

Preservación digital

Conviene comenzar por aclarar a qué nos referiremos en esta unidad con la expresión “preservación digital”. Se habla de la “preservación digital” de objetos culturales y de nuestro patrimonio cuando los “digitalizamos”. Por ejemplo, cuando digitalizamos un manuscrito (lo escaneamos), o incluso cuando hacemos una reproducción con una impresora digital de un objeto antiguo con valor arqueológico. Desde luego, esto también es preservación digital de nuestro patrimonio, pero no es el concepto que ahora nos interesa explicar. El problema del que queremos tratar ahora es el de la preservación digital de los mismos objetos o productos digitales. Lo que aquí nos interesa es la segunda fase, es decir, la preservación de lo digital, de los archivos digitales que implican creaciones culturales nacidas digitales u objetos culturales que se han digitalizado para su preservación. Imaginemos que hemos creado una obra de arte digital, o bien que hemos digitalizado un libro o una pintura, o que hemos creado una compleja base de datos sobre la Guerra Civil, o una edición crítica que combina texto, imágenes y diferentes funciones complejas. Esos objetos digitales (datos conservados en archivos de diferentes formatos, estructuras, software de soporte), en tanto que tales, con su grado de complejidad y arquitectura, deben ser preservados para que puedan seguir consultándose. De esta preservación digital que requieren estos objetos digitales, en su diversidad, es de la que nos ocuparemos en esta sección. Podríamos expresarlo de otro modo: nos ocuparemos de la preservación de LO digital (Ortiz Ancona 2012).

Uno de los grandes retos de los proyectos de Humanidades Digitales es su preservación. No podemos dar por seguro que nuestras investigaciones y nuestros proyectos digitales sigan estando vivos, sigan pudiéndose consultar y utilizar con el paso del tiempo; es algo de lo que nos debemos de ocupar activamente. En el terreno digital, no son sólo centrales los datos, que se pueden conservar en documentos estáticos y cerrados, sino la funcionalidad de páginas web, bases de datos, blogs... Lo que importa no son sólo los datos, sino también su interrelación y presentación digital: ese conjunto es el producto digital. Y el hecho de que esos productos digitales deban seguir siendo funcionales es problemático, porque veces implican estructuras complejas, combinaciones de hardware, formatos y software que pueden caducar o quedar obsoletos. Incluso documentalistas e ingenieros especialistas en este tema, cuestionan la posibilidad de mantener vivos determinados proyectos digitales en un escenario

realista, por ejemplo, sin una financiación específica, o sin la implicación de la persona que los impulsó.

Si visitamos y exploramos la web de la edición crítica digital de La dama boba, de Lope de Vega (damaboba.unibo.it), realizada por Marco Presotto en la Università di Bologna, podremos observar cómo, en ella, no se debe considerar resultado de investigación exclusivamente el texto de la obra literaria, sino también la manera de combinarlo y presentarlo con otros archivos y funciones. Ciertamente, algunos proyectos pueden plantearse como efímeros y pueden conformarse con una existencia garantizada sólo a corto plazo, pero lo más normal es que los proyectos humanísticos, especialmente si son artística o científicamente relevantes, y con mayor razón si han sido financiados públicamente, pretendan e incluso estén obligados a perdurar, a sobrevivir en el tiempo. Para ello es necesario plantearse desde el principio del proyecto sus necesidades y posibilidades de preservación, de seguridad e incluso de mantenimiento (Ortiz Ancona 2012). Aunque preservación, seguridad y mantenimiento no son exactamente lo mismo, en cierto modo están interrelacionados. Por mantenimiento entendemos las necesidades o posibilidades de que un proyecto necesite ser actualizado ya sea desde un punto de vista científico (continuidad en la alimentación o necesidad de actualización de datos) o desde un punto de vista tecnológico (formatos, *software*, *hardware*). Los conceptos son distintos, pero todo aconseja una buena planificación y documentación del proyecto.

Centrándonos en la preservación, ¿podrá preservarse nuestro producto digital por sí solo o será necesario plantearse políticas y estrategias orientadas a tal preservación? ¿Qué podemos hacer nosotros para facilitarla? ¿Quién se va a ocupar de ello? ¿Y quién de mantenerlo activo desde un punto de vista tecnológico o también humanístico si es necesario? ¿Con qué financiación? ¿Podemos confiar su preservación a alguna institución?

EL PROBLEMA: ALGUNOS FIASCOS HISTÓRICOS EN PRESERVACIÓN

En el ámbito de las Humanidades Digitales, uno de los proyectos más ambiciosos, a la vez que uno de los más grandes fiascos fue la “Domesday Machine”. Es historia, aunque haya adquirido rango de leyenda. El hecho no podía pasar desapercibido para filólogos como Daniel Paul O'Donnell [2004], medievalista interesado en recursos digitales cuyo relato seguimos. El “Domesday Project”, dotado de un presupuesto de dos millones y medio de libras, pretendía difundir y celebrar el noveno centenario del Libro de Domesday (una especie de censo de Inglaterra elaborado en el año 1086) en su contexto histórico gracias al uso de la realidad virtual y de sofisticadas máquinas y software especialmente diseñados

para ello (incluso miniordenadores y reproductores de discos láser *ad hoc*). Al final, todo el material quedó obsoleto y resultó prácticamente ilegible en el transcurso de quince años... a diferencia del propio Libro de Domesday, código por el que habían pasado novecientos años, en perfectas condiciones de consulta en la biblioteca del Archivo de Kew. El escándalo estalló cuando se publicó la noticia en *The Observer* en marzo de 2002. Un experto informático declaró, sarcástico, que «tenemos suerte de que Shakespeare no escribiera en un ordenador personal». Ni el peor papel ácido de publicación de novelas de terror o prensa amarilla a inicios del siglo xx ha sufrido ese deterioro, llegó a decirse. Al presupuesto inicial, luego hubo que añadir más proyectos y gastos para poder recuperarlo en labores de lo que también se conoce como “arqueología digital”, que consiste en la recuperación de archivos informáticos caducados o estropeados. Las posibilidades prácticas y los costes económicos de la preservación de los materiales digitales, con todo lo que ello comporta en términos de hardware y software, está todavía por descubrir, según mantenía O'Donnell.

Por supuesto, este problema no afecta sólo a las Humanidades. Otros famosos fiascos la planificación de preservación y seguridad afectaron nada menos que a las imágenes emitidas por la misión Viking enviada a Marte por la NASA, que sólo fue capaz de recuperar tras costosas operaciones y dos años de trabajo de arqueología digital 3.000 de las 56.000 imágenes emitidas. Durante la producción de la película *Toy Story 2*, estuvieron a punto de perder todo un mes de trabajo. Es decir: el diseño de la seguridad y respaldo, la preservación y el mantenimiento digital es un aspecto complejo al que conviene prestarle la máxima atención.

SOLUCIONES: ¿QUÉ PODEMOS HACER NOSOTROS? ALGUNAS DECISIONES Y ESTRATEGIAS BÁSICAS

Si pretendemos preservar nuestro objeto digital a largo plazo, ello afectará, según Suzanne Keene, especialista en conservación digital de la British Library, en el uso de determinados estándares y lenguajes extendidos en comunidades científicas, tecnología de *software*, un protocolo de actualización de dicho software, así como el equipo y la tecnología (también *hardware*) necesarios para conservar ese objeto y esos datos funcionales en el futuro.

O'Donnell propone tres reglas:

1. No escribir para software o hardware específicos ni utilizar lenguajes no-standard.

2. Mantener la diferenciación entre contenidos (los datos) y presentación (la visualización).

3. Evitar innovaciones técnicas si es que éstas son innecesarias.

Podríamos decir que también otros principios, que son propios de las Humanidades Digitales, producto de la experiencia, aconsejan el uso de códigos abiertos y datos abiertos; interesa asimismo la máxima disseminación y réplica de nuestros productos. No se trata solo de copias de seguridad, sino que nuestros proyectos estén alojados y publicados en más de un lugar, en repositorios colectivos y gratuitos. Es decir, si nuestros productos son reutilizados en otros proyectos, eso también ayuda a su preservación. Y además contribuimos de modo colaborativo al arte y la ciencia. También conviene que nuestro proyecto esté muy bien documentado y que detalle a través de sus metadatos todo el proceso de elaboración tanto humanística como tecnológica. Qué software se ha utilizado, en qué versión, qué versiones de nuestros propios objetos digitales y resultados de investigación se han generado, perfectamente identificadas. Se constituirá así una especie de “guía del usuario” que permitirá a otros reconstruir el proceso e incluso la estructura de software que lo sustenta y lo hace funcionar.

SOLUCIONES: ¿QUIÉN NOS PUEDE AYUDAR?

Más allá de las decisiones que nos competen en cuanto a qué lenguajes más extendidos en comunidades científicas, qué estándares, qué *software*, etc., lo cierto es que la preservación digital de nuestros proyectos no depende sólo de nosotros y será conveniente contar con todos los apoyos que podamos obtener. Estamos todavía en un momento de definición de necesidades y roles y no siempre está claro quién nos puede o nos debe ayudar y, a veces, definir una planificación de preservación digital con implicación de instituciones supone un camino por recorrer juntos. Igual que es clave elegir formatos, *software* y *hardware* para una buena preservación digital, conviene aclarar qué organizaciones, instituciones o qué contempla en torno a ella nuestro marco legislativo.

Existen organizaciones e instituciones especializadas e interesadas en la preservación digital. Conviene tener muy presente que esas organizaciones e instituciones públicas pueden e incluso deben, en el caso de estas últimas, prestar apoyo, por ley, a la preservación de materiales digitales de interés cultural.

Desde el ámbito de las organizaciones se puede mencionar, por ejemplo Internet Archive, con sede en Estados Unidos, en California, o la Digital Preservation Coalition, en el Reino Unido. Estas instituciones no sólo ofrecen repositorios para

objetos digitales con enorme capacidad y servicio gratuito, sino que también ofrecen manuales, como el *Digital Preservation Handbook*. Plataformas como GitHub, por otro lado, permiten conservar gratuitamente toda la documentación de proyectos digitales.

Pero posiblemente el primer paso y más razonable sea dirigirse a la institución u organización más próxima. La preservación de objetos artísticos y relacionados con el patrimonio cultural está comenzando a ser asumida por las bibliotecas. Puede ser una biblioteca o centro municipal con MediaLab, o la institución u organización más cercana con intereses en el ámbito digital. Si perteneces a una universidad, la biblioteca universitaria debería asumir esa obligación y ser capaz de orientarte y prestarte ayuda.

En este sentido, también conviene que conozcas cuál es tu contexto de infraestructuras y legislativo. Por ejemplo, si desarrollas investigación en Humanidades Digitales en España o sobre su patrimonio cultural, conviene que sepas qué instituciones y organismos pueden prestarte apoyo. En Cataluña, por ejemplo, existe el programa PADICAT, o en el País Vasco, OndareNet. Y también conviene que sepas que la ley del depósito legal, de ámbito estatal, obliga a las bibliotecas a preservar tu producto en cuanto se le dé un ISBN. Según el Real Decreto 635/2015, por el que se regula el depósito legal de las publicaciones en línea: “Serán objeto de depósito legal, junto con los metadatos que incluyan, todo tipo de sitios web y las publicaciones en ellos contenidas –tanto de acceso libre como restringido–; cualquiera que sea el procedimiento de producción, edición o difusión; cualquiera que sea el soporte o medio no tangible por el que sean distribuidas o comunicadas; cualquiera que sea la localización física del servidor o servidores a partir de los cuales se difunden a las redes electrónicas; y cualquiera que sea el dominio que albergue la publicación; siempre que contengan patrimonio bibliográfico, sonoro, visual, audiovisual o digital de las culturas de España; y siempre que cumplan alguna de las condiciones siguientes:

- a. Que estén en cualquiera de las lenguas españolas oficiales;
- b. Que estén producidas o editadas por cualquier persona física o jurídica que tenga su domicilio, residencia o establecimiento permanente en España;
- c. Que estén producidas o editadas bajo un nombre de dominio vinculado al territorio español.”

Bibliografía de referencia

Corrado, Edward M., y Heather Mouilaison Sandy, *Digital Preservation for Libraries, Archives, and Museums*, Lanham (Maryland), Rowman & Littlefield, 2017 (2ª ed.; 1ª ed. 2014).

Digital Preservation Coalition, *Digital Preservation Handbook*, Digital Preservation Coalition, 2015 (2ª edición).

< <http://handbook.dpconline.org/> >

Keene, Suzanne, «Now you see it, now you won't. Preserving digital cultural material», en D. Bearman y J. Trant, eds. *Museums and the Web 2002*.

Archives & Museum Informatics, Toronto, 2002.

< <http://www.slideshare.net/ashtreex/now-you-see-it-now-you-wont> >

Mc Kie, Robin, & Vanessa Thorpe, "Digital Domesday Book lasts 15 years not 1000", *The Observer*, 3 de marzo de 2002.

< <https://www.theguardian.com/uk/2002/mar/03/research.elearning> >

O'Donnell, Daniel Paul, «The Domsday Machine, or, "If you build it, will they still come ten years from now?" What Medievalists working in digital media can do to ensure the longevity of their research», *Electronic Medievalia*, 7, 2004.

< <http://www.mun.ca/mst/heroicage/issues/7/ecolumn.html> >

*Ortiz Ancona, Dante, «Preservación Digital: problemáticas, estrategias, metadatos, infraestructura y políticas», brevísimo manual inédito de acceso libre depositado en ELIS, E-prints in Library & Information Science, 2012.

< <http://eprints.rclis.org/18590/1/PreservacionDigital2012.pdf> >

Enlaces de interés

Digital Preservation Coalition: < <https://www.dpconline.org/> >

GitHub: < <https://github.com/> >

Internet Archive: < <https://archive.org/> >

How Toy Story 2 Almost Got Deleted: Stories From Pixar Animation: ENTV

< https://www.youtube.com/watch?v=8dhp_20j0Ys >

Museums and the Web

< <https://www.museumsandtheweb.com/> >

Ondarenet: < <http://www.ondarenet.kultura.ejgv.euskadi.eus:8085/ondarenet/> >

PADICAT: < <https://www.padicat.cat/es> >

Panzarino, Matthew, “How Pixar’s Toy Story 2 was deleted twice, once by technology and again for its own good”.

< <https://thenextweb.com/media/2012/05/21/how-pixars-toy-story-2-was-deleted-twice-once-by-technology-and-again-for-its-own-good/> >

Real Decreto 635/2015, por el que se regula el depósito legal de las publicaciones en línea:

< <https://www.boe.es/boe/dias/2015/07/25/pdfs/BOE-A-2015-8338.pdf> >

Vega, Lope de, *La dama boba: edición crítica y archivo digital*. Bajo la dirección de Marco Presotto y con la colaboración de Sònia Boadas, Eugenio Maggi y Aurèlia Pessarrodona. PROLOPE, Barcelona; Alma Mater Studiorum - Università di Bologna, CRR-MM, Bologna, 2015. ISBN 9788898010257.

doi:10.6092/UNIBO/LADAMABOBA

< <http://damaboba.unibo.it/> >

Autor: Ramón Valdés
Departamento de Filología Española
UAB

