

Gestión de los documentos digitales: estrategias para su conservación

Por Jordi Serra Serra



Jordi Serra Serra

Resumen: La voluntad de conservar de forma permanente los documentos digitales ha estado presente desde los primeros tiempos de la revolución informática. Sin embargo, han sido pocas las experiencias de crear archivos digitales, debido principalmente a la dificultad para encontrar soluciones técnicas a los problemas de obsolescencia, al alto coste económico y a la ausencia de instituciones que asuman esta responsabilidad. Este artículo analiza las estrategias de conservación aplicadas hasta el momento, y reflexiona sobre los planteamientos metodológicos que han inspirado las mismas, las tendencias actuales y su viabilidad futura. A continuación expone las dos opciones para identificar el documento digital y establecer sus límites, operación previa a la aplicación de cualquier estrategia de conservación. Finalmente describe cómo se puede facilitar esta identificación aplicando metodologías de evaluación y selección propias de la archivística, y propone las reglas de evaluación documental como la herramienta fundamental que permite definir una política de conservación de los documentos digitales.

Palabras clave: Documentos digitales, Archivos digitales, Preservación digital, Documentos electrónicos, Evaluación y selección, Expurgo.

Title: *Electronic records management: strategies for long term preservation*

Abstract: The idea of long term preservation of electronic records has been present since the early days of the computer revolution. However, there have been few documented experiences of creating digital archives, mainly due to the difficulty of finding technical solutions to the problems of obsolescence, high costs and a lack of institutions able to assume this responsibility. This paper begins by analysing the preservation strategies applied to date and reflecting upon the methodological underpinnings of these strategies, current trends and future viability. Next it

presents two options for identifying electronic records and defining their limits, a necessary step prior to applying any preservation strategy. Finally it describes how archival appraisal can facilitate this identification process and proposes appraisal schedules as the main tools for defining an electronic records preservation policy.

Keywords: Digital documents, Digital files, Digital archive, Digital preservation, Electronic records, Appraisal, Appraisal schedules.

Serra Serra, Jordi. "Gestión de los documentos digitales: estrategias para su conservación". En: *El profesional de la información*, 2001, septiembre, v. 10, n. 9, pp. 4-18.

Introducción

El interés por el archivo de los documentos digitales está creciendo a medida que se generaliza el *e-business*. El comercio electrónico es el hecho que hace más evidente la necesidad de capturar y perpetuar las transacciones realizadas telemáticamente. ¿Qué sentido tiene garantizar la seguridad de una transacción electrónica y la autenticidad de sus intervinientes y de su contenido, si esta información no puede ser utilizada como evidencia un tiempo después? La firma digital ha sido la llave que ha hecho posible la expansión del comercio electrónico, pero al mismo tiempo ha subido el listón de las dificultades para la conservación a largo plazo de los documentos digitales¹, un campo con un desarrollo, tanto técnico como legislativo, todavía insufi-

ciente. Es paradigmático el hecho de que los autores de la directriz europea sobre la firma electrónica tan sólo regularon su utilización, pero obviaron que los documentos firmados digitalmente requieren unos procedimientos controlados de archivo cuando deben conservar su valor legal durante todo el tiempo que la legislación estipula². El *Real decreto ley 14/1999*³, que desarrolla la directriz europea, tampoco recoge esta necesidad, con más gravedad por cuanto que desde la Ley 30/1992 la conservación se considera una garantía del valor legal de los documentos digitales⁴.

El objetivo de este artículo es comentar las diferentes tendencias dentro de las estrategias para la conservación a largo plazo de los documentos digitales.

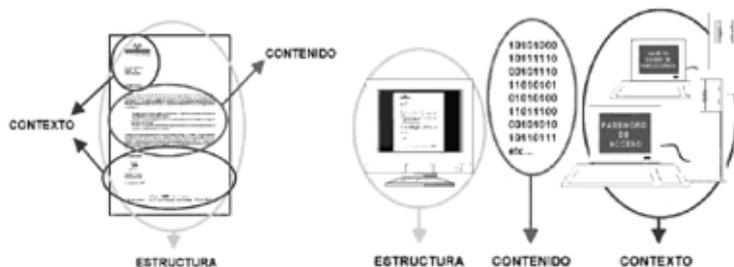


Figura 1. El documento digital ha perdido la unidad física del documento tradicional

Iniciativas existentes

Son numerosos los proyectos internacionales que están estudiando los varios aspectos que condicionan la conservación de los documentos digitales y de su valor evidencial. Sin ánimo de exhaustividad, destacamos en primer lugar el proyecto *Interpares* (*International Research on Permanent Authentic Records in Electronic Systems*), compuesto por seis grupos de investigación de Canadá, USA, Norte de Europa, Italia, Australia, y Este de Asia, y que tiene como base el estudio *The preservation of the integrity of electronic re-*

ords realizado entre 1994 y 1997 por la *Universidad de British Columbia*. Otra iniciativa a destacar es el *Cerar* (*Center for Electronic Recordkeeping & Archival Research*), que recoge los trabajos de la *Universidad de Pittsburgh* (*Functional requirements for evidence in recordkeeping*), puestos en práctica en el *Philadelphia electronic records project* (*Perp*). En el Reino Unido, el *Public Record Office* desarrolla desde 1995 el proyecto *Eros* (*Electronic records from office systems*). El organismo *National Archives of Australia* dispone de una sección dedicada a los documentos digitales, que experimenta con su gestión y publica interesantes trabajos. También es importante el *Dutch Digital Repository* (*Digiduur*), un proyecto holandés que tiene por objetivos adquirir experiencia concreta en la preservación de los documentos digitales y conseguir una toma de conciencia a nivel político.

Conceptos básicos

La conservación de los documentos digitales presenta una doble problemática⁵. En una primera fase es necesario crear y mantener documentos activos y semiactivos fiables y auténticos. Esto significa preservar los valores primarios de los documentos, es decir el valor del documento como evidencia de un hecho a efectos de su utilización dentro de un procedimiento administrativo o proceso de negocio⁶. En una segunda etapa se deben conservar los documentos inactivos auténticos por sus valores secundarios, es decir por el valor intrínseco de la información que contienen⁷.

«¿Qué sentido tiene garantizar la fiabilidad de una transacción electrónica y la autenticidad de sus intervinientes y de su contenido si esta información no puede ser utilizada como evidencia un tiempo después?»

En el contexto de la preservación de los valores primarios la problemática va asociada a la dificultad para identificar el documento electrónico, y al establecimiento de sus límites y partes componentes (figura 1). El objetivo es conservar el documento como una entidad intelectual u objeto lógico, no como entidad física⁸. En el entorno digital la mayoría de los métodos de preservación de la integridad que se aplican (firma digital, contraseñas, *watermarks* o filigranas, etc.) están basados en una concepción unitaria del documento. Sin embargo los estándares *de facto* basados en

Universidad de British Columbia

Propone un modelo de actuación que tiene por objetivos:

- Establecer cómo reconocer e identificar un documento en el entorno digital.
- Determinar qué tipos de sistemas informáticos generan documentos.
- Formular los criterios para segregarse los documentos del resto de conjuntos de datos generados por los sistemas informáticos.
- Definir los requerimientos conceptuales para garantizar la fiabilidad y la autenticidad de los documentos en los sistemas electrónicos.

Este modelo parte de la definición de un documento tradicional completo (que tiene todas las partes: contenido, estructura y contexto administrativo y documental), fiable (que su contenido es cierto) y auténtico (que es original, si ha sido transmitido de forma controlada y conservado y custodiado sin alteraciones). Esta definición se extrapola a lo que debe considerarse un documento electrónico completo, fiable y auténtico, al que se añade un contexto tecnológico. El modelo establece tres espacios de trabajo (individual, de grupo y general o corporativo), y define los estados del documento electrónico en relación con estos espacios. Un documento generado en el espacio individual es un borrador, y sólo se convierte en original cuando atraviesa una frontera electrónica. Es a partir de este momento cuando se debe controlar cualquier modificación para que el documento se considere auténtico.

sgml, firmes candidatos a formatos de conservación a largo plazo, se basan en un modelo de documento **compuesto** de diferentes documentos o partes de ellos, **distribuido** entre diferentes redes y sistemas informáticos, y que puede ser actualizado con gran facilidad y sin dejar rastro (**dinámico**). Podemos tomar como ejemplo xml, en el que las hojas de estilo permiten almacenar la estructura fuera del contenido del documento, e incluso definir una estructura común a un conjunto de contenidos.

En el ámbito de la conservación de los valores secundarios, la preocupación principal es tanto la conservación física de los soportes como la legibilidad de su contenido, un problema que afecta conjuntamente a la documentación administrativa y a la cognitiva. Por esta razón los archivos y las bibliotecas nacionales están trabajando en proyectos de conservación de los soportes y documentos digitales. Es muy interesante consultar los numerosos trabajos publicados por el *Council on Library and Information Resources (Clir)* y las actividades de la *European Commission on Preservation and Access (Ecpa)*, así como la obra: National Preservation Office. *Digital culture: maximising the nation's investment*. Mary Feeney (ed.). London: British Library Board, 1999.

<http://www.clir.org>

<http://www.knaw.nl/ecpa/ecpatex/intro.htm>

A medida que se ha asumido la independencia del documento electrónico de sus distintos soportes físicos ha disminuido la problemática de la conservación física de los soportes puesto que, considerando que una copia en el entorno informático es siempre una copia idéntica, un cambio de soporte físico no implica ninguna transformación en el documento. Ciertamente sigue existiendo el reto de la conservación física de los soportes digitales, pero su tratamiento no es el objeto de este artículo, teniendo en cuenta que la duración de los soportes actuales, tanto ópticos como magneto-ópticos, proporciona un margen de tiempo suficiente como para realizar un recopiado periódico (*refreshing*).

«A medida que se ha asumido la independencia del documento electrónico de sus distintos soportes físicos, ha disminuido la problemática de la conservación física de los soportes»

El problema de la legibilidad está asociado a la obsolescencia tecnológica, tanto de hardware como de software. Como todos hemos experimentado, el plazo

de legibilidad de un documento de uso cotidiano puede oscilar entre los 10-20 años en el aspecto de hardware, y unos 5-10 años en el de software. Si este plazo puede ser admisible para documentos con vigencia limitada, no permite una planificación de la conservación a medio o a largo plazo, por lo que se imponen soluciones previas. Vamos a comentar con más detalle cada una de estas soluciones.

1. Estrategias de conservación

Podemos identificar cuatro métodos para garantizar la accesibilidad futura de los documentos digitales⁹:

1.1. Preservación de la tecnología

La primera solución ha sido conservar la tecnología de la cual depende el documento. Para mantener el contexto tecnológico de accesibilidad de los documentos digitales es necesario “congelar” el estado de la técnica en un momento determinado. Esto implica conservar tanto el hardware como el software originales, y ocuparse de tenerlos en funcionamiento incluso cuando su obsolescencia sea ya completa, lo que en algunos casos ha conducido a la creación de verdaderos museos cibernéticos. Esta opción ha demostrado ser económicamente inviable, y tan sólo es aplicable a corto y medio plazo, o como una medida extrema cuando se trate de una información de gran importancia y no sea posible migrarla a un nuevo entorno tecnológico actualizado.

1.2. Migración de la información digital

Se ha impuesto como la solución más aceptada para la creación y el mantenimiento de archivos digitales. Parte de la idea de que los documentos deben ser accesibles a partir de los sistemas informáticos existentes en cada momento, lo que exige su migración periódica a formatos inteligibles por los sistemas actuales.

Existen diversas estrategias de migración, cada una de ellas adaptada a un determinado tipo de documento digital.

—**Cambio de medio.** Esta estrategia pretende subsanar la obsolescencia física, y consiste en un recopiado periódico de los documentos digitales a soportes más estables (por ejemplo, de soporte magnético a óptico). En su forma más extrema incluye la impresión en papel o la microfilmación de los documentos digitales con tal de conservar la accesibilidad o legibilidad del contenido, aun a costa de perder todas las funcionalidades del documento digital original.



El método *Constance*

En 1978 los archivos nacionales de Francia, a través del *Centre des Archives Contemporaines*, empezaron a trabajar en la conservación de los documentos digitales generados por la Administración Central francesa. Esta actividad se concretó en 1983 con la creación de un archivo digital y de un conjunto de métodos de trabajo que recibió el nombre de *Constance* (*Conservation et stockage des archives nouvelles constituées par l'électronique*)¹⁰. El principal objetivo de este proyecto era la conservación **permanente** de los documentos digitales con valor evidencial, mediante una evaluación previa que eliminaba los documentos sin interés histórico o con información repetitiva.

Inicialmente se usó la cinta magnética como medio de almacenaje, puesto que era el tipo de soporte de uso más extendido. Sin embargo las condiciones de conservación eran exigentes y se hacía necesario un recopiado periódico de las cintas (*refreshing*). Su anunciada obsolescencia obligó, a partir de 1995, a sustituir el soporte magnético por el soporte óptico, iniciándose un laborioso proceso de migración, que finalizó en 1997¹¹. Esta operación se realizó con la previsión de unos diez años de relativa estabilidad, tras los cuales sería necesario iniciar una nueva migración. Pero en 1998 se vio que la opción escogida había sido un error estratégico, y fue necesario volver a migrar toda la documentación a cintas magnéticas DLT 4000.

Un punto importante para el mantenimiento de un archivo digital es la regulación del proceso de transferencia de documentos. Actualmente la operativa del método *Constance* es la siguiente:

—Selección de los documentos a conservar en las mismas oficinas productoras, con el asesoramiento directo de miembros del equipo *Constance*.

—Tratamiento de los documentos transferidos, migrándolos a un formato estándar (preferentemente *ascii*) y elaborando los siguientes elementos descriptivos¹²:

- Del sistema informático o aplicación que genera el documento.
- Del proceso administrativo.
- De la estructura del documento.
- Del contenido del documento.
- Ficha técnica indicando el peso, número de registros, longitud, etc.
- Diccionario de datos (cuando se trata de una base de datos).

—Prueba del documento en su nuevo formato usando la descripción anterior.

—Fijación en un soporte definitivo y archivo seguro del documento, con las correspondientes copias de seguridad.

La mayoría de los documentos conservados son ficheros estadísticos y censales o de gestión de personal a partir de 1960. Los períodos de publicidad y las características del acceso a los documentos archivados son establecidos por la *Commission Nationale pour l'Informatique et les Libertés*. Sin embargo, dado que la mayoría de documentos conservados contienen datos personales que, según la legislación francesa, no pueden ser públicos hasta pasados 100 años, la difusión de los documentos se ha limitado hasta el momento a los usuarios internos de la propia Administración. Paradójicamente es muy difícil justificar la inversión realizada en un archivo histórico cuando existen tales restricciones a la difusión de su contenido.

Las conclusiones del proyecto *Constance* son las siguientes:

—Los costes de mantenimiento de un archivo digital son **imprevisibles**, puesto que dependen de factores externos al archivo, tanto de la evolución tecnológica como de la estructura de las organizaciones que producen los documentos.

—La principal inversión para el mantenimiento de un archivo digital debe realizarse en la formación de los **recursos humanos**, puesto que los problemas tecnológicos serán cambiantes y su resolución dependerá de las habilidades del personal del archivo.

—Las decisiones sobre la metodología y la financiación de un archivo digital público deben ser tomadas al más alto nivel de gobierno.

La principal aportación del proyecto *Constance* es la experiencia obtenida en referencia a los **costes de mantenimiento** de un archivo digital de gran envergadura. Con anterioridad al establecimiento de unos estándares de conservación, la voluntad del *Centre des Archives Contemporaines* era conservar ejemplares de todos los programas y tipos de ordenadores utilizados por la Administración Central francesa, y en esta dirección dirigieron sus esfuerzos. El coste de mantenimiento de este archivo crecía de forma exponencial, y pronto adquirió dimensiones prohibitivas para los fondos públicos que, coincidiendo con un recorte presupuestario (1993), condujeron a la Administración a reducir la plantilla de 15 a 4 técnicos y a limitar los gastos en hardware. El proyecto, que hasta el momento había sido impulsado por el *Ministerio de Cultura*, pasó a depender directamente de los archivos nacionales, y se creó un grupo de trabajo interministerial para su seguimiento. En la actualidad, sin embargo, el coste del archivo todavía es financiado íntegramente por el *Ministerio de Cultura*. El archivo se ha tenido que adaptar a la nueva situación, trabajando sólo con formatos de bajo nivel. Pero incluso con estas limitaciones el mantenimiento de este archivo sigue siendo muy caro, y la cantidad necesaria de personal altamente cualificado muy elevada.

Existe otro factor que aumenta los costes: mientras en un archivo administrativo los fondos documentales pueden mantener estables sus dimensiones gracias al expurgo, en un archivo histórico los fondos crecen constantemente. Por esta razón, considerando que aproximadamente cada diez años será necesario migrar todo el fondo documental a nuevos soportes y formatos, el volumen de documentos para cada migración será mucho mayor y sus costes mucho más elevados.

Center for Electronic Records¹³

Es un organismo dependiente de la *Nara* que tiene como misión garantizar el acceso a los documentos digitales de conservación permanente generados por las oficinas del *Gobierno Federal* de los Estados Unidos. Este archivo digital constituye una vasta fuente de información, al mismo tiempo que sirve de evidencia y testimonio de cómo las oficinas del *Gobierno Federal* han llevado a cabo su acción política y administrativa. El centro recoge todo tipo de documentos generados directamente en formato digital, tales como bases de datos, documentos de texto o sistemas de información geográfica. Los documentos digitales conservados por la *Nara* constituyen asimismo un testimonio de la evolución de la tecnología informática. Los más antiguos datan de la Segunda Guerra Mundial, aunque la mayor parte de los documentos son posteriores a 1960.

El *Center for Electronic Records* ha establecido unas normas para la transferencia de documentos electrónicos a su archivo digital por parte de las oficinas federales. En re-

ferencia al **soporte**, admite tanto cintas magnéticas como cd-rom, y establece las normas de grabación de los datos, pero no contempla la transferencia telemática realizada directamente a los servidores del *Center*. Estas normas especifican que los soportes transferidos pueden contener ficheros de software y de datos en un mismo soporte, pero que los documentos que se consideren históricos deben estar en un soporte que contenga únicamente ficheros de conservación permanente.

En referencia al **formato**, los documentos transferidos deben estar en un formato no dependiente de un hardware o software concretos. El documento debe estar codificado en *ascii* o *ebcdic*, habiendo eliminado de él todos los caracteres de control, a excepción de los delimitadores de campo y registro y de las etiquetas *sgml*. El fichero no puede tener ningún tipo de compresión que no haya sido aprobado previamente por la *Nara*. Los ficheros de base de datos se deben transferir en *ascii*, *sql* o *irds¹⁴*, desvinculados de la tecnología original, únicamente con

las tablas de datos y la documentación de las relaciones, y respetando la primera forma normal. En principio no se contempla la transferencia de bases de datos no relacionales.

Cuando se recibe una transferencia electrónica es necesario un proceso de validación antes de incorporar los documentos al archivo digital. La validación es el proceso de comparar el contenido de los documentos digitales recibidos de una agencia federal con la descripción del contenido de estos documentos que los acompaña¹⁵. Inicialmente la *Nara* realizaba una validación manual, imprimiendo una parte de los datos y comparando su estructura con la expresada en la documentación. Desde 1992 se utiliza la aplicación informática *Aeric (Archival electronic records inspection and control)*, desarrollada en *Oracle*, que analiza y depura las tablas de datos, generando un informe con las anomalías que deben ser solventadas antes de su ingreso en el archivo digital.

—**Compatibilidad retroactiva.** Los programas comerciales permiten a menudo leer documentos creados con versiones anteriores de una misma familia de programas, así como su migración a la versión actual. La compatibilidad retroactiva, sin embargo, está sujeta a los cambios en la estrategia comercial de la empresa productora del software. Además, este tipo de migración puede realizar alteraciones en la estructura interna de los documentos digitales no deseables desde el punto de vista legal, o que atenten a la autenticidad del documento original.

—**Interoperabilidad.** Programas de distintas empresas pueden compartir los mismos formatos de documento digital. Esta estrategia comercial, practicada por los principales productores de software, permite visualizar documentos de un formato propietario con el programa de otra empresa y convertirlos a nuevos formatos propietarios. En este caso se reproducen los mismos problemas que en la compatibilidad retroactiva: la dependencia de las políticas de empresa, y el elevado porcentaje de pérdida de datos.

—**Conversión a formatos estándar.** Podemos considerar esta opción una versión extendida de la interoperabilidad.

La existencia de formatos estándar es la principal garantía de la conservación de la funcionalidad de los documentos digitales. Sin embargo, los formatos estándar válidos como formato final para la conservación permanente no siempre pueden ser un formato de trabajo, como sucede por ejemplo en la edición de documentos con *sgml* o *xml*. Esta estrategia, por lo tanto, debe contemplar la utilización de un formato seguro (fiable y auténtico) durante el uso del documento, cuando tiene valores primarios, e incluir la posibilidad, cuando se quiera disponer el documento para su con-

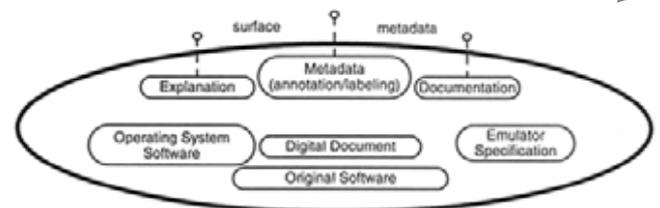


Figura 2. Un documento digital encapsulado

Teoría de la emulación

El uso de emuladores para la conservación permanente de los documentos digitales ha sido propuesto por **Jeff Rothenberg**¹⁶. Esta teoría parte de que la única forma de poder recrear la apariencia y la funcionalidad originales de un documento digital es mediante el software con el que fue creado. La propuesta consiste en utilizar aplicaciones que emulen el funcionamiento de los programas originales cuando éstos sean ya obsoletos.

Sin embargo no tendría sentido conservar emuladores de cada software utilizado, puesto que con el tiempo estos emuladores también serían obsoletos. Lo que se debe conservar asociado al documento es **la información sobre cómo desarrollar un emulador** que pueda reproducir ese documento. Esta fórmula debe ser suficientemente abierta para sobrevivir a la incertidumbre tecnológica de la conservación indefinida.

La propuesta consiste en **encapsular** tres tipos de información con cada documento digital (figura 2)¹⁷:

—La explicación en un formato permanentemente inteligible del software y del hardware emulado,

metadatos sobre el documento, su ciclo de vida y su contexto de creación, anotaciones, etc. Consiste en un fichero *readme*, que contiene las instrucciones para el proceso de abrir y poner en marcha un documento encapsulado.

—El emulador de la plataforma informática original. No consiste en un programa ejecutable, sino en la especificación de los atributos tecnológicos considerados relevantes para recrear el comportamiento del documento original.

—El documento y su entorno de software. Consiste en el documento digital, compuesto de uno o más ficheros en el formato original, más los ficheros que representan la cadena de bits ejecutable original que permitía reproducir el documento, y los ficheros que representan la cadena de bits del sistema operativo que permitía que se ejecutase el software.

En la práctica los elementos comunes a más de un documento pueden ser almacenados de forma centralizada, encapsulando con el documento tan sólo un enlace al elemento correspondiente¹⁸.

El proceso completo comprende cuatro etapas:

—**Anotación:** crear las explicaciones sobre el contexto del documento y sobre cómo abrir y usar la encapsulación.

—**Encapsulación:** construir la estructura lógica que contenga todos los elementos citados.

—**Transliteración:** periódicamente actualizar las anotaciones o explicaciones textuales para mantenerlas inteligibles.

—**Emulación:** cuando sea necesario acceder al documento se deberá abrir la encapsulación, crear el emulador específico, y hacerlo funcionar en el ordenador futuro. Obviamente no será necesario crear un emulador para cada documento, tan sólo para cada tipo o formato de documento.

Para desarrollar este modelo será necesario investigar en el desarrollo de un metalenguaje que sirva para describir las especificaciones del emulador con relación a una plataforma de hardware concreta, así como de técnicas para hacer siempre inteligibles las explicaciones y los metadatos, y de métodos de encapsulación que protejan la información digital de posibles manipulaciones o del deterioro y que aseguren su cohesión¹⁹.

Universal preservation format (UPF)²⁰

La *Wgbh Educational Foundation*, con una subvención de la *National Historical Publications and Records Commission*, ha impulsado el desarrollo de un formato universal de preservación a largo plazo de los documentos digitales que asegure la accesibilidad de un amplio abanico de formatos con independencia de cualquier plataforma.

La idea del formato universal de preservación es la de un contenedor en el que no sólo se almacenan los documentos digitales, sino también toda la información necesaria para poderlos reproducir. Esta idea la encontramos desarrollada en el *Bento container format*²¹, creado por *Apple*

Computer y conocido también como *OpenDoc standard interchange format*. Un contenedor *Bento* es un tipo de fichero que contiene uno o más objetos, junto con la información sobre estos objetos (metadatos). Dicho contenedor permite reproducir e intercambiar documentos entre diferentes plataformas, aunque los objetos que contiene no correspondan a un mismo formato propietario ni hayan sido diseñados para utilizar los mecanismos *Bento*. Los objetos pueden ser simples o complejos, y pueden pesar desde unos pocos bytes hasta 264 bytes. Actualmente el código *Bento* funciona sobre *Macintosh*, *DOS*, *Microsoft Windows*, *OS/2* y dife-

rentes variantes de *Unix*.

El formato de intercambio *Open media framework (OMF)*²², de *Avid Technology*, actualmente un estándar para el intercambio de documentos digitales entre plataformas heterogéneas, ha adoptado el uso de contenedores *Bento*. *OMF* encapsula toda la información necesaria para la edición y la reproducción de los documentos digitales que contiene.

Aunque no se consiga un único formato estándar de preservación, el uso de estos contenedores compatibles con una gran variedad de formatos reduce notablemente la diversidad de los documentos que es necesario conservar.

Tabla 1

	Migración: cambio de medio	Migración: compatibilidad retroactiva	Migración: interoperabilidad	Migración: conversión a estándar	Emulación
Conjuntos de datos				X	
Texto estructurado				X	
Documentos ofimáticos	X	X		X	
Gráficos de datos	X	X	X	X	
Presentaciones	X	X		X	
Imágenes	X	X		X	
Sonido	X	X		X	
Vídeo	X	X		X	
Datos geográficos y mapas		X	X	X	
Publicaciones multimedia interactivas					X

servación a largo plazo, de congelar sus valores primarios en un formato estándar.

A pesar de ser el método más aceptado, la migración presenta numerosos problemas. Migrar la documentación de un archivo digital es caro —en recursos materiales y humanos—, laborioso y lento. También es una tarea peligrosa que comporta un riesgo de pérdida de datos, tanto por errores en la grabación como por incompatibilidades entre formatos. Hay que añadir que el curso de la tecnología y su ritmo de obsolescencia son imprevisibles, por lo que cada nueva migración presenta una problemática distinta.

1.3. Eliminación de la dependencia de los documentos de una determinada tecnología

Este método consiste en eliminar las dependencias no esenciales que los documentos mantienen con los programas con que han sido creados. Estas dependencias pueden ser de dos tipos:

—Funcionalidades para editar o actualizar el documento electrónico (por ejemplo, el sistema gestor de una base de datos, sgbd). Cuando se plantea la conservación a largo plazo de un documento digital auténtico se infiere que debe ser inmodificable. Por lo tanto, no es necesario conservar más que las funcionalidades de visualización y los datos contextuales.

—Vestigios de la tecnología original, como por ejemplo los formularios y listados de una base de datos.

Esta opción necesita una tecnología estándar para la recuperación de la información, una vez eliminada su dependencia original. Hasta el momento se ha utilizado en documentos digitales simples (texto plano o

datos estructurados de bases de datos relacionales), pero puede ser aplicable a documentos más complejos siempre que éstos estén sujetos a normas abiertas y se disponga de las herramientas adecuadas para su migración al formato estándar con la menor pérdida de datos posible.

1.4. Emulación de la tecnología

El objetivo de este método es conservar la funcionalidad que tienen los documentos cuando son reproducidos por el software original. Esta emulación se puede limitar a conservar la presentación visual de los documentos mediante visores (*viewers*) o bien conservando la imagen electrónica del documento, con formatos como pdf o djvu. Pero se puede avanzar hacia una opción más completa, intentando conservar una versión reducida del software operacional y los sistemas originales, o bien desarrollando un **emulador** que reproduzca el funcionamiento de dichos sistemas originales. Esta opción garantiza la autenticidad de los documentos conservados y la fidelidad de la reproducción, pero al mismo tiempo presenta un importante inconveniente: tanto la versión reducida del software original como el emulador también están sujetos a una progresiva obsolescencia, y su mantenimiento y actualización pueden ser muy difíciles, especialmente cuando dependan de una firma comercial que actúe exclusivamente en función de los intereses del mercado.

Cada tipo de documento digital tiene unas características específicas de conservación y, sobre todo, unas necesidades de utilización por parte de la organización productora, que son establecidas durante su evaluación²³. Por esta razón la política de preservación de un archivo digital no debe basarse tan sólo en una de las estrategias citadas. En la tabla 1 se indican los métodos más adecuados para aplicar a cada tipo de documento digital²⁴.

2. Objetos versus procesos

Hasta este momento hemos analizado la problemática de la conservación a largo plazo de los documentos digitales considerándolos unidades mínimamente identificables e independientes. Ya hemos comentado que esta no es, en la mayoría de los casos, la naturaleza real de los documentos digitales. Su carácter compuesto, distribuido y dinámico proviene de la relación que mantienen con los sistemas informáticos que los generan, los cuales a su vez están inextricablemente ligados a los procesos operativos de una organización. En este horizonte metodológico se plantea si la conservación debe **orientarse al documento** como enti-

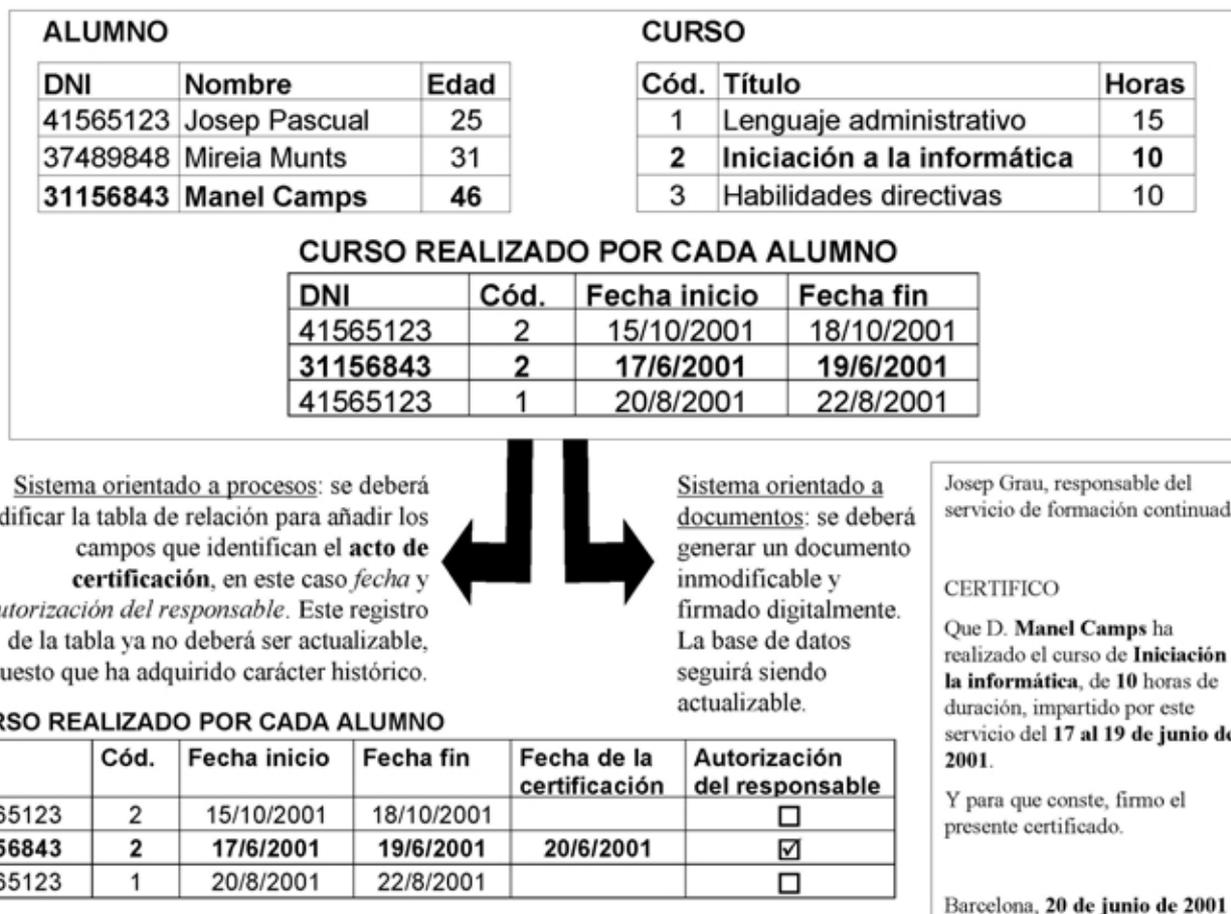


Figura 3

dad unitaria, o si debe seguir un enfoque **orientado a procesos**. Vamos a trabajar algunos ejemplos.

«En el contexto de la preservación de los valores primarios, la problemática va asociada a la dificultad para identificar el documento electrónico, y para establecer sus límites y partes componentes»

El caso paradigmático de este dilema son las bases de datos (bdds), las cuales se pueden concebir como documentos en su conjunto, o únicamente como herramientas generadoras de documentos²⁵. Tomemos el ejemplo de una base de datos relacional típica, que entendemos como un conjunto de datos estructurados en una o más tablas que mantienen entre sí relaciones de distinto grado, formando un núcleo compuesto por el contenido (datos) y la estructura (tablas y relaciones). Estos datos se actualizan mediante diversas interfaces de usuario (formularios), y su explotación se realiza a través de un conjunto de consultas o vistas y formatos de salida (informes). La actualización y la explotación de los datos no dependen de éstos ni de su estructura,

sino del sistema gestor de la bdd que se utilice. Cuando nos planteamos la conservación a largo plazo de los contenidos de una bdd, debemos ser capaces de identificar sobre qué partes de la misma se deben aplicar las estrategias de preservación. El usuario que trabaja con ella accede a los datos mediante un conjunto de formularios e informes, y lo hace de manera que la forma como se estructuran los datos en una interfaz de entrada o de salida no se corresponde con la estructura interna de las tablas, sino con las necesidades del usuario. Por ejemplo, una base de datos que gestiona la formación del personal de una empresa le sirve al usuario para generar los certificados para los alumnos que han asistido a los cursos. En último término, al usuario lo que más le interesa es el conjunto de datos que constituye el certificado de asistencia, documento identificable y al cual podemos atribuir valores administrativos y jurídicos. Los datos de este certificado provienen de las tablas de *alumnos* y de *cursos*, y de la relación “muchos a muchos” *cursos realizados por cada alumno*. Sin embargo, al usuario no le interesa conservar sólo una de estas tablas. Lo que necesita es poder reproducir siempre que sea necesario el agregado de datos que constituye el certificado. Ante esta situación es

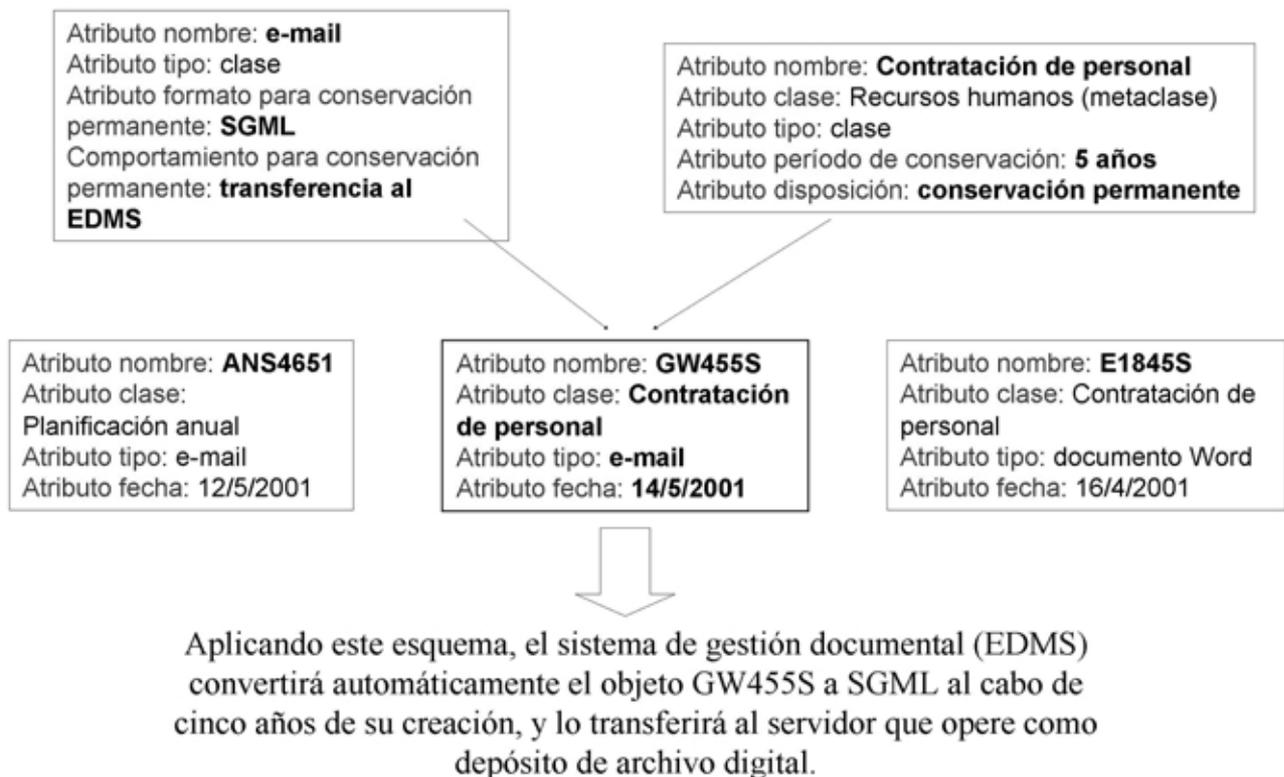


Figura 4

cuando se pueden adoptar dos estrategias distintas (figura 3).

«El problema de la conservación de los documentos electrónicos reside tanto en la diversidad de soluciones tecnológicas como en la falta de cohesión de las soluciones metodológicas que se aplican»

Si nos referimos a una estrategia orientada a documentos, el usuario se conformará con poder convertir el certificado que tiene en pantalla en un documento digital completo, extrayéndolo de la base de datos y archivándolo junto con su contexto de creación mediante un proceso controlado y autenticado. Podría, por ejemplo, convertir el certificado en un documento pdf, con sus correspondientes metadatos, y con algún tipo de certificación digital. Si optamos por una estrategia orientada a procesos, será toda la base de datos que automatiza el proceso de gestionar cursos la que deberá ser conservada, tomando la prevención de que su estructura responda a la necesidad de poder “recrear” el certificado, es decir, protegiendo y fijando los datos necesarios a partir del momento en que adquieren carácter “histórico”. Esta decisión se debe tomar en la fase de análisis y diseño de la base de datos.

El mismo razonamiento podemos hacerlo extensible a todos los sistemas de gestión de documentos: Aunque un sistema *workflow* trabaje orientado a documentos ¿no es la información de los procesos en su conjunto lo que realmente refleja el funcionamiento de la organización? ¿en qué medida un *data warehouse* puede identificarse, por su carácter exhaustivo y acumulativo, con el archivo histórico de una organización?

Los mismos parámetros podemos aplicar al **correo electrónico**. Los mensajes y su contexto residen dentro del sistema de mensajería, pero dicho sistema no puede garantizar su conservación ni funcionar a modo de archivo digital, puesto que la mayoría de programas de correo electrónico están diseñados para un corto período de conservación de los mensajes en los buzones digitales. Por lo tanto, los mensajes deben transferirse al sistema de gestión documental corporativo (*edms*), una vez seleccionados aquellos que tengan valor evidencial²⁶.

Como veremos más adelante, la evaluación es clave para escoger la mejor política de preservación del correo electrónico. En Estados Unidos, cuando el presidente abandona su cargo, todos los mensajes de correo electrónico se transfieren en bloque para su conservación permanente, puesto que toda la correspondencia de la Casa Blanca se considera un documento histórico. En cambio la mayoría de las organizaciones deben adoptar una estricta política de selección que

permita que los mensajes que se consideran documentos evidenciales sean transferidos a un sistema seguro de conservación junto con su contexto de transmisión (registros de fechas, ficheros adjuntos o *attachments*, etc.), considerando el sistema de correo electrónico tan sólo un instrumento para realizar las transacciones telemáticas. La *Universidad de California* incluye los mensajes de correo electrónico en sus reglas de evaluación de conservación (*records disposition schedules*), y transfiere los mensajes con valor evidencial a formatos más estables para su conservación permanente²⁷.

Las administraciones públicas y las empresas utilizan cada vez más internet en sus relaciones con sus usuarios o clientes, así como intranets para el intercambio de información dentro de la organización²⁸. Estos desarrollos tienen un gran impacto en la necesidad

del archivo a largo plazo de las páginas web, del que en otras ocasiones se ha tratado ya²⁹. Desde la utópica propuesta de **Brewster Kahle**³⁰ hasta los enfoques más realistas del proyecto *Pandora*³¹ se ha planteado cuáles son los elementos de una página web que forman parte de una misma unidad documental y que en conjunto pueden reunir el valor evidencial que justifique su conservación³². Las organizaciones deben saber qué ha estado disponible, para quién y en qué momento. **Stephen Harries** define las dos perspectivas para archivar las páginas web:

—**Modelo documento** (u orientado a objetos): una página web contiene elementos que pueden considerarse documentos de archivo. En este caso es preciso evaluar toda la página para seleccionar estos elementos, que se integraran en el *edms* corporativo y serán

Estrategia orientada a objetos

El *National Partnership for Advanced Computational Infrastructure*, con sede en el *San Diego Supercomputer Center (Sdsc)* está desarrollando una parte del proyecto *Distributed object computation testbed (Doct)* de infraestructura de redes de alta velocidad. La *Nara* colabora con el *Sdsc* en el desarrollo de un modelo de preservación de los documentos electrónicos, centrándose en la conservación y el acceso de acuerdo con los estándares archivísticos³⁵.

La metodología orientada a objetos se basa en asignar una serie de propiedades (atributos y comportamientos)³⁶ a cada objeto de un conjunto, y relacionarlos a partir de una determinada propiedad. De esta manera habrá tantas clasificaciones posibles como atributos comunes compartan un determinado grupo de objetos. Estas clasificaciones o **clases** son a su vez objetos, con sus propiedades, que se agrupan en **metaclases**. Todos los objetos que pertenecen a una clase **heredan** automáticamente los atributos y comportamientos aso-

ciados a la misma. Teniendo en cuenta que desde el punto de vista archivístico no se deben capturar solamente los documentos sino también las relaciones entre los documentos de un mismo fondo, la herencia de las propiedades y la jerarquía de clases y subclases se adecua sin problemas a la estructura de un fondo documental.

El primer paso en esta estrategia consiste en identificar qué propiedades tiene cada objeto, y cuáles de éstas se heredarán para cada clase. Y en segundo lugar, se deben aplicar las soluciones siempre a partir del nivel jerárquicamente superior. Por ejemplo, los metadatos sobre el período de conservación de un documento digital deben estar incluidos a nivel de clase. Todos los objetos relacionados con dicha clase heredarán esta información, que determinará su comportamiento cuando se apliquen automáticamente métodos de selección y eliminación sobre un conjunto documental (figura 4).

El sistema incluye dos componentes: un depósito digital,

donde se conservan los documentos y se realiza virtualmente el trabajo archivístico, y un conjunto de herramientas utilizadas para acceder a los documentos.

Durante el proceso de transferencia al archivo digital, la colección o serie se descompone en todos los documentos que contiene, y el contexto y la estructura de cada documento son caracterizados mediante metadatos. Hasta el momento se ha utilizado xml para agregar los metadatos a los objetos, creando una *document type definition (dtd)* para cada clase de objeto identificada durante esta descomposición. Con esta operación la clase se convierte en un nuevo objeto, la *dtd*.

Las colecciones de objetos, encapsuladas en formato xml, son notablemente independientes de la infraestructura informática utilizada. El uso de xml sobre una gran variedad de tipos de documentos digitales produce una colección de objetos muy homogénea de cara a la conservación a largo plazo, facilitando procesos de migración o emulación.

gestionados conjuntamente con el resto de documentos digitales.

—**Modelo base de datos** (u orientado a procesos): toda la página web es un documento de archivo. En este caso es preciso almacenar periódicamente una especie de “fotografía” digital de la sede web, fijando su contenido en un momento dado y almacenándolo en un depósito digital, o en un sistema que permita un control de versiones. Aun así, la funcionalidad de muchos elementos de la página web —así como la vigencia de los enlaces— es muy difícil de capturar, por lo que en la fase de evaluación hay que establecer:

a. **Los límites del documento**, es decir hasta dónde hay que seguir los enlaces, qué elementos externos forman parte del documento y qué elementos multimedia deben archivarse conjuntamente.

b. **El contexto del documento**. Se debe crear un registro (*log*) del contexto de creación y utilización de la página web. Estos registros también forman parte, conceptualmente, del mismo documento.

c. **La estructura y la presentación**. En las bases de datos dinámicas es necesario conservar también los datos de origen, o las sucesivas presentaciones de estos datos.

Si el mantenimiento de una página web está centralizado y depende del filtro de un *webmaster* es más fácil preparar los contenidos para su conservación a largo plazo que si todos los usuarios tienen la posibilidad de publicar directamente en internet o en la intranet. La actuación sobre las fuentes de información es siempre más efectiva que la captura de los contenidos una vez publicados. Así, el *webmaster* puede extraer periódicamente una copia de archivo de cada página publicada, y mantener un registro histórico del mapa de la sede web y de los enlaces³³. Existen en el mercado programas que facilitan la automatización de esta tarea, incluso en entornos descentralizados donde todos los usuarios pueden publicar directamente los documentos.

Entre las dos estrategias reseñadas, actualmente se impone la tendencia **orientada a documentos** (véase xml), cada vez más autosuficientes, independientes de un software concreto, que podrán ser modificados y visualizados por distintos programas a lo largo de su ciclo de vida. La *Victorian electronic records strategy (Vers)*³⁴, Australia, se decanta por esta opción. Según ella, todos los documentos digitales deben estar un formato que permita:

—Soportar una conservación a largo plazo, es decir que sean accesibles, que se puedan visualizar en su forma original y que incluyan su contexto de creación.

—Servir de evidencia para saber quién creó el documento, cuándo lo hizo, y que el documento no ha sido modificado desde ese momento, es decir que está íntegro.

—Si el documento está evaluado, saber cuándo se debe transferir o destruir y ejecutar dicho proceso.

—Debe ser posible aumentar o cambiar la información asociada a un documento sin alterar de ninguna manera su valor evidencial ni su integridad. Para cumplir con este requerimiento, en el aspecto estructural los documentos digitales deben ser **autodocumentados** (que sea posible reconstruir el contexto de creación a partir del mismo documento, sin necesidad de disponer del sistema que lo creó), **autocontenidos** (de manera que toda la información sobre el documento esté contenida en el mismo documento) y **extensibles** (de forma que se pueda ampliar la estructura del documento para añadir nuevos metadatos o enlaces sin afectar su estructura básica).

«La herramienta conceptual que se convierte en la clave para la preservación de los documentos electrónicos son las reglas de evaluación documental»

Es interesante remarcar la **extensibilidad** como característica de la conservación orientada a documentos. Cuando un documento digital es firmado digitalmente o enviado por correo electrónico se convierte en un original, en un documento definitivo y auténtico con valores primarios. Pero aquí no finaliza su utilización. El documento puede pasar por trámites o validaciones que le añadan información, marcas de autenticación, control de versiones, o referencias a otros documentos posteriores (por ejemplo durante su tramitación dentro de un sistema *workflow* o con herramientas *groupware*). Asimismo, el entorno tecnológico del documento cambiará, y el documento tiene que ser compatible con los nuevos sistemas sin perder sus características primigenias. Por esta razón la *Vers* define dos partes en un documento digital:

—Una parte invariable o **contenido**, que es aquella que debe ser conservada en su codificación y estructura originales (pdf, xml, etc.).

—Una parte variable o **contexto**, que acumula todos los metadatos que describen la estructura, el contenido, el contexto documental o las relaciones con otros documentos, y las características técnicas y de reproducción o visualización.

3. La importancia de la evaluación

Todas estas propuestas eminentemente tecnológicas exigen una actuación previa en la fase de creación de los documentos. Con el fin de preservar sus valores primarios el documento debe ser **creado de forma controlada**, de manera que el sistema garantice su fiabilidad y su autenticidad (recordemos que el sistema informático es parte del contexto que da sentido al documento). Pero también debe ser creado **en el formato que mejor responda a su futura disposición**, sea esta la conservación o la destrucción. De este modo, la herramienta conceptual que se convierte en la clave para la preservación de los documentos electrónicos son las **reglas de evaluación documental** (*appraisal schedules*)³⁷.

Estas reglas no corresponden al modelo tradicional de expurgo (la selección realizada sobre conjuntos de documentos preexistentes), sino a la **evaluación por funciones**, que parte del análisis del contexto general de la organización (macroevaluación) para llegar al detalle en cada función o proceso (microevaluación). Estas reglas no se limitan a un calendario de conservación, sino que especifican el formato de creación y utilización del documento acorde con su período de conservación. Asimismo establecen su contenido informativo en función de su participación en un determinado proceso, su régimen de acceso y sus sistemas de autenticación, permitiendo así calcular los costes de la conservación de los documentos digitales. En definitiva, las reglas de evaluación proporcionan toda la información que se debe introducir en el sistema informático de una empresa o institución para que su gestión documental sea verdaderamente eficaz.

La localización física de los documentos informáticos es aleatoria, de modo que la clasificación es la única forma de relacionar archivísticamente dichos documentos. Los documentos virtuales (compuestos, distribuidos y dinámicos) no pueden ser conservados como tales más que a partir de la clasificación, que permite relacionar intelectualmente las distintas partes del documento. Las reglas de evaluación parten pues de una herramienta previa, el **cuadro de clasificación**, que permite identificar y referenciar todos los procesos de una empresa o institución susceptibles de ser documentados³⁸. El diseño de esta herramienta va íntima-

El modelo finlandés aplicado a la evaluación de los documentos electrónicos

El modelo finlandés de gestión documental se ha caracterizado siempre por su pragmatismo en la búsqueda de soluciones. Según este modelo⁴², las reglas de evaluación (*records management schedules* en este caso) deben incluir cierta información básica sobre los sistemas informáticos que generan los documentos, como por ejemplo:

- Funciones básicas del sistema.
- Si la información en el sistema está clasificada o no.
- Información técnica (software, hardware).
- Personas responsables de la información y del mantenimiento técnico del sistema.

Desde 1994 en los *Archivos Nacionales de Finlandia* un grupo de trabajo estudia los materiales digitales y propone métodos para su tratamiento. Una aportación interesante es la distinción entre conservación permanente y conservación continua. Los documentos de **conservación permanente** son aquellos que se conservan por su valor histórico, y deben permanecer inalterables, mientras que los documentos de **conservación continua** son aquellos que se conservan por sus valores primarios y deben permanecer siempre accesibles. Esta distinción debe ponerse en relación con las dos necesidades especificadas por el proyecto de la *Universidad de British Columbia*, cuando distingue entre unas necesidades primarias de integridad y unas necesidades secundarias de conservación.

mente ligado a proyectos de reingeniería de procesos (*bpr*), e incluso el proyecto de la *Universidad de British Columbia* establece que para garantizar la correcta gestión y preservación de los documentos electrónicos es necesario integrar los procedimientos documentales con el resto de procedimientos de la empresa. Así se ha puesto en marcha en el estado de Kentucky (EUA), donde el programa para impulsar las transacciones electrónicas entre las empresas y unificar los servicios de las oficinas públicas (*Enterprise information technology strategic plan*) incorpora el requerimiento de que una vez rediseñados los procesos se elaboren las herramientas conceptuales básicas del sistema de gestión documental antes de proceder a la informatización.

David Bearman³⁹ expone cuatro principios para la evaluación de los documentos digitales:

—Ya no es posible evaluar únicamente documentos. Se deben evaluar las funciones y los procesos que producen documentos.

—Son los archiveros⁴⁰ quienes deben definir los requerimientos funcionales para la gestión del ciclo de vida de los documentos digitales de una organización.

—La evaluación debe realizarse en la fase de diseño de los sistemas informáticos.

—Los conjuntos de metadatos que se asocian a los documentos deben incluir elementos que sirvan para describir las diferentes etapas de su ciclo de vida.

Quiero insistir en la afirmación de que las reglas de evaluación deben introducirse en el sistema informático de la organización antes de que se empiecen a generar documentos. El objetivo es asegurar que los documentos con valores secundarios serán conservados en una forma auténtica, y esto no es posible si el formato original no facilita la migración, si no se han recogido los metadatos suficientes o si no se han conservado también los documentos y conjuntos de datos que configuran el contexto documental. La elección de la estrategia de conservación a largo plazo (tipo de migración, estándares, metadatos, etc.) es una de las decisiones que se deben tomar durante la evaluación, e implica participar, junto con los analistas informáticos, en el diseño de sus sistemas, pasando de la tradicional evaluación enfocada a los documentos a evaluar la estructura del sistema de información de la organización⁴¹.

Para gestionar con eficacia la documentación administrativa tradicionalmente se ha utilizado una periodización por etapas orientada a optimizar los recursos invertidos. La más habitual consiste en distinguir una fase de utilización administrativa o activa, una fase semiactiva y una fase inactiva. Con los documentos

digitales es difícil seguir manteniendo la misma división, puesto que las decisiones que antes se debían tomar en cada una de las fases las tomará automáticamente el sistema a partir de la información que se le proporcione durante el análisis y diseño. Parece necesario sustituir esta periodización por modelos basados en considerar el ciclo de vida de los documentos administrativos como un *continuum*, un destino que tienen escrito los documentos evaluados, y que puede ser reescrito si cambian las condiciones del entorno de utilización⁴³. En este caso se debe admitir la posibilidad de reevaluar los documentos periódicamente o con motivo de la migración a un nuevo entorno tecnológico⁴⁴ (tabla 2)⁴⁵.

4. Conclusión

El problema de la conservación de los documentos electrónicos reside tanto en la diversidad de soluciones tecnológicas como en la falta de cohesión de las soluciones metodológicas que se aplican. A modo de conclusión quiero citar el acertado colofón conceptual que pone **Margaret Hedstrom** a esta problemática, cuando detalla las condiciones básicas para una estrategia de preservación digital⁴⁶. Estas condiciones se resumen en la necesidad de disponer de un sistema de gestión documental que garantice la conservación de los documentos digitales, y que sea:

—Implantado: Las posibilidades de conservar los documentos electrónicos van asociadas al nivel de implantación real del sistema de gestión documental. El conocimiento detallado de los documentos que se deben conservar es la premisa básica para articular una estrategia de preservación digital.

—Integrado: Es necesario que la gestión de los documentos que pertenecen a un mismo fondo sea integrada, con independencia de su soporte y ubicación.

	Dinamarca	Finlandia	Francia	Alemania	Suecia	Reino Unido
Fecha de creación	1973	1995	1986	1991	1983	1998
Volumen de datos	6500 docum. 4500 cintas 200 CD-ROM		6000 docum.	23000 docum.	2 Terabytes de información 11500 cintas	25 bases de datos
Tipo de soporte	CD-R	DAT	Cinta DLT4000 DLT7000 DAT Óptico	Cinta DAT CR-ROM	Cinta DAT CR-ROM CD-R	Cinta DAT CR-ROM
Formatos de texto	ASCII	ASCII	ASCII	ASCII EBCDIC	ASCII EBCDIC SGML	PostScript
Formatos de bases de datos	ASCII	ASCII	ASCII	ASCII EBCDIC	ASCII EBCDIC	ASCII EBCDIC
Formatos de imagen	TIFF 6.0		TIFF	TIFF 6.0 CCITT/TSS		TIFF 6.0

Tabla2. Archivos digitales nacionales en Europa

Los documentos digitales deben gestionarse conjuntamente con los documentos en soportes tradicionales.

—En la fase de diseño: Las reglas de evaluación deben incorporarse a los sistemas informáticos en la fase de diseño.

—Controlado: Se deben establecer políticas que eviten que la decisión sobre la conservación o eliminación de documentos, o sobre la aplicación de las reglas de evaluación, recaiga sobre el usuario del sistema.

—Normalizado: La tan demandada normalización no se debe restringir a los formatos. Los métodos de trabajo y las herramientas conceptuales deben seguir también un proceso de normalización que permita su integración con los sistemas informáticos⁴⁷.

Bibliografía y notas

1. En los documentos firmados digitalmente es necesario adoptar una política para la conservación de las claves de encriptación.

2. **Kampffmeyer, Ulrich**. “E-documents: It’s all legal, or is it?”. *E-doc*, 2000, septiembre-octubre, p. 31-33.

3. El *Real decreto ley 14/1999*, de 17 de septiembre, sobre la firma electrónica regula las funciones de los prestadores de servicios de certificación (las “notarías digitales”), los cuales autentican las transacciones telemáticas privadas o públicas mediante un sistema de claves asimétricas. Cabe destacar el artículo 3.1. “La firma electrónica avanzada, siempre que esté basada en un certificado reconocido y que haya sido producida por un dispositivo seguro de creación de firma, tendrá, respecto de los datos consignados en forma electrónica, el mismo valor jurídico que la firma manuscrita en relación con los consignados en papel y será admisible como prueba en juicio, valorándose ésta según los criterios de apreciación establecidos en las normas procesales”.

4. *Ley 30/1992 de Régimen jurídico de las administraciones públicas y del procedimiento administrativo común*, de 26 de noviembre de 1992. En el artículo 45.5 dice: “Los documentos emitidos, cualquiera que sea su soporte, por medios electrónicos, informáticos o telemáticos por las Administraciones Públicas, o los que éstas emitan como copias de originales almacenados por estos mismos medios, gozarán de la validez y eficacia de documento original siempre que quede garantizada su autenticidad, integridad y conservación y, en su caso, la recepción por el interesado, así como el cumplimiento de las garantías y requisitos exigidos por esta u otras Leyes”. Este artículo fue desarrollado por el *Real decreto 263/1996*, de 16 de febrero, por el que se reguló la utilización de técnicas electrónicas, informáticas y telemáticas por la Administración General del Estado.

5. *The preservation of the integrity of electronic records* [en línea]. **Luciana Duranti, Terry Eastwood, Heather MacNeil**. Vancouver: University of British Columbia; School of Library, Archival & Information Studies, 1997. Consultado: 14-07-2001. Disponible en: <http://www.interpares.org/UBCProject/>

6. El valor de un documento como prueba de un hecho y su capacidad para, a raíz de este valor, dar origen a nuevas acciones están sujetos a que el documento cumpla dos condiciones:

—Que sea fiable, es decir, que su contenido sea cierto. La fiabilidad es, por lo tanto, una noción relativa, y por esto se habla del grado de fiabilidad de un documento.

—Que sea auténtico, es decir, que su grado de fiabilidad inicial no se haya alterado desde su creación y durante su utilización y conservación.

Gestión de los documentos digitales: estrategias para su conservación

La fiabilidad depende de las condiciones de creación del documento (la fiabilidad del autor, el grado de control ejercido durante su creación y la integridad o nivel de completitud del documento), mientras que la autenticidad depende de las condiciones de utilización y, posteriormente, de conservación (los mecanismos de transmisión y de custodia).

7. Sobre la formulación de la teoría de los valores véase: **Schellenberg, T. R.** *Archivos modernos: principios y técnicas*. La Habana: Instituto Panamericano de Geografía e Historia, 1958, p. 163-210.

8. **Hofman, Johannes**. “Shooting at a moving target: The development of a repository for the preservation of digital information”. En: *European Commission. Proceedings of the DLM-Forum: European citizens and electronic information: the memory of the information society*. Luxembourg: Office for Official Publications of the European Communities, 2000, p. 81.

9. Comité sur les documents électroniques. *Guide pour la gestion archivistique des documents électroniques*. Paris: Conseil International des Archives, 1997. (Études, 8), p. 53-56.

10. **Conchon, Michèle**. “Traitement et conservation des archives informatiques en France: la méthode Constance”. En: *Commission Européenne. Actes du DLM-Forum sur les données lisibles par machine*. Luxembourg: Office des publications officielles des Communautés européennes, 1997, p. 146-151. También en: **Dhérent, Catherine**. “The Archives of France: policy on storing electronic documents”. En: *European Commission. Proceedings of the DLM-Forum: European citizens and electronic information: the memory of the information society*. Luxembourg: Office for Official Publications of the European Communities, 2000, p. 172-179.

11. Las 4.500 cintas existentes fueron reemplazadas por 40 discos ópticos, con un coste de tres millones de francos. Desde mayo de 1999, 20 cintas han reemplazado a los 40 discos ópticos.

12. Hasta el momento todos los metadatos recopilados sobre cada documento digital, y las fichas descriptivas asociadas, se conservan preferentemente en papel. Se está experimentando en la digitalización de estas descripciones.

13. *Transfer of electronic records (from 36 CFR 1228.270 - formerly 36 CFR 1228.188)* [en línea]. Washington, DC: National Archives and Records Administration, last updated April 5, 2001. Consultado: 11-07-2001. Disponible en: <http://www.nara.gov/nara/electronic/transfer.html>

14. La aceptación de sql e irids como formatos de base de datos es una propuesta del informe presentado en mayo de 1989 por el *National Computer Systems Laboratory del National Institute of Standards and Technology (Nist)*, a petición de la *Nara*.

15. Básicamente el diccionario de datos y el conjunto de relaciones.

16. **Rothenberg, Jeff**. *Avoiding technological quicksand: finding a viable technical foundation for digital preservation*. Amsterdam: European Commission on Preservation and Access; Washington DC: Council on Library and Information Resources, 1999. Véase también **Rothenberg, Jeff**. *An experiment in using emulation to preserve digital publications*. Den Haag: The Koninklijke Bibliotheek; Rand-Europe, 2000.

Otro estudio interesante es: **Rothenberg, Jeff; Bikson, Tora K.** *Carrying authentic records through time in a usable and understandable way*. Disponible en: <http://www.archief.nl/digiduur>

17. **Rothenberg, Jeff**, op. cit., 1999.

18. Véase el comentario de la metodología orientada a objetos en este mismo artículo.

19. La *Universidad de Michigan*, junto con el proyecto *Cedars (Curl exemplars in digital archives)*, desarrolla una iniciativa centrada en la aplicabilidad de la emulación para ciertos tipos de objetos digitales en los

que la presentación es esencial. Otra experiencia importante en emulación es la que realiza el proyecto *Nedlib* (*Networked European Deposit Library*), bajo la coordinación de la Koninklijke Bibliotheek. <http://www.leeds.ac.uk/cedars/> <http://www.konbib.nl/nedlib/>

20. Se puede obtener más información del *Formato Universal de Preservación* en: <http://info.wgbh.org/upfl>

21. *Bento specification*. Apple Computer Inc., c1993. <ftp://cil.org/pub/cilabs/tech/bento/>

22. *Open media framework interchange specification*. Avid Technology Inc., c1995. <http://www.avid.com/omfi>

23. Véase el apartado correspondiente en este mismo artículo.

24. National Preservation Office. *Digital culture: maximising the nation's investment*. **Mary Feeney** (ed.). London: British Library Board, 1999.

25. Comité sur les documents électroniques, op. cit., p. 48-52.

26. Australian Archives. *Managing electronic messages as records*. También: **Peterson, Trudy H.** *Nara electronic mail regulations*, donde se proponen los criterios para discriminar los mensajes con valor evidencial del resto de comunicaciones. Se incluyen en esta consideración otros objetos digitales típicamente asociados a los programas de correo electrónico, como pueden ser los calendarios y agendas.

27. *Electronic mail policy*. University of California; Office of the President, 1998.

28. **Harries, Stephen**. "Capturing and managing electronic records from websites and intranets in the government environment". En: *European Commission. Proceedings of the DLM-Forum: European citizens and electronic information: the memory of the information society*. Luxembourg: Office for Official Publications of the European Communities, 2000, p. 72-79.

29. **Serra, Jordi**. "Archivar internet". *El profesional de la información*, 1998, núm. 7-8, julio-agosto.

30. **Kahle, Brewster**. "Archiving the internet". *Scientific American*, March, 1997. Es interesante visitar la página de *The internet archive: building an internet library*, que actualmente cuenta ya con 4 billones de páginas web almacenadas, 16 millones de mensajes de Usenet y 360 películas que en total ocupan más de 40 terabytes de memoria. Consultado: 14-07-2001. <http://www.archive.org/index.html>

31. <http://pandora.nla.gov.au/pandora/>

32. **Berninger, Peter**. *Website electronic records management (Werm) in the European Commission*. January 2000.

33. Como ejemplo, la *Nara* ha evaluado todas las páginas web estáticas del *Gobierno Federal* disponibles en internet a partir del 20 de enero de 2001 como documentos de conservación permanente. Para su preservación captura periódicamente una "instantánea" de las páginas web gubernamentales.

34. <http://www.prov.vic.gov.au/vers/welcome.htm>

35. **Thibodeau, Kenneth**. "Limites et transformations: une stratégie orientée objet pour la conservation des documents électroniques". En: Commission Européenne. *Actes du DLM-Forum sur les données lisibles par machine*. Luxembourg: Office des publications officielles des Communautés européennes, 1997, p. 161-167. También en: **Thibodeau, Kenneth; Moore, Reagan; Baru, Chaitanya**. "Persistent object preservation: advanced computing infrastructure for digital preservation". En: European Commission. *Proceedings of the DLM-Forum: European citizens and electronic information: the memory of the information society*. Luxem-

bourg: Office for Official Publications of the European Communities, 2000, p. 113-118.

36. Los atributos de un objeto constituyen el estado en que se encuentra. Un comportamiento es una acción o un proceso que opera sobre un estado del objeto.

37. **Eastwood, Terry**. *Appraisal of electronic records: a review of the literature in English*. Appraisal Task Force. Interpairs project, 2000.

38. La identificación de un documento ya forma parte del proceso de evaluación, puesto que permite el estudio del ciclo de vida de dicho documento dentro del sistema informático que lo genera.

39. **Bearman, David**. "Archival strategies". *The American archivist*, 1995, 58:4.

40. Es este caso la traducción corresponde al *records manager*, el responsable de gestionar la documentación administrativa de una organización. En nuestro contexto esta figura corresponde al archivero de un archivo administrativo que actúa sobre la fase activa de los documentos.

41. **Canela, Montserrat; Campos, Isabel; Domingo, Joan; Serra, Jordi**. "Le processus d'évaluation comme moyen d'intégrer le point de vue archivistique dans la planification, la création et l'utilisation de documents électroniques et de systèmes automatisés. Étude de cas". En: Commission Européenne. *Actes du DLM-Forum sur les données lisibles par machine*. Luxembourg: Office des publications officielles des Communautés Européennes, 1997, p. 113-117.

42. **Pohjola, Raimo**. "Appraisal and disposal of electronic records". En: *Principles of appraisal and their application in electronic environment: European models and concepts*. DLM-Experts meeting in Tampere, Finland, 1999.

43. **Hofman, Johannes**, op. cit., p. 81.

44. **Kowlowitz, Alan**. "Appraising in a vacuum: electronic records appraisal issues: a view from the trenches". En: **David Bearman** (ed.). *Archival management of electronic records*. Archives and museum informatics technical report, 1991, n. 13.

45. **Schürer, Kevin**. "Survey on the relationship between public administration and archives services concerning electronic records management in the EU Member States." En: *European Commission. Proceedings of the DLM-Forum: European citizens and electronic information: the memory of the information society*. Luxembourg: Office for Official Publications of the European Communities, 2000, p. 219-235.

46. **Hedstrom, Margaret; Kowlowitz, Alan**. "Meeting the challenge of machine readable records: a state archives perspective". *Reference services review*, 16(1/2), p. 31-40.

47. Véase:
US Department of Defense 5015.2 Standard:
<http://jtitc.fhu.dis.mil/reemgt/>

Standard for the management of electronic records, del Vers:
<http://www.prov.vic.gov.au/vers/>

General records schedule 20 de los archivos de los Estados Unidos.
<http://www.nara.gov/records/grs20/>

En Europa se está trabajando en un estándar de gestión documental (*ISO/CD 15489*).
Oais (*Open archive information system*)

Jordi Serra Serra. *Facultat de Biblioteconomia i Documentació*. *Universitat de Barcelona*.
DLM-Forum on electronic records, Monitoring committee.
serra@fbd.ub.es
jordiserra@correu.gencat.es



En primera línea en Sistemas de Información y Gestión del Conocimiento.

- Informática Documental
- Internet, Intranet
- Edición de Bases de Datos en CD-Rom
- Base de Datos PRENSA BARATZ (Internet y CD-Rom)
- Sistema Integrado de Gestión Bibliotecaria Absys
- Catalogación Retrospectiva
- Sistema Integrado de Gestión de Centros Archivísticos Albalá



BARATZ
SERVICIOS DE
TELEDOCUMENTACION S.A.

Raimundo Fernández Villaverde, 28
28003 Madrid (España)
Teléfono +34 91 456 03 60 - Fax +34 91 533 09 58
www.baratz.es - E-mail: informa@baratz.es

