

Cuánto cuesta conservar documentos electrónicos: ¿Es posible saberlo?

José Ramón Cruz Mundet, Carmen Díez Carrera

UNIVERSIDAD CARLOS III DE MADRID
DEPARTAMENTO DE BIBLIOTECONOMÍA Y DOCUMENTACIÓN

Resumen

El objeto central de esta ponencia es argumentar las dificultades existentes en el cálculo de los costes de la preservación digital, dada la ausencia de un modelo válido. Se analiza la situación actual de la conservación digital y los factores de costes sobre los que existe general acuerdo, para concluir con la incertidumbre que, también en este aspecto, caracteriza al entorno digital.

Palabras clave: Preservación digital; Costes; Modelos de cálculo de costes; Documentos electrónicos; Conservación

Cita recomendada: Cruz Mundet, J. R., y Díez Carrera, C. (2018). Cuánto cuesta conservar documentos electrónicos ¿Es posible saberlo?. En *El archivo electrónico en la administración digital: 23 Jornadas de Archivos Universitarios, 21-23 de junio de 2017, A Coruña* (pp. 75-82).

DOI capítulo: <https://doi.org/10.17979/spudc.9788497496803.075>

DOI libro: <https://doi.org/10.17979/spudc.9788497496803>

Este texto pretende recoger de la mejor manera posible las palabras pronunciadas en las XXIII Jornadas de Archivos Universitarios organizadas por la Conferencia de Archiveros de Universidades de la CRUE y la Universidade da Coruña. Tales palabras es imposible recordarlas con fidelidad, aunque sí resultará posible acercarse a ellas con la ayuda del hilo conductor de la presentación empleada en los veinte minutos de charla. Por la misma razón, este texto carece de aparato crítico y, en consecuencia, no tiene más crédito que el que las suscribe, en cuanto quien las pronunció, no de su autoría, que lo fue al alimón con la investigadora que acompaña en el encabezamiento y es forzoso reconocerlo en honor a la verdad. Los organizadores se pusieron primeramente en contacto con la profesora Díez, y al coincidir las fechas con otro compromiso académico, lo hicieron con el que encabeza estas líneas.

Queremos expresar, en primer lugar, nuestro agradecimiento a la Conferencia de Archiveros de Universidad de España (CAU), por la gentileza de invitarnos a participar una vez más en sus jornadas, y por darnos la oportunidad de compartir unos días con tan queridos colegas.

1. En el último cuarto de siglo

En el último cuarto de siglo se viene produciendo un *crescendo* en la preocupación de cuantos nos vemos afectados, en un sentido o en otro, por el tema (sector público, empresas, especialistas, instituciones de memoria, investigadores, etc.) acerca de si es posible abrigar la certeza de que conservamos, y lo haremos en el futuro, documentos digitales, de la misma manera que los analógicos. La respuesta es, hoy por hoy, negativa. En efecto, aunque cada vez son mayores los esfuerzos por conservar un

patrimonio documental digital a largo plazo, y todo parece indicar que el resultado se verá algún día coronado por el éxito, la verdad es que por el momento lo es sólo en parte.

En el último cuarto de siglo se viene produciendo un *crescendo* en la preocupación de cuantos nos vemos afectados, en un sentido o en otro, por el tema de los costes de la preservación digital, pero en esta ocasión el grupo se reduce notablemente, en el sentido de que casi nadie parece preocupado por saber si es posible conocer con certeza cuánto cuesta guardar para siempre algo (documentos digitales) que parece incompatible con el propio objetivo: siempre. En cualquier caso, la respuesta es rotunda, y no está flotando con el viento. No, no es posible hoy por hoy saber a ciencia cierta lo que nos cuesta conservar, no ya permanentemente, sino a largo plazo, los documentos electrónicos.

En el último cuarto de siglo se viene produciendo un *crescendo* en la preocupación de cuantos nos vemos afectados, en un sentido o en otro, aunque seamos menos aún, por el tema de que estamos obligados a conservar los documentos electrónicos, con independencia de lo que cueste; de modo que la preocupación aumenta por la imposibilidad o la antítesis: no se puede sostener una actividad a cualquier coste.

2. De dónde venimos y adónde vamos

Este tema cae de lleno en las preguntas de la vida, que se atribuyen al genial Woody Allen, aunque sea más que dudosa su autoría: de dónde venimos, a dónde vamos... y cuánto cuesta.

Venimos de un mundo que llaman analógico, de lámparas y tubos de rayos catódicos que tardaban minutos en reproducir una imagen, un mundo lento pero lleno de certezas. Y vamos hacia un mundo digital, de

inmediatez y también de incertidumbre. En el ámbito archivístico, venimos del papel como soporte por antonomasia de los documentos, indisociable de su contenido, accesible, disponible, legible a ojos vista sin más intermediarios, razonablemente fácil de guardar por siglos. Y vamos a lo digital, entorno donde parecen diluirse los documentos, primero legibles por máquina, luego electrónicos, más tarde digitales, y finalmente subsumidos en el concepto omnicomprendivo, por no decir ambiguo, de activos digitales.

Por lo que nos toca ahora, los activos digitales deben preservarse de modo que permanezcan siendo auténticos, íntegros, fiables, seguros y accesibles. Al menos en lo que respecta a los documentos, condiciones también exigibles a los analógicos, donde la custodia segura preserva la autenticidad, la integridad y, en consecuencia, la fiabilidad de los documentos archivados, que al estarlo son accesibles. Permítasenos un inciso acerca del concepto de documento, que en ocasiones parece que quiere ser dado de baja de la lista. La humanidad conserva información de todo tipo y, siendo digital, lo hace con los mismos medios, con independencia de su naturaleza y finalidad. Pues bien, dentro de ese magma de información, hay una parte que cumple con tres requisitos inexcusables y simultáneos: es una información interna, producida como resultado de procesos que sustentan la actividad de su productor; es además previsible, por cuanto se sabe con antelación el tipo de información que va a contener; y está regulada, pues es resultado de actividades sujetas a normas, como lo está la forma de estructurarse esa información, su régimen de uso, acceso y conservación. Pues bien, esa información digital interna, previsible y reglada es el documento... digital, sea suelto o agrupado en un expediente, como en el papel, o en base de datos como lo ha sido hasta hace poco el formato libro. La contabilidad, los censos, listas... se asentaban en libros, como ahora lo hacen en bases de datos.

El reto de la preservación digital y, por consiguiente, el de sus costes, se debe a cambios formales y también en la naturaleza de la información. Ya hemos hecho referencia a los derivados de los medios, veamos sucintamente otros de naturaleza más profunda, si cabe. Uno de ellos es la contradicción en los términos que existe entre conservación permanente y medios tecnológicos efímeros por naturaleza. Si permanecer es, según la primera acepción del Diccionario de la Real Academia Española (DRAE), mantenerse sin mutación en un mismo lugar, estado o calidad; y su adjetivo, permanente, lo es sin limitación de tiempo, está claro que resulta imposible respecto de los documentos digitales. ¿Qué es lo que cabe entonces? Creemos honestamente que preservar, o sea, resguardar anticipadamente los documentos electrónicos para evitar daños y peligros. Y una vez dados a la luz, lo que cabe es curar, en el sentido que le da el verbo intransitivo (13.^a acepción del DRAE) y en nuestro caso: cuidar de los documentos, poner cuidado en que sigan siendo los mismos, aun habiendo sufrido alguna mutación, y sobre todo que sigan siendo accesibles. Aspecto este en el que la diferencia papel-digital resulta patente. Y este es el fondo del reto digital, mantener los documentos accesibles en el tiempo y que, aun habiendo variado su estado o calidad en algo, continúen siendo auténticos, íntegros, fiables y seguros.

Y estamos obligados a hacerlo en un entorno en el que el factor tiempo, además de gran enemigo de la permanencia, incrementa de manera imparable el volumen a conservar; de modo y manera que a día de hoy la humanidad produce cuatro veces más información de la que puede contener en sus sistemas... lo que conduce a una vida media decreciente de la información.

Se da la paradoja de que la gestión de unos documentos producidos, usados y conservados por medios tecnológicos, sigue haciéndose de modo manual en gran medida. Sí, los procesos de organización, captura, transferencia y gestión en los repositorios digitales son altamente dependientes de la acción humana directa, hasta el punto de que constituye uno de los principales factores de costes y, por si fuera poco, una de las variables que más hace honor a su apellido. Así, *mutatis mutandis*, nos vamos acercando al objeto de esta disertación.

3. Cuánto cuesta

Desde una fecha tan lejana como 1998, en que Beagrie-Greenstein y Hendley trajeron a colación el planteamiento de los costes, hasta 2015 en que se publicó el modelo 4C (Colaboration to Clarify the Costs of Curation), último de los conocidos, ha habido al menos diez intentos serios y cabales de modelizar los costes de la preservación digital... con otros tantos resultados distintos.

Pero ¿es que resulta tan difícil ponerse de acuerdo para medir una misma realidad? La respuesta es sí, más que difícil parece imposible, porque la realidad es variada. No es lo mismo conservar datos científicos, por más estructurados que estén, que documentos dependientes de su estructura para surtir efectos. Por otra parte, sabemos lo que cuesta almacenar en la nube, por byte o por petabyte, como sabemos que no es lo mismo si la nube es propietaria o un servicio, como que no es igual hacerlo en local que en una red LOCKSS, por ejemplo.

En lo que hay consenso es en los factores de costes, aunque a partir de aquí su peso y valor en el cálculo varían enormemente, como ya hemos tratado en diversos trabajos publicados en los últimos dos años. Dichos factores son, escuetamente expresados:

- 1) los costes de desarrollo o de puesta en marcha y el plan de amortización;
- 2) la infraestructura tecnológica;
- 3) los costes operativos no laborales (instalaciones, suministros, mantenimiento...);
- 4) los costes laborales: directos e indirectos;
- 5) los inherentes a la adquisición o, en nuestro caso, creación de la información;
- 6) los derechos de propiedad;
- 7) los seguros;
- 8) la gestión de riesgos;
- 9) las pérdidas (datos, documentos); y
- 10) los costes inherentes a la actividad específica: identificación y selección; transferencia y control; conversión de formatos; descripción/incorporación de metadatos; suministro de medios de acceso; generación, almacenamiento y servicios de copias de acceso.

A partir de aquí, las piezas encajan de forma que dan lugar a distintos y hasta cierto punto incompatibles modelos de costes. La dificultad para llegar a una conclusión de validez general tiene diversas caras. Hay pocas experiencias todavía y, por lo tanto, pocos datos. Además, los datos son difíciles de extrapolar por su diversa naturaleza, de modo que no es lo mismo conservar datos que imágenes, documentos que libros. Los costes dependen también del rango de servicios que se preste. Todo coste debe ser contrastado con su beneficio y, en este caso, muchas veces los beneficios no se pueden monetizar. Sin olvidar que se trata de una actividad de alto riesgo, asociada a elevadas probabilidades de fallos y

pérdidas. Por último, se trata de una actividad, la conservación, carente de incentivos económicos, de modo que mientras producir sea más barato que conservar, la preservación será difícilmente sostenible.

¿Hay, para concluir, lugar para la esperanza? Según los más optimistas para el 2020 –¡ah, las fechas emblemáticas!– habremos alcanzado un modelo de costes operativo, aunque posiblemente lo que suceda, por lógico y también deseable, es que, en un problema derivado de la tecnología, sea esta quien lo resuelva. En tal sentido, no parece descabellado aventurar que el desarrollo de la inteligencia artificial y de los sistemas expertos permita adquirir la capacidad de hacer permanente lo que hoy resulta efímero.