



Digitalización de libros

Revista de Comunicación, Arte y Literatura

Principal

Artículos opinión

Estudios / Informes

Formación

Comunicación

Entrevistas

Literatura / Ensayo

Reseñas

Relatos breves

Exposiciones

Fotografía

Enlaces

Dar de Alta / Baja

La digitalización de libros

El pasado mes de diciembre tuvo lugar en la [Residencia de Estudiantes de Madrid](#) un interesante seminario dirigido por [José Antonio Millán](#) sobre “El archivo de la Internet española” que reunió a varios especialistas en la materia para analizar si merece la pena o no preservar lo digital. En formato digital hoy discurre buena parte de la creación y difusión de la creación cultural en español, pero tan sólo un 4% del contenido en la Red está redactado en español. Los nueve especialistas que han formado parte del seminario se han preguntado si basta con la labor de los buscadores como [Google](#) o de iniciativas de diversas bibliotecas públicas para preservar las obras y las webs en nuestra lengua. Desde nuestro punto de vistas, las editoriales deberían actuar rápidamente para preservar todo aquello sobre sus autores que sea útil para los investigadores del mañana.

Sin lugar a dudas, la iniciativa de digitalización de libros de [Google](#) representa el principal desafío de las editoriales a la hora de definir su modelo de negocio en los próximos años, pero desgraciadamente, las editoriales españolas no se sienten muy cómodas con las nuevas tecnologías. ¿Cuántas editoriales españolas y/o escritores guardan cada una de las versiones de una determinada obra para preservar todo lo que sea útil para los investigadores del mañana? ¿Cuántas editoriales graban y archivan las presentaciones de libros de sus autores? ¿Cuántas editoriales archivan las reseñas publicadas en medios digitales sobre los libros de sus autores? Con la llegada de Internet, la manera de trabajar de los medios de comunicación y de los escritores está experimentando una gran transformación, pero la mayoría de las editoriales españolas no tienen en cuenta estos cambios.

La intervención de Luis García Melero del [Ministerio de Administraciones Públicas](#) destacó la necesidad de recopilar, preservar y conservar los originales de un escritor y de sus documentos particulares creados en forma digital como fuentes para el estudio actual y de un inmediato futuro. En la actualidad, la mayoría de los escritores redactan sus obras empleando un sistema informático. Cuando finalizan su redacción, proporcionan al editor una copia en soporte informático. Esta copia equivaldría al original manuscrito o manuscritos mecanografiados de otras épocas. Según este especialista, ninguna institución pública ni privada se ha preocupado por la recopilación y conservación de los originales electrónicos o “manuscritos” digitales como los denominan algunos autores. La preservación del patrimonio digital debería comenzar desde el momento de su creación. ¿Qué pasa si la editorial desaparece o los ficheros del autor se pierden o destruyen?

Dado su interés, publicamos a continuación la intervención completa de Luis García Melero.

EL ARCHIVO DE DOCUMENTOS PERSONALES DIGITALES

© Luis Ángel García Melero, Área de Publicaciones, Biblioteca y Documentación
Ministerio de Administraciones Públicas

ORIGEN DEL PATRIMONIO DIGITAL

Aunque es de sobra conocido, no está de más recordar que la aparición de la microelectrónica, la popularización de los ordenadores personales y el desarrollo de las redes telemáticas a partir de mediados de los años 1980 ha supuesto una revolución en el proceso creador, editor y distribuidor de toda clase de documentos afectando a la educación y a la cultura y, especialmente, a la industria de los bienes culturales. Este cambio se incrementa a medida que se mejoran las capacidades y funcionalidades de los sistemas informáticos y de telecomunicación, se expande la denominada "alfabetización digital" y se reducen los costes de producción y venta.

El resultado de la utilización de esta tecnología es un conjunto de objetos digitales con un contenido informativo, formativo, recreativo, artístico, administrativo, etc. merecedor de ser conservado, sobre todo cuando únicamente existe en formato electrónico.

El valor de estos objetos ha sido reconocido, además de por organismos profesionales nacionales e internacionales, por la UNESCO y por la Unión Europea que promueven acciones para conservar lo que se denomina patrimonio digital con el objeto último de que pueda ser consultado a lo largo del tiempo.

CLASES DE PATRIMONIO DIGITAL

Dentro del patrimonio digital y con independencia del tipo de información o de la forma de publicación, se suelen diferenciar dos clases de patrimonio digital atendiendo a su origen: 1) el creado directamente de forma digital al usar las tecnologías de la información y documentación y 2) el resultante de la digitalización de materiales analógicos (manuscritos, impresos, imágenes fijas o en movimiento, etc.)

OBJETO DE LA INTERVENCIÓN

La intervención pretende suscitar el interés por la recopilación, preservación y conservación de los originales de un escritor y de sus documentos particulares (archivos personales) creados de forma digital como fuentes para el estudio y la investigación actual y de un inmediato futuro. No contempla los archivos personales convencionales ni el proceso de su digitalización con fines de conservación y de difusión como los existentes, entre otras instituciones, en la Residencia de Estudiantes

LOS ORIGINALES ELECTRÓNICOS DE UNA PUBLICACIÓN IMPRESA

En la actualidad, la gran mayoría de los escritores redactan sus obras empleando un sistema informático. Cuando finalizan su redacción, proporcionan al editor una copia en soporte informático lo que contribuye a agilizar el proceso de impresión. Esta copia equivaldría al original manuscrito o mecanoscrito o manuscritos mecanografiados de otras épocas.

Durante el proceso de creación, el autor puede actuar con el original electrónico, al menos; de dos maneras:

1. Guardar cada versión o estadio de una obra en un fichero informático y entregar al editor la copia de la última versión, original o "texto definitivo".
2. Escribir el texto en un único archivo sobre el que va modificando el contenido y del que entrega al editor una copia cuando da por finalizado el original.

El primer procedimiento permite conocer la evolución del acto de creación y la evolución estilística de un texto desde su comienzo.

Una vez impresa la obra, el autor, de acuerdo con lo establecido con la editorial, todavía puede introducir cambios en las galeradas y primeras pruebas de imprenta. Por lo general, ni el autor ni la editorial suelen grabar estos cambios en sus respectivas copias informáticas del original.

Cuando se publica el texto, una copia del original en soporte informático se conserva en los archivos de la editorial a la espera de una posible reimpresión y otra en el domicilio del autor hasta la desaparición de una y otro.

A partir de este momento es cuando entra en juego la incertidumbre. Dejando a un lado las técnicas de

conservación a medio y largo plazo (refresco de soportes, transferencia a formatos normalizados, condiciones medioambientales adecuadas a los soportes digitales...), hay algo tan sencillo, como las averías graves de los equipos informáticos, la certeza de que el autor y / o la editorial realicen copias de seguridad, la obsolescencia de los soportes y de las aplicaciones informáticas, que puede poner en peligro la recuperación y adquisición de los textos originales y de sus variantes por una institución pública interesada en su conservación y acceso por la comunidad investigadora.

De esta forma, las copias informáticas con las distintas versiones de una obra y/o con el texto definitivo pueden alcanzar cierta importancia para el estudio de un escritor y/o la crítica textual y estilística de su obra; son susceptibles de ser equiparadas con el valor cultural y el interés de un manuscrito de Miguel de Cervantes, de Federico García Lorca, de Isaac Newton o Albert Einstein.

Así pues, el patrimonio bibliográfico digital se origina al emplear las tecnologías de la información en la escritura o diseño de documentos utilizando programas informáticos aunque luego se publique en un medio analógico como un libro o revista en soporte papel.

LA CONSERVACIÓN DE LOS ORIGINALES ELECTRÓNICOS

Hasta el presente, que yo sepa, ninguna institución pública o privada se ha preocupado por la recopilación y conservación de los originales electrónicos o "manuscritos digitales" como los denominan algunos autores. Tampoco se han ocupado de llevar a cabo acciones similares con las obras de arte elaboradas utilizando las tecnologías de la información. En ambos casos se ha conservado el producto final, el libro o el cuadro, pero no los originales ni los programas informáticos desarrollados para componer la representación artística. Ha faltado y continúa faltando perspectiva y sensibilidad respecto a estos bienes culturales, muchos de los cuales se han perdido definitivamente debido a la obsolescencia tecnológica, al fallecimiento de los autores y a la desaparición de las editoriales privadas que las publicaron.

En algunos casos, los libreros se han ocupado de la localización, identificación, adquisición, tasación y oferta en venta de las pruebas de imprenta de obras literarias. Una muestra fue la salida al mercado de las pruebas de una novela de la relevancia de *Cien años de soledad* de Gabriel García Márquez de la que, según el propio autor, no se ha guardado ni el manuscrito ni el mecanoscrito. Sin embargo, la tasación fue tan desorbitada en relación con las novedades aportadas, que no se llegó a adquirir.

Aunque, en España (un buen ejemplo es la de la Residencia de Estudiantes) y en otros muchos países, ya existen bibliotecas digitales, compuestas por materiales analógicos digitalizados y por alguna publicación digital (tesis doctorales, revistas, actas de congresos, etc.), aún se echan de menos las versiones y los originales de los originales electrónicos.

A este respecto, resultaría conveniente que las instituciones culturales (la Residencia de Estudiantes, entre otras), los coleccionistas y/o el sector del comercio del libro antiguo detectaran la relevancia del hecho al que me vengo refiriendo y se encargaran de la adquisición o, en el caso de las entidades culturales, de la gestión de donativos y legados de materiales bibliográficos digitales además de los manuscritos e impresos. Esto supone establecer una red informativa sobre autores y editoriales desaparecidas, en trance de disolverse o en plena actividad. En cualquier supuesto, habría que suscribir acuerdos para la conservación, acceso, consulta y eventual reproducción de los "manuscritos" electrónicos.

Sólo de esta manera se podría recopilar y preservar este subconjunto del patrimonio bibliográfico digital y promover proyectos como *Digital Variants (DV)* ^[1], entre otros, que contribuyan a la comprensión de la génesis textual.

LA COLABORACIÓN CON LOS PRODUCTORES COMO PRIMER PASO PARA CONSERVAR EL PATRIMONIO DIGITAL

La preservación del patrimonio digital debería comenzar desde el momento de su creación. Para ello conviene que los programas de conservación y las instituciones que los promueven y realizan, lleven a cabo campañas de concienciación con los productores (término utilizado para designar los autores, diseñadores, revisores, editores y difusores de los materiales digitales). Estas campañas tendrían como objetivo informarles de las normas y buenas prácticas cuya aplicación facilitaría la utilización, accesibilidad y conservación de los objetos del patrimonio digital.

Un posible planteamiento para conseguir la colaboración de los productores y los programas de conservación del patrimonio digital, consiste en elaborar un código de buenas prácticas. En él se deben definir, de forma consensuada por ambas partes, las funciones, responsabilidades, alcance y las condiciones de una cooperación duradera. Debería establecer, además, directrices y especificaciones escritas, entre otros, sobre los siguientes aspectos:

1. El empleo de soportes informáticos sólidos desde el punto de vista industrial, capaces de subsistir hasta que los datos se transfirieran a otro soporte.
2. La utilización de formatos de ficheros y estándares validados (no variantes que compliquen la conservación) de datos apropiados. Resulta más fácil conservar datos de formatos sencillos que usan programas de fuentes abiertas y de dominio público.
3. La creación de metadatos, que identifiquen, describan y faciliten la recuperación del objeto digital por motores de búsqueda, conforme a sistemas ampliamente aceptado y normalizados, como el Dublin Core.
4. La adopción de medidas de seguridad para proteger los ficheros de daños o pérdidas.
5. La autenticidad del objeto, identificando todos los archivos y documentando su procedencia e historia como prueba de su autenticidad.
6. La protección de la propiedad intelectual permitiendo la realización de ciertas operaciones que aseguren la conservación de los ficheros informáticos (copia a otros soportes, transferencia a formatos y datos normalizados, etc) y que permitan su consulta en determinadas condiciones en los centros depositarios.

LOS ARCHIVOS PERSONALES

El estudio sobre los escritores de obras literarias y los investigadores precisa, además de los originales electrónicos y la versión definitiva impresa, de los documentos que aportan información sobre su biografía. Nos referimos a la correspondencia familiar, con otros autores o instituciones públicas y privadas relacionadas con actividades formativas, culturales o comerciales, sin olvidar la documentación generada desde el nacimiento hasta la defunción de cada individuo y sus relaciones con las

administraciones públicas. Aunque en los archivos personales suele haber una copia en papel de estos documentos, su temática sale del objeto de esta intervención, pues compete a la archivística y a la administración electrónica. Simplemente apuntemos que toda investigación de una persona debe complementarse con la búsqueda y consulta de la documentación existente en los archivos de las entidades públicas.

EL IMPACTO DE LAS TECNOLOGÍAS EN LAS COMUNICACIONES INTERPERSONALES

Aunque todavía se escriben cartas manuscritas, la sucesiva mejora y expansión de la telefonía fija, de la móvil y del correo electrónico ha modificado la comunicación y las relaciones individuales.

El correo escrito cada vez más está siendo sustituido por la conversación telefónica, los mensajes telefónicos, el correo electrónico y la videoconferencia debido a su inmediatez y a la reducción de los costes, si no económicos, sí temporales: ya no hay que aguardar días o semanas para obtener una respuesta.

Los autores continúan escribiendo cartas manuscritas o utilizando un procesador de textos cuyo texto imprimen, firman y remiten por correo postal, pero también emplean estas tecnologías para comunicarse con sus seres queridos, con otros escritores, con sus editores y con otras instituciones públicas y privadas.

En el caso de los documentos confeccionados mediante un procesador de textos, su conservación y las incidencias técnicas sobre la misma, sería análoga a lo expuesto sobre los textos originales digitales. El archivo de una copia en papel no es garantía de perdurabilidad debido a la acidez del papel y a la escasa permanencia de las tintas.

En el supuesto de las comunicaciones por medio de las restantes tecnologías enumeradas, apenas queda constancia del contenido, excepto en el correo electrónico, lo que va a incidir en el estudio de las biografías y en la génesis y desarrollo de las ideas de las personalidades actuales.

LA CONSERVACIÓN DE LOS MENSAJES DE CORREOS ELECTRÓNICOS

La información de los mensajes de correo electrónico es de naturaleza muy diversa. Por lo tanto, lo primero que habría que establecer son unos criterios de selección. Parece claro que el correo basura, los mensajes de prueba y la propaganda no sería necesario conservarlos en el sistema informático de una persona física, aunque tal vez sí en una persona jurídica, como una empresa comercial. En este último caso es información sobre sus productos, sobre el diseño de la publicidad y sobre la cartera de clientes potenciales.

Por otra parte, hay que tener en cuenta las implicaciones legales (derecho a la privacidad y a la intimidad) que afectan al correo electrónico. A ello habría que añadir, que no existe garantía absoluta de que el mensaje haya sido escrito por el remitente, al menos que vaya acompañado de una firma o certificación digital emitida por una entidad legalmente establecida. Finalmente, recordemos el mensaje que emite el sistema informático cuando recibimos el acuse de recibo a un correo electrónico que hemos enviado: *Este acuse de recibo sólo acredita que el mensaje se mostró en el ordenador del destinatario. No hay garantía de que el destinatario haya leído o comprendido el contenido del mensaje.* La buena fe de las personas no tienen en consideración estos aspectos que sí se deberían tener presentes en los archivos personales digitales.

Sobre el correo electrónico también incide la diversidad de tipos de datos y de programas informáticos que gestionan el correo electrónico, así como la obsolescencia tecnológica y la falta de aplicación de los estándares pueden dificultar su conservación. Si se cambia de aplicación, puede que la nueva no sea capaz de leer los correos gestionados por la anterior, excepto si se eligen aplicaciones que utilicen normas aceptadas internacionalmente y programas de fuentes abiertas. En el caso de cambio de programa informático, una posible solución consistiría en guardar los mensajes "antiguos" como archivos de texto, que puedan ser interpretados por un procesador de textos. En este supuesto, se tratarían como un objeto digital textual más a efectos de conservación.

La realidad pone de manifiesto que los particulares no suelen tener presente estos aspectos técnicos y que no suelen guardarlos indefinidamente en el disco duro de un sistema informático o en memorias magnéticas u ópticas externas. En este caso, también cabe recordar que los mensajes de correo electrónico impresos en papel se borran al cabo de cierto tiempo debido a las causas apuntadas más arriba: la acidez del papel y la falta de permanencia de las tintas, máxime si no se guardan en unas condiciones medioambientales adecuadas.

Apuntemos que no hay uniformidad de criterio en las normas y prácticas del depósito legal, el principal procedimiento para recopilar, conservar y facilitar el acceso al patrimonio bibliográfico a lo largo del tiempo, sobre las denominadas "comunicaciones públicas". No obstante, debido a la importancia de este medio de comunicación, existen algunos proyectos e iniciativas para la conservación de los correos electrónicos como :

- *Testbed Digitale Bewaring* de los Países Bajos que ha elaborado un programa informático prototipo basado en una personalización de Microsoft Outlook para permitir la comunicación con un servidor central encargado de la recolección de metadatos y la conversión y almacenamiento de los mensajes y metadatos en XML.
- *DAVID: Digitale Archivering in Vlaamse Instellingen en Diensten* (Archivos Digitales de las Instituciones y Administraciones Flamencas). Este proyecto nace de la colaboración de los Archivos de la ciudad de Antwerp y el ICRI. Su objeto es la investigación de la durabilidad digital en el entorno de los organismos oficiales, así como la elaboración de buenas prácticas para conservar datos electorales y demográficos, correos electrónicos y sitios web.

REFERENCIAS BIBLIOGRÁFICAS

DIFFUSE: Guide to Archiving. http://www.diffuse.org/archive_guide.html

DAVID (Digitale Archivering in Vlaamse Instellingen en Diensten. <http://www.antwerpen/be/david/>

DIGITAL VARIANTS. <http://www.digitalvariants.org/>

ECPA (European Commission on Preservation and Access. Actualizado periódicamente. <http://www.knaw.nl/ecpa/about.html>

E-MAIL-XML DEMONSTRATOR: Technical Description. Testbed Digitale Bewaring. <http://www.digitaleduurzaamheid.nl/bibliotheek/docs/email-demo-en.pdf>

ERPANET (Electronic Resource Preservation and Access NETWORK. Comisión Europea. Actualizado

periódicamente. <http://www.erpanet.org/>

ESPAÑA. Consejo Superior de Informática y para el Impulso de la Administración Electrónica *Criterios de seguridad, normalización y conservación de las aplicaciones utilizadas para el ejercicio de potestades*. - Madrid : Ministerio de Administraciones Públicas, Secretaría General Técnica, 2004. - 212 p.

GARCÍA MELERO, Luis Ángel *Algunas iniciativas y técnicas relacionadas con la recopilación y conservación del patrimonio bibliográfico digital*. - En: Foro Biblioteca y Sociedad (2004. Murcia) Actas del Foro Bibliotecas y sociedad: experiencias de innovación y mejora: Murcia, 14 a 16 de octubre de 2004. - Madrid: Confederación de Asociaciones de Archiveros, Bibliotecarios, Museólogos y Documentalistas (ANABAD), 2004. - P. 429-456

GARCÍA MELERO, Luis Ángel *La recopilación y conservación del patrimonio digital*. - En: Jornadas sobre bibliotecas nacionales (2005. Valencia) Las bibliotecas nacionales del siglo XXI: Biblioteca Valenciana, 18, 19, 20 y 21 de mayo de 2005. -<http://bv.gva.es/documentos/Ponencias/Melero.pdf>

GUIDELINES for digitization projects for collections and holdings in the public domain, particularly those held by libraries and archives. - Paris: UNESCO, 2002. Existe una traducción al español, realizada por el Grupo de Trabajo de Colecciones Digitales de las Comunidades Autónomas y el Ministerio de Cultura, disponible en http://travesia.mcu.es/documentos/pautas_digitalizacion.pdf.

GUIDELINES for the preservation of digital heritage / prepared by the National Library of Australia. - [Paris]: Information Society Division, United Nations Educational, Scientific and Cultural Organization (UNESCO), 2003. - También disponible en <http://unesdoc.unesco.org/images/0013/001300/130071s.pdf>

IFLA / IPA Steering Group *Preserving the Memory of the World in Perpetuity: a joint statement on the archiving and preserving of digital information*. <http://www.ifla.org/V/press/ifla-ipa02.htm>

INTERNATIONAL COUNCIL ON ARCHIVES. Actualizado periódicamente. <http://www.ica.org/>

INTERNET ARCHIVE. Actualizado periodicamente. <http://www.archive.org/>

LONG Term Preservation Study. Koninklijke Bibliotheek. Actualizado periodicamente. <http://www.kb.nl/kb/ict/dea/ltp/ltp-en.html>

NDIIPP (National Digital Information Infrastructure and Preservation Programmes). Library of Congress. Actualizado periódicamente. <http://www.digitalpreservation.gov/ndiipp>

NHPRC (National Historical Publications and Records Commission. Actualizado periódicamente. <http://www.archives.gov/grants/index.html>

PADI (Preserving Access to Digital Information) Biblioteca Nacional de Australia. Actualizado periódicamente. <http://www.nla.gov.au/padi/>

RESOLUCIÓN 2002 / C 162 /02 del Consejo de 25 de junio de 2002 sobre "Conservar la memoria del mañana. Conservar los contenidos digitales para las generaciones futuras". - En: Diario Oficial de las

Comunidades Europeas C 162, 45º año, 6 de julio de 2002

VÁZQUEZ DE PARGA, Margarita *Documentos electrónicos: estándares para su creación*. - En: Boletín de la Confederación de Asociaciones de Archiveros, Bibliotecarios, Museólogos y Documentalistas (ANABAD), 2003, LIII, n º 4, p. 137-199

[1] . El Proyecto *Digital Variants*, accesible en <http://www.digitalvariants.org/>, es un archivo digital de variantes de autores contemporáneos fundado en 1996 por el Departamento de Italiano de la Universidad de Edinburgo. Colaboran en él las Universidades de Roma, de Girona y la Autónoma de Madrid

Nota: Esta intervención fue expuesta en el Seminario Archivo de la Internet Española organizado por la Residencia de Estudiantes de Madrid <http://www.residencia.csic.es/>

Grupo Dosdoce de Comunicación - ISSN 1697.4891 - Aviso legal - Optimizada en Explorer - 1024x768