities, Choices and Progress to Date”. En: *D-Lib Magazine*, 2002, v. 8, n. 4. Consultado en: 30-10-2009. <http://www.dlib.org/dlib/april02/friedlander/04friedlander.html>
- Kenney, Anne R.; Entlich, Richard; Hirtle, Peter B.; McGovern, Nancy Y.; Buckley, Ellie L.** *E-journal archiving metes and bounds: a survey of the landscape*. Washington, DC: Council on Library and Information Resources, 2006. ISBN 978-1-932326-26-0. Consultado en: 30-10-2009. <http://www.clir.org/pubs/reports/pub138/pub138.pdf>
- Hoorens, Stijn; Rothenberg, Jeff; van Oranje-Nassau, Constantijn; van der Mandele, Martin, Levitt, Ruth.** *Addressing the uncertain future of preserving the past. Towards a robust strategy for digital archiving and preservation*. Santa Monica (EUA): RAND Corporation, 2007. Consultado en: 30-10-2009. [http://www.rand.org/pubs/technical\\_reports/2007/RAND\\_TR510.pdf](http://www.rand.org/pubs/technical_reports/2007/RAND_TR510.pdf)

**MoReq2.** *Model Requirements for the management of electronic records. Update and extensión 2008. MoReq2 Specification.* Luxembourg: Office for Official Publications of the European Communities, 2008. ISBN 978-92-79-09772-0. Consultado en: 30-10-2009. <http://www.moreq2.eu/>

*Reference model for an open archival information system (OAIS). Blue Book.* Washington DC: Consultative Committee for Space Data Systems, 2002. Consultado en: 30-10-2009. <http://public.ccsds.org/publications/archive/650x0b1.pdf>

**Shirky, Clay.** *Library of Congress Archive Ingest and Handling Test (AIHT). Final Report.* [Washington DC]: National Digital Information Infrastructure Preservation Program, 2005. 46 p. Consultado en: 30-10-2009. [http://www.digitalpreservation.gov/partners/aiht/high/ndiipp\\_aiht\\_final\\_report.pdf](http://www.digitalpreservation.gov/partners/aiht/high/ndiipp_aiht_final_report.pdf)

**Siegfried Hackel.** “The ArchiSafe Project - legally secure and scalable long-term record keeping”. En: *DLM Forum 2005*, 2005. Consultado en: 30-10-2009. [http://ec.europa.eu/transparency/archival\\_policy/dlm\\_forum/doc/26\\_hackel\\_06-10-05pm.pdf](http://ec.europa.eu/transparency/archival_policy/dlm_forum/doc/26_hackel_06-10-05pm.pdf)

**Thibodeau, Kenneth.** “Building the Archives of the Future. Advances in Preserving Electronic Records at the National Archives and Records Administration”. En: *D-Lib Magazine*, 2001, v. 7, n. 2. Consultado en: 30-10-2009. <http://www.dlib.org/dlib/february01/thibodeau/02thibodeau.html>