GUIA PARA DIGITALIZAR DOCUMENTOS

María Teresa Bermúdez Muñoz[1]

INTRODUCCION

Se dice que estamos en el siglo de la información y del conocimiento y que, tanto en las empresas como en las instituciones del estado, las tecnologías de la Información y de las Telecomunicaciones deben ocupar un lugar destacado con el fin de brindar servicios eficientes y eficaces.

Y dónde se encuentra la información en las instituciones? Pues bien, la información está en los documentos que producimos o recibimos todos los días en el ejercicio de nuestras funciones.

Los documentos nos permiten, entre otras cosas, las siguientes:
- Conducir los asuntos de una manera ordenada, eficaz y responsable.
- Dar coherencia, continuidad y productividad a las labores de la entidad.
- Mejorar los resultados de las actividades que se realizan en la entidad.
- Suministrar pruebas en caso de litigio.
- Proteger los interese y derechos de los colaboradores y de los usuarios.
- Proporcionar evidencia de las actividades que se realizan.

Por otra parte, los directivos de las instituciones públicas dicen estar de acuerdo con el discurso teórico que concede una gran importancia a la gestión de la información pero en muchos casos no saben cómo abordar los problemas de la información que se dan con soluciones prácticas.

En una gran mayoría de casos, se quiere solucionar los problemas de información desde una perspectiva informática, esto debido a la publicidad que constantemente nos bombardea con frases como las siguientes: "cualquier día de estos su empresa podría hundirse en un mar de papeles", "alivie los requerimientos de bodega, digitalice sus documentos", "prohibido el papel", etc.

Sin embargo la experiencia ha probado que confundir la herramienta con la solución del problema puede generar frustraciones como hacer grandes inversiones económicas con resultados poco efectivos, hundirnos en un mar de registros, cambiar bodegas reales por bodegas virtuales, etc. Como dice una consultora española "Todos sabemos que comprarnos el mejor martillo del mundo no equivale a poder clavar doscientos clavos en una hora y todos derechos..."

---

1. Jefe del Departamento Servicios Archivísticos Externos del Archivo Nacional. Profesora de la Sección de Archivística de la Universidad de Costa Rica

Para conseguir una buena gestión de la información, ya sea manual o informatizada, se debe contar primero con una buena gestión documental.

La gestión documental persigue que el trabajo con los documentos sea más fácil, de manera que una persona sabe qué documentos guardar, cuándo, cómo, y dónde, con el fin de localizarlos cuando los necesita.

Concretamente una buena gestión documental permita:

- Definir cuáles documentos deben ser creados según cada actividad y cuál información deben contener.
- Decidir bajo cual forma y estructura deben ser producidos los documentos.
- Seleccionar cual tecnología debe utilizarse para la conservación documental.
- Definir cuales metadatos conviene establecer para cada documento y cómo se van a conservar y administrar durante la vigencia administrativa, legal e histórica.
- Determinar las necesidades de recuperación, utilización y de transmisión de los documentos entre los usuarios inmediatos y los otros usuarios.
- Establecer los plazos de conservación de los documentos de acuerdo con su vigencia administrativa-legal e histórica con el fin de conservarlos el tiempo necesario y no saturar los locales de archivo.
- Determinar el modo de organización de los documentos con el fin de recuperarlos eficaz y eficientemente.
- Evaluar los riesgos que se pueden presentar en caso de que los documentos no estén disponibles cuando se requieran por parte de la administración y de los usuarios en general.
- Asegurarla adecuada conservación de los documentos en locales que cumplan las condiciones legales de sanidad y seguridad.

ASPECTOS FUNDAMENTALES DE UN PROGRAMA DE DIGITALIZACION [2]

Todo proyecto de digitalización debe contemplar tres aspectos fundamentales.

- Los objetivos que se pretenden conseguir: La definición de estos objetivos dependerá de lo que cada institución pretenda realizar. Se debe establecer la relación costo-beneficio del proyecto y comparar la situación actual de la gestión documental en soporte papel con la situación futura que se presentará al digitalizar los documentos, por ejemplo: recurso humano, material y financiero requeridos en uno y otro caso, confrontación de los niveles resultado real y esperado, pertinencia del cambio, etc. También se debe tener claro si se quiere digitalizar para sustituir los documentos en soporte papel, para agilizar los trámites o para difusión.

---

2. Muchos de los conceptos aquí enunciados fueron tomados de la "*Guide d'imagerie numérique. Numerisation de l'information consignée sur des supports traditionnels*", elaborada por Micheline Bélanger y otros, del Archivo Nacional de Québec y adaptados a la realidad archivística de Costa Rica.

- El ámbito de aplicación del proyecto: Implica integrar los archivos en papel, los documentos electrónicos y las bases de datos, en el caso de que esté trabajando en esas tres áreas por separado.

- La responsabilidad de las partes involucradas en el proyecto: Un proyecto de digitalización requiere la participación de administradores, archivistas e informáticos, de manera que se debe establecer claramente, desde el principio del proyecto, la responsabilidad de cada uno de los participantes.

Asimismo, en un programa de digitalización se deben cubrir tres fases diferentes pero estrechamente ligadas entre sí:

1. FASE ARCHIVÍSTICA: Preparación y conducción de la digitalización de documentos según las reglas archivísticas.

2. FASE TECNOLÓGICA: Identificación y aplicación de un cierto número de reglas directamente ligadas a la tecnología de la digitalización.

3. FASE LEGAL: Cumplimiento de las prescripciones legales para garantizar el valor de prueba de los documentos digitalizados.

1. FASE ARCHIVÍSTICA

Si se digitalizan los documentos sin una preparación adecuada, es decir, sin orden, método y sin una descripción precisa para recuperar la información, lo que se obtendrá como resultado es un caos electrónico costoso y peligroso, siendo más difícil acceder a una información almacenada en un disco óptico que abarca varias decenas de millares de imágenes que a documentos conservados en cajas debidamente rotuladas.

La digitalización propiamente dicha es un asunto de segundos, pero la preparación de los documentos seleccionados, la puesta en marcha del proceso y el control de calidad de los resultados requiere de mucho tiempo.

1.1 La preparación material de los documentos

Algunos documentos pueden ser digitalizados de forma más rápida, fácil y provechosa que otros. Partiendo de los altos costos que implica digitalizar, se debe tener presente algunos parámetros para seleccionar los documentos que se van a someter a ese proceso, estos son:

- El soporte de transacciones costosas.
- Frecuentemente son consultados, manipulados y reproducidos.
- Deben ser conservados a largo plazo.
- Deben ser consultados por múltiples usuarios.
- Son homogéneos en su forma, tamaño y contenido.
- Representan una gran masa documental.
- Han sido clasificados, expurgados y descritos con precisión para facilitar su acceso.

Según la naturaleza y el estado de los documentos así como su contexto organizacional, la preparación material de los documentos se puede hacer antes, durante o después de su preparación intelectual. Esta preparación incluye los pasos siguientes:

- Retirar y eliminar los duplicados así como documentos rutinarios y formularios no útiles. Esto se debe hacer de acuerdo con las Tablas de Plazos de Conservación Documental o de algún informe de valoración parcial. Se debe tener presente que ningún documento producido o recibido por una entidad pública, se puede eliminar sin el aval del Comité Interinstitucional de Selección respectivo y sin la aprobación de la Comisión Nacional de Selección y Eliminación de Documentos. En el caso del expediente clínico, existe una Comisión de Expediente Clínico que está trabajando en una Tabla Modelo de Plazos de Conservación del Expediente Clínico.

- Separar cada página, distinguiendo cada documento en relación con el precedente y el siguiente, es decir, que debe tener una continuidad cronológica entre un documento y otro.

- Foliar los documentos de un mismo expediente.

- Llevar un control estricto de los documentos que se van digitalizando.

- Guardar los originales, cuando se van a conservar permanentemente, en cajas adecuadas para su conservación.

1.2 La descripción y los medios para recuperar los documentos

Antes de proceder a la digitalización, los documentos se deben inventariar ya sea utilizando formularios manuales o informatizados. Los inventarios deben contener los siguientes datos:

- Unidad administrativa que produjo el documento.
- Nombre completo del paciente
- Número de expediente o de cédula, en caso de que existan dos números, se deben incluir los dos
- Fecha y lugar de nacimiento
- Fechas extremas (fecha del inicio y del cierre del expediente)
- Vigencia administrativa y legal
- Cualquier otro dato que se considere pertinente

Si la entidad posee un inventario de los documentos que se quieren digitalizar, debe verificar antes de iniciar la digitalización que la descripción hecha corresponde a cada documento, en este caso, a cada expediente clínico.

Si el método de trabajo empleado y los programas informáticos utilizados lo permiten, se puede insertar al principio y al final de cada expediente una hoja con un código de barras, lo que permitirá recuperar más fácilmente el expediente. También, si la naturaleza y la forma de los documentos lo permiten, se puede utilizar

un programa de reconocimiento óptico de caracteres con el fin de identificar y recuperar más rápidamente los documentos digitalizados.

1.3 Conservación

Se deben tener en cuenta los siguientes lineamientos en aspectos ambientales para la conservación y la manipulación de soportes electrónicos:

- Temperatura - 16° C a 20° C

- Humedad – 30% a 40%

- Evitar los cambios bruscos de temperatura y humedad. Lo fundamental es no llegar a la condensación de la humedad. De ahí la importancia de conservar estos soportes en lugares en donde la temperatura y la humedad relativa sean constantes todo el tiempo

- La colocación de los soportes en la estantería debe ser en forma vertical

- Se deben conservar estos soportes en lugares libres de polvo y suciedad

- La manipulación debe ser muy cuidadosa porque estos soportes se deforman con facilidad y esto ocasiona deficiencias en la recuperación de los datos.

- No se deben colocar estos soportes cerca de campos magnéticos porque producen deterioro en los mismos

- Se debe mantener al menos una copia de seguridad de los documentos electrónicos en depósitos ubicados en localizaciones diferentes

- Se debe tener cuidado especial en las labores de consulta y recuperación de documentos electrónicos y en la manipulación de sus soportes para garantizar la integridad y conservación de los mismos

-Se debe garantizar el buen funcionamiento de los dispositivos con los que se gestionan los documentos electrónicos, que tengan un mantenimiento adecuado y estén debidamente protegidos contra fluctuaciones de voltaje

- Se debe dedicar tiempo y recursos para efectuar investigaciones periódicas de mercado, en cuanto a los dispositivos de almacenamiento electrónico de información y sus soportes, para hacer las adquisiciones y ajustes de equipo necesarias

- Se debe tener el personal debidamente capacitado para efectuar la migración de los documentos electrónicos a nuevos soportes cuando el acceso a los mismos se vea comprometido por la obsolescencia de los soportes en que se encuentran.

2. FASE TECNOLOGICA

En materia informática se deben tener en cuenta los siguientes aspectos:

- La selección del sistema adecuado
- Los objetivos del sistema
- Los estándares tecnológicos
- El acceso a las imágenes y a los índices de datos
- La calidad de las imágenes
- La funcionalidad del sistema
- La protección de la información digitalizada
- El control del acceso
- La evolución de las tecnologías de la información
- Las auditorías

2.1 La selección del sistema adecuado

Debe efectuarse después de un análisis costo-beneficio. Se debe tomar en consideración la infraestructura tecnológica de la entidad para que el sistema que se va a adquirir sea compatible con el ya existente, así como revisar el conjunto de las necesidades de información de la entidad y el proceso de gestión de la información electrónica. De otra forma, el sistema podría multiplicar las lagunas del actual.

2.2 Los objetivos de un sistema

Los responsables de elegir un sistema para digitalizar deben visualizar los objetivos siguientes:

- El acceso a largo plazo de la información digitalizada
- La calidad de las imágenes digitalizadas
- La seguridad y la confidencialidad de la información
- La funcionalidad continua del equipo y de los programas
- El control del deterioro de los soportes de almacenamiento de la información
- La anticipación y la planificación del desarrollo tecnológico

2.3 Los estándares tecnológicos

El respeto de un conjunto de estándares tecnológicos contribuye mucho a lograr los objetivos antes enunciados. Por ejemplo: el Archivo Nacional de Québec recomienda tener en cuenta los siguientes estándares:

- La Norma ISO 9660 para el registro y lectura de los datos en discos ópticos con el fin de asegurar su migración a diferentes ambientes tecnológicos.
- El formato TIFF para la toma de imágenes y los estándares CCITT grupo 3 y grupo 4 para comprimir datos textuales.
- El formato ASCII para la recogida de índices.
- La interfaz SCSI-2 entre el computador, el digitalizador y el grabador o quemador.
- Un monitor de visualización de alta resolución.

Por su parte la Dirección de los Archivos de Francia recomienda los siguientes formatos:

- TXT para textos
- XML para textos estructurados
- UIP recomendación T4 para imagen fija en blanco y negro
- PNG para imagen fija en color
- Texto plano para base de datos

2.4 El acceso a las imágenes y a los índices de datos

El acceso a la información conservada en discos ópticos supone que esa información permanecerá legible, inteligible y recuperable.

La legibilidad significa la posibilidad de acceder en todo momento, presente o futuro a la información digitalizada sin ninguna pérdida de datos. Para garantizar esa legibilidad, es necesario conservar los discos ópticos en las condiciones apropiadas y verificarlos regularmente para detectar errores de lectura, con el fin de recopiarlos inmediatamente.

Generalmente, la imposibilidad de acceso a la información proviene de la obsolescencia del material.

La inteligibilidad por su parte significa que la información producida y recuperada por un sistema de digitalización permanece comprensible para otro sistema. Además, implica que la información leída con la ayuda de un computador debe ser comprendida por el ser humano.

Las barreras de la inteligibilidad de la información son los campos preestablecidos para la recuperación de los datos de los sistemas comerciales, las técnicas de compresión y la obsolescencia de los programas informáticos.

La recuperación de la información implica la posibilidad de encontrar la información almacenada en los discos ópticos y accederla fácilmente. Esa posibilidad está basada en los elementos que establecen las ligas entre la estructura lógica de los registros y su ubicación en el disco óptico Las ligas se encuentran habitualmente en los encabezados de los ficheros y son interpretadas por el sistema del computador. Para mantener a largo plazo la capacidad de recuperación de la información, se debe asegurar el funcionamiento del sistema y de los dispositivos periféricos (device, drivers).

Para garantizar el acceso a la información y el índice de los datos, se deben respetar las siguientes reglas:

-Establecer una estrategia de migración que permita la evolución del equipo, los programas informáticos y los soportes de almacenamiento de los datos al ritmo de los cambios tecnológicos.

- Optar por las etiquetas de encabezamiento estándar que dan los sistemas abiertos y que pueden ser leídos por otros sistemas.

Si se utiliza un programa comercial, es difícil conservar la inteligibilidad de la información a largo plazo, sobre todo cuando se tiene que modificar el sistema o se debe instalar una nueva versión. En este caso se debe tomar la precaución de exigir al proveedor lo siguiente:

- Una interfaz que permita el acceso a la información con la ayuda de un sistema abierto.
- Una definición detallada de la estructura de las etiquetas de encabezamiento y si es posible el lenguaje fuente, la documentación completa y todos los elementos que se requieren para asegurar el funcionamiento del sistema.

2.5 La calidad de las imágenes

La verificación de la calidad de las imágenes se hace mediante tres operaciones:
-La toma de las imágenes
- su mejoramiento
- la compresión y descompresión de los datos digitalizados.

Para obtener una buena calidad de la imagen se deben respetar cuatro reglas:

1. El equipo (unidad de disco óptico y escánner) debe ser calibrado correctamente de acuerdo con las instrucciones del productor y ajustado de acuerdo con las necesidades del usuario.

Para los documentos que tienen valor a largo plazo o valor científico-cultural o para aquellos que se eliminarán después de digitalizados, se recomienda verificar la calidad de cada una de las imágenes digitalizadas. (Se debe recordar que de acuerdo con la Ley 7202, ningún funcionario público puede eliminar documentos sin la autorización de la Comisión Nacional de Selección y Eliminación de Documentos, ya que hay pena de prisión de 6 meses a 3 años).

Para los otros documentos se deben toma las precauciones de:

- Clasificar los documentos originales de acuerdo con los problemas de digitalización que se presenten
- Efectuar una prueba con un muestreo de documentos antes de iniciar la digitalización de todo un fondo documental
- Obtener la autorización de la persona encargada del proyecto.

2. La densidad de la digitalización se da en función de la calidad de la imagen y de la capacidad de almacenamiento que se desean. Se debe tener en cuenta la capacidad de memoria de los discos ópticos.

3. Para mejorar una imagen digitalizada se puede utilizar un algoritmo que limpia la imagen, suprimiendo los datos de una forma selectiva o automática. Esto

puede provocar una pérdida significativa de detalles. Cuando se utiliza este procedimiento, se debe conservar una imagen no mejorada o se debe asegurar que el algoritmo usado es reversible. Para los documentos que tienen valor de prueba o valor científico-cultural, se debe conservar siempre una imagen no retocada.

4. La compresión y la descompresión de los datos. La compresión reduce enormemente el volumen de las imágenes digitalizadas por medio de fórmulas matemáticas. Estas fórmulas pueden ser comerciales o estándar. Si se utiliza una fórmula comercial, se debe obtener del proveedor un algoritmo de compresión y descompresión sin pérdida de datos.

2.6 La funcionalidad del sistema de digitalización

La funcionalidad de un sistema de digitalización depende principalmente de una arquitectura abierta y de una compatibilidad descendente.

Una arquitectura abierta permite
- Mejorar los componentes del sistema sin degradar sus funciones y sin perder información
- Importar y exportar los datos
- Funcionar con componentes materiales de orígenes diferentes sin efectos negativos en los programas informáticos

Si se trabaja con una arquitectura comercial, se debe obtener del proveedor una interfaz que permita la compatibilidad con los estándares de los sistemas abiertos.

Un sistema con una compatibilidad descendente permite que las nuevas versiones se apoyen en versiones anteriores. Es decir, que se puede acceder a la información ya digitalizada y convertirla a la nueva versión, permitiendo la transferencia de la información sin que sufra ninguna modificación. En consecuencia, cuando se va a seleccionar un sistema de digitalización es necesario:

- Exigir al proveedor que toda actualización del sistema o todo nuevo sistema sea compatible con el sistema ofrecido.
- Transferir la información ya digitalizada para que pueda ser tratada en el nuevo sistema.

Por otra parte la interfaz de comunicación debe permitir un ligamen entre equipos provenientes de diferentes casas comerciales. Por lo tanto se hace necesario:

- Analizar cuidadosamente la interfaz ofrecida por el proveedor.
- Obtener del proveedor la documentación completa de la interfaz, tomando en cuenta tanto el equipo como los programas informáticos.

El índice sirve para recuperar las imágenes almacenadas en los soportes ópticos. El índice puede ser manual o automático. En el caso de un índice automático, este

se debe conservar en un disco magnético, pero es más recomendable conservarlo en el mismo soporte óptico que se conservan las imágenes. Asimismo, se recomienda duplicar el índice en varios soportes como una medida de seguridad.

La documentación técnica

Se debe obtener y conservar toda la documentación técnica relacionada con todos los componentes del sistema de digitalización con el fin de asegurar la funcionalidad del sistema a largo plazo. Esta documentación comprende:

- Programa fuente
- Modelo de base de datos
- Diagrama de flujo de datos
- Manual del usuario
- Manual técnico
- Versión beta por un tiempo limitado para probar la funcionalidad del sistema
- Toda la documentación que se considere necesaria sobre la aplicación y el mantenimiento del sistema

2.7 Protección de la información digitalizada

Para proteger los documentos digitalizados de una forma adecuada, es necesario:

- La adquisición de soportes y tecnología de registro y almacenamiento de alta calidad
- El control de la cantidad de los datos almacenados en el soporte
- La protección del soporte conservándolo en condiciones ambientales adecuadas
- Generar respaldos de la información documental de forma periódica, utilizando formatos homogéneos en los datos y la documentación, procurando establecer un solo medio de respaldo electrónico, por ejemplo, los discos compactos no regrabables (tipo Worm).

2.8 El control del acceso

En el caso de los hospitales que desean digitalizar los expedientes médicos, se debe asegurar la confidencialidad de la información mediante un control del derecho y del procedimiento del acceso, el cual debe darse sobre el lugar donde se produce la digitalización, los equipos y los programas que se van a utilizar. Es decir, que se debe regular el acceso a las instalaciones de cómputo, los archivos, los datos, procesos o servicios del sistema por parte de usuarios y establecer el rango de acción de cada usuario. Se deben establecer diferentes niveles de acceso de acuerdo de acuerdo con las necesidades de los usuarios y determinar si el acceso se va a dar por medio de palabras, nombres o números claves, tarjetas, firmas, voz, etc. Asimismo, se debe llevar un histórico de todos los usuarios que acceden a la información digitalizada.

2.9 Evolución de las tecnologías de información

Este aspecto exige una constante actualización tecnológica con el fin de conocer los nuevos equipos y programas que aparecen en el mercado para asegurar la migración de la información digitalizada. Además, se debe asegurar una perfecta adecuación entre los equipos y los programas para proteger la información que tiene valor legal o científico-cultural.

2.10 Las auditorías

Con el fin de garantizar la autenticidad, confiabilidad e integridad de los documentos digitalizados se deben realizar auditorías periódicas al sistema y a todos los procedimientos de digitalización. Esas auditorías pueden ser internas o externas.

Los objetivos de las auditorías se pueden resumir en:

- Determinar la conformidad de los elementos del sistema de almacenamiento y de los procedimientos según las exigencias prescritas.
- Determinar la eficacia global del sistema y sus procedimientos.

Para lograr los objetivos el auditor deberá revisar el acceso del personal autorizado a los datos, los controles de procesamiento de datos, la integridad de los datos, los tipos de respaldos, la manipulación de los datos por personal confiable, los soportes de almacenamiento de los datos, etc.

La normativa internacional establece hacer una auditoría una vez al año.

3. FASE LEGAL

En varias leyes nacionales se le da al documento electrónico el mismo valor que al documento en papel, incluyendo el dictamen C-283-98 de la Procuraduría General de la República. Sin embargo, se debe tener presente que en casi todas las leyes se estipula que el documento electrónico tiene valor legal y probatorio siempre y cuando se pueda demostrar su autenticidad, integridad y confiabilidad. Dado lo anterior, es necesario esperar cuales son las directrices que va a emitir la Comisión de Gestión de Documentos en Soporte Electrónico.

A modo de ejemplo, se pueden citar algunos aspectos que en países como Canadá y Francia se han tomado como base para establecer el valor legal de los documentos conservados en soporte electrónico, a saber: una adecuada gestión documental, los componentes del sistema de digitalización, los controles del acceso, la funcionalidad del sistema, el control de calidad de las imágenes, la seguridad del sistema, etc.

BIBLIIOGRAFIA

AFNOR. Norme francesa NF Z 42-013 Recommandations relatives á la conception et a l'exploitation de systemes informatiques en vue d'assurer la conservation et l'integrité des documents stokés dans ces systémes. París. 1999.

BÉLANGER, Micheline, Jean Maurice Demers y Claude Minotto. Guide d'imagerie numérique. Numérisation de l'información consignée sur des supports traditionnels. Archives Nationales du Québec. 1998. Traducción libre realizada por María Teresa Bermúdez Muñoz.

CONSEJO INTERNACIONAL DE ARCHIVOS. Guide pour la gestion de documents éléctroniques du point de vue archvistique. París. 1996.

JACQUESSON, Alain y Alexis Rivier. Bibliothéque et documents numériques. Edition du Cercle de la Libraire. París. 1999.

MTIC. Guide pour la conservation des informations et des documents numériques. París. 2001.

PASCON, Jean Louis e Isabelle Pottier. Archivage électronique. París. 2000.