



SISTEMA NACIONAL DE DOCUMENTACIÓN HISTÓRICA

GUÍA GENERAL DE DIGITALIZACIÓN DE DOCUMENTOS

Versión 1.0

Agosto de 2019

Indice

0. Presentación y objetivos

- 0.1 Breve presentación
- 0.2. Finalidad
- 0.3. Destinatarios
- 0.4. Alcance y contenido

1. PRIMERA PARTE. Tareas previas

- 1.1. Criterios de selección de la documentación a digitalizar
- 1.2. Tareas previas archivísticas
- 1.3. Tareas previas de preservación de la documentación analógica
- 1.4. Análisis y evaluación de la tecnología requerida
 - 1.4.1. Dispositivos de captura
 - 1.4.2. Dispositivos de almacenamiento de los objetos digitales

2. SEGUNDA PARTE. Requisitos para la captura de imágenes

- 2.1. Especificaciones técnicas para la creación de imágenes máster
- 2.2. Especificaciones para la creación de copias derivadas para la web
- 2.3. Metadatos
- 2.4. Nombramiento de documentos
- 2.5. Tratamiento de las imágenes
- 2.6. Control de la calidad del proceso
- 2.7. Recepción y devolución del material original y sus instrumentos de clasificación y descripción
- 2.8. Reconocimiento Óptico de Caracteres (OCR)
- 2.9. Marca de agua
- 2.10. Organización de los archivos digitalizados

3. TERCERA PARTE. Estrategias de preservación de los objetos digitales

0. Presentación y objetivos

0.1. Breve presentación

El Sistema Nacional de Documentación Histórica (en adelante, SNDH) presenta esta **Guía General de Digitalización de Documentos** como parte de un conjunto de acciones para la sistematización del tratamiento de la documentación histórica de los archivos en nuestro país, en las distintas etapas comprendidas en este proceso. Si bien hay importantes aportes de instituciones y subsistemas que han diseñado guías directrices ad-hoc, es importante señalar la carencia de una política nacional explícita respecto de la producción y gestión de documentos digitales, tanto nativos como producto de digitalizaciones. En consecuencia, tampoco se han consensuado criterios profesionales, funcionales y tecnológicos volcados en instrumentos técnicos (manuales, reglamentos y demás) para la digitalización del patrimonio documental que cada Archivo tiene bajo su responsabilidad.

Con la intención de minimizar debilidades o problemas ya detectados en las prácticas y experiencias de proyectos de digitalización en distintas instituciones, a los efectos de avanzar en la sistematización de las actividades que surgen de la administración electrónica en general y de los procesos de digitalización en particular, y sin dejar de capitalizar el aprendizaje extraído de los distintos trabajos realizados con anterioridad, consideramos de vital importancia esta **Guía General de Digitalización de Documentos** para el desarrollo de proyectos por parte de las instituciones adherentes al SNDH, así como también un instrumento didáctico y sistémico, las directrices con las cuales poder transitar la organización dinámica y efectiva de distintos tipos de instituciones que custodian fondos de archivo.

Esta Guía es un producto desarrollado por el Grupo de Digitalización del SNDH, tomando como base las tareas desarrolladas por el Grupo de Trabajo de Tecnologías de la Información y las Comunicaciones (TICs) del Archivo General de la Nación (AGN), así como los valiosos aportes de reconocidos expertos argentinos en ese campo, en el marco de las reuniones de los grupos especializados/técnicos del SNDH.

0.2. Finalidad

La finalidad de esta **Guía** es dotar tanto a los archivos que se adhieran al SNDH, así como a aquellos que carezcan de protocolos normalizados, de una herramienta que les permita adaptar sus procedimientos a normas consensuadas o institucionalizadas. Los contenidos de esta **Guía** están dirigidos tanto a las instituciones como a sus profesionales con el objetivo de diseñar criterios generales comunes para la planificación, desarrollo y supervisión de proyectos de digitalización del patrimonio documental que cada institución tiene bajo su responsabilidad.

En particular, esta guía apunta a la digitalización de fondos de archivo, entendiendo que no podrá desarrollarse ningún proyecto de digitalización sin ser acompañado por un correcto desarrollo archivístico. Es por eso que esta guía se divide en tres partes: en la primera se encuentran los lineamientos previos indispensables en materia archivística antes de llevar adelante cualquier

¹ Fernando Boro, Matías Butelman, Claudia Patricia Cabouli, Luis Alberto Farías, Carina Frid, Mariana Nazar, Martín Oliver, Andrés Pak Linares, Graciela Teresa Perrone y María Eugenia Sik

² Guía de Digitalización del Archivo General de la Nación. Fecha: 19/02/2018

proyecto de digitalización (parte 1), las especificaciones técnicas recomendadas para digitalizar distintos tipos de material (parte 2) y las estrategias generales para la preservación de los objetos digitales (parte 3). Se incluyen cuatro anexos para poder complementar la información de esta guía. Éstos se irán actualizando y diversificando en la medida en que avance la sistematización del conocimiento en esta materia.

Esta **Guía**, con sus directrices, debe tenerse en cuenta tanto para desarrollar flujos de trabajo al interior de las instituciones, como para los procesos de digitalización que se decidan tercerizar. A la hora de redactar instrumentos de contratación y/o proyectos para la obtención de financiamiento en el marco del SNDH, las condiciones y productos esperables deberán ajustarse a lo establecido en esta **Guía**. Los anexos servirán como sugerencias técnicas para la elaboración y el desarrollo de los proyectos que se presenten.

0.3. Destinatarios

Destinatarios principales

- Instituciones adherentes al SDNH
- Instituciones que custodien fondos de archivo en su acervo y que necesiten guías para su gestión documental institucional,
- Directores de área y/o personal con capacidad de decisión dentro de las instituciones

Destinatarios secundarios

- Personal de las instituciones adherentes y de aquellas que utilicen esta Guía de Digitalización
- Ciudadanía en su conjunto (Ley 27.275, de Acceso a la Información Pública)

0.4. Alcance y contenido

Aunque pueden hallarse en la actualidad otras finalidades para la digitalización, resulta evidente que la principal responde a la necesidad de realizar copias de documentos. Las fotocopias en papel y en microfilm han dado paso al formato digital que, si bien aún tiene como asignatura pendiente demostrar su perdurabilidad en el tiempo en el largo plazo, presenta ventajas lo suficientemente significativas como para que tanto las instituciones públicas o privadas hayan depositado en el entorno digital todas sus expectativas.

Para los archivos históricos, en consecuencia, es una poderosa herramienta para permitir una mayor rapidez en la recuperación de la información deseada, así como también el acceso simultáneo y/o remoto de diversos interesados a la misma, al tiempo que preserva al material original de una excesiva manipulación.

Por otro lado, el entorno digital requiere garantías básicas de preservación, integridad, autenticidad y accesibilidad a lo largo del tiempo, lo que implica un compromiso institucional integral, que debe poder ser verificado en la asignación de recursos materiales y humanos, así como en la definición de políticas y herramientas operativas y técnicas, tales como manuales, reglamentos, protocolos, etc.

En ese sentido, la presente **Guía** tiene por objetivo sentar las bases comunes para la normalización de los proyectos de digitalización que se desarrollen en las distintas instituciones adherentes al SNDH. Para su implementación, es ineludible y necesario el compromiso institucional de directivos, profesionales y personal operativo.

Asimismo, es preciso enfatizar que los proyectos de digitalización deben integrarse con otros procesos de la gestión del patrimonio documental tales como los de clasificación, descripción, conservación, difusión y atención de usuarios por un lado (que también deben normalizarse, si aún no lo estuvieran), y con las áreas de apoyo a la gestión, esto es: las que se relacionan con redes, sistemas, equipamiento, monitoreo y servicio técnico de las herramientas de TICs necesarias para la mencionada implementación.

1. PRIMERA PARTE. **Acciones previas al proceso de digitalización**

1.1. Criterios para la selección del material

La determinación final respecto de la selección de la documentación a digitalizar debe ser una decisión de la autoridad máxima de cada institución, de acuerdo con su misión institucional y las funciones que de ella se derivan. Cuando la institución conste de departamentos o áreas que eleven proyectos para la aprobación de la dirección, esta Guía será la que estandarice dicho proceso.

Esto es así, por un lado porque de esa decisión depende la asignación de recursos humanos y materiales para los proyectos, en sintonía con la política general de la institución; pero, por el lado de los departamentos o áreas existentes, porque es en esos ámbitos en donde se registran las variables que dan cuenta del valor histórico y patrimonial, el estado de conservación y el nivel de consulta de los fondos, secciones y series documentales que tienen a su cargo.

Además de la legislación vigente, a la hora de proyectar la digitalización para hacer accesible el patrimonio documental así digitalizado, cada institución deberá desarrollar un análisis DAFO / FODA (ver modelo en el anexo 5.2) que permita generar los debidos consensos para establecer factibilidades y prioridades y factibilidad de las propuestas.

Esto, por supuesto, debe hacerse en el contexto normativo, organizacional y tecnológico de cada institución.

Esta guía propone una serie de criterios a ponderar para el desarrollo de proyectos de digitalización que deberán evaluarse en conjunto para realizar un diagnóstico más certero de la necesidad de digitalización de un acervo documental y/o el establecimiento de prioridades a la hora de planificar la digitalización a largo plazo:

- **a) Contenido:** debe evaluarse la relevancia del contenido intelectual y/o el valor patrimonial del material que se propone digitalizar, en relación con la posibilidad de mejorar la difusión y el acceso al mismo.
- **b) Demanda:** una alta frecuencia de consulta del material original por un lado supone -más allá de los cuidados implementados- un cierto nivel de degradación. Por otra parte, esa misma

frecuencia es indicativa de la importancia que el material tiene para la comunidad de usuarios de cada institución. Por ambos motivos, podría concluirse en la necesidad de digitalizar el material tanto para disminuir su deterioro como para mejorar las condiciones de acceso, incluido aquel de carácter remoto y simultáneo.

- **c) Acceso:** en caso de contar con material con limitaciones de accesibilidad física, ya sea por malas condiciones de conservación, difícil lectura, escucha, visualización u obsolescencia inminente de las herramientas de reproducción, la digitalización puede ser una alternativa para favorecer, o incluso permitir, el acceso al mismo. Si las condiciones de preservación del material son muy malas es imprescindible la restauración como paso previo a la digitalización.
- **d) Derechos de autor y protección de datos personales:** se priorizarán aquellos documentos en dominio público o con total potestad de los derechos por parte de la institución. Las limitaciones al acceso de orden legal, así como la normativa referida a la propiedad intelectual debe ser analizada con anterioridad a la propuesta de digitalización, dado que si éstas pusieran límites a la difusión de las imágenes digitalizadas, quizás los recursos invertidos en el proceso terminen careciendo de justificación, a menos que se trate exclusivamente de copias de resguardo.
- **e) Estado del tratamiento archivístico previo:** la adecuada identificación, clasificación y descripción archivística del material a digitalizar es condición excluyente para proponer el proceso de digitalización. En el caso del material bibliográfico, corre el mismo presupuesto pero bajo los estándares que se correspondan con ese tipo de documentos.
- **d) Equipamiento requerido y recursos vinculados:** es necesario evaluar el costo de adquisición tanto del equipamiento para la obtención de imágenes digitalizadas como de todos los recursos vinculados (capacidad de las computadoras, servidores, etc.). También puede considerarse, la tercerización del servicio y/o el desarrollo de un convenio interinstitucional, siempre en los términos de la presente **Guía**.

1.2. Tareas previas archivísticas: identificación, clasificación y descripción

Como se mencionaba anteriormente, toda la documentación cuya digitalización se proponga debe estar debidamente identificada, clasificada y descripta, además de estar en condiciones de conservación adecuadas para su digitalización (ver 1.3).

Una adecuada clasificación y los instrumentos de descripción, además de dotar del necesario contexto a los documentos, representan conocimiento del material, control del mismo y las referencias necesarias para su usabilidad; en términos del proceso de digitalización, además:

- **a)** Permite, con el conocimiento del volumen exacto de la documentación, evaluar tiempos del proyecto.
- **b)** Permite la trazabilidad de los documentos originales durante el proceso: salida y retorno de los documentos originales a su lugar de conservación
- **c)** Permite la creación y asegura el seguimiento de los archivos digitales en forma normalizada, incluyendo la adecuada elección de los dispositivos de captura en función de los requisitos que imponga cada tipo de original por su estado, formato, soporte, etc.
- **d)** Colabora con la aplicación de medidas de control de calidad y supervisión del proceso
- **e)** Facilita el control, la gestión y la relación de los documentos originales con sus imágenes digitales, en el marco de un sistema integral de administración del patrimonio documental.

El Cuadro de Clasificación y los instrumentos de descripción se entregarán, junto con la documentación original, a la persona encargada de la digitalización. Es preciso reiterar que toda documentación que sea seleccionada para su digitalización, independientemente del soporte, deberá ser analizada y descripta en contexto.

Además de las necesidades particulares de descripción para cada agrupación documental, en los proyectos de digitalización deberá consignarse, como mínimo, la información correspondiente a los elementos obligatorios de la Norma ISAD(G) que, además, serán parte de los metadatos obligatorios.

- Código de referencia
- Título
- Fecha
- Nivel de la descripción
- Volumen
- Soporte
- Nombre del productor

1.3. Tareas previas de preservación de la documentación analógica

Con anterioridad a la digitalización de los documentos se verificará el estado de conservación y la integridad de las unidades documentales, a los efectos de establecer si están en condiciones de ser sometidos a un proceso de digitalización o, en su defecto, cuánto tiempo demorará su preparación y qué recursos serán destinados a tal fin.

Si bien esto requiere instrumentos operativos específicos del área correspondiente, en líneas generales los documentos a digitalizar deben estar libres de cualquier elemento que obstaculice la digitalización. Si corresponde, se deben intervenir los documentos (restauración, limpieza, etc.) siempre de acuerdo a las características específicas de cada tipo de soporte, a los efectos tanto de proteger el documento original como de mejorar o posibilitar la captura, y evitar daños en el equipamiento de digitalización manteniendo el buen estado de funcionamiento de los mismos.

En ningún caso y bajo ningún punto de vista el personal que diagnostica el estado de conservación de los documentos para evaluar la posibilidad de su digitalización, ni el que interviene preparándolos para ese procesamiento, puede alterar el orden de los mismos, su integridad, o los instrumentos de descripción que los acompañan; sin embargo sí puede -y en ciertos casos, debe- agregar información relativa a esa intervención en éstos últimos.

Toda información relevante que surja de la intervención en términos de conservación debe consignarse en forma específica, y en relación con la demás información de los documentos que integran el proyecto y, si corresponde, volcarse en los metadatos.

1.4. Análisis y evaluación de la tecnología requerida

1.4.1. Dispositivos de captura

La selección y utilización del hardware se hará en función del soporte y las características de los documentos originales. Se optará por el equipo más adecuado para llevar las tareas digitalización, teniendo en cuenta como premisa básica que los originales sufran la menor manipulación y desgaste posible.

En base a los parámetros de resolución y calidad establecidos para cada tipo de formato, se optará por el más adecuado, si bien puede haber equipos que tengan mayores resoluciones ópticas reales se deberá evaluar si ese tipo de equipo es el mejor, teniendo en cuenta lo antes mencionado.

Es recomendable que en la captura de imágenes masters no se utilicen las herramientas de “mejora de captura” (corrección de histograma, eliminación de rayas, mejora de sonido, etc.) que traen en su mayoría los softwares controladores de los escáneres ya que al tratarse de la obtención de un master la mejor opción es buscar la mejor calidad en base a ajustes manuales.

Respecto del software para la visualización, post proceso, conversión de archivos por lotes, es recomendable optar por programas con licencia de software libre. Esto significa un ahorro en los costos de obtención y mantenimiento de licencias y permite contar con el código fuente de los mismos, lo que contribuirá directamente a la preservación digital.

1.4.2. Dispositivos de almacenamiento de los objetos digitales

Los objetos digitales obtenidos pueden ser conservados en diversos soportes físicos. Estos pueden ir desde discos duros internos, discos duros externos, pendrives, CD, DVD, hasta cintas magnéticas. Por supuesto cada soporte tiene sus ventajas o desventajas por lo que quizás lo más recomendable sea una mezcla de ellos. Esto es utilizar el disco interno de una computadora o un servidor y varias copias de lo mismo en discos externos o en otra computadora. Esto también puede incluir almacenamiento en soportes ubicados físicamente localmente o en ubicaciones remotas.

Esta idea se basa en el sistema LOCKSS (Lots of Copies Keep Stuff Safe, o en español muchas copias mantienen a las cosas seguras) utilizado por ejemplo en proyectos del Public Knowledge Project como el Open Journal System (OJS). Esto implica que al encontrarse las colecciones de objetos digitales diseminadas en diversos medios y ubicaciones, se garantizaría una mayor probabilidad de una conservación a largo plazo.

PARTE 2. Requisitos para la captura de imágenes

Si bien el desarrollo de los ítems subsiguientes debe volcarse en herramientas operativas concretas, la presente Guía orienta sobre los criterios generales a tener en cuenta para las mismas. En los anexos a la presente guía se encontrará información adicional complementaria que profundiza los puntos desarrollados debajo.

2.1. Especificaciones técnicas para la creación de imágenes máster

Como resultado de la digitalización de documentos se obtendrán dos tipos de copias, a saber:

1. *Máster*, son reproducciones digitales de alta calidad realizadas con fines de preservación, en formatos sin compresión (o, si es factible, con compresión sin pérdida) que afecte a la calidad del objeto digital
2. *Derivada*: a partir de cada máster deberá crearse un archivo apto para la consulta en el formato que corresponda según el tipo de documento

Cabe puntualizar que por archivos de consulta se entiende aquél que se genera para la consulta web o de usuarios. Para el caso, de archivos solicitado bajo demanda, el peso del archivo dependerá de las necesidades del interesado (publicaciones, exposiciones, uso particular, etc.).

Para dar estructura a los archivos digitales y para adecuarla a la que se ha creado en el repositorio, se establece una estructura interna de carpetas que debe responder a la estructura del Cuadro de Clasificación del material digitalizado y que se indica a continuación.

- MASTER

Nombre del fondo

Nombre del subfondo (si corresponde)

Nombre de la sección (si corresponde)

Nombre de la serie

Nombre de la unidad documental compuesta o simple

Nombre archivo. Extensión

- CONSULTA

Nombre del fondo

Nombre del subfondo (si corresponde)

Nombre de la sección (si corresponde)

Nombre de la serie

Nombre de la unidad documental compuesta o simple

Nombre archivo. Extensión

Si bien cada proyecto puede tener determinadas particularidades, tanto de los materiales originales como de las necesidades respecto de la imagen digital obtenida, a continuación, la presente **Guía** brinda el siguiente cuadro orientativo para el desarrollo general de los proyectos de digitalización que se propongan al SNDH.

| Características materiales de originales | Resolución de captura/ tipos de archivo a obtener | Modos de Color (Escala de grises/color RGB) |
|--|---|---|
| PUBLICACIONES PERIODICAS (formato hasta A4) | 300 dpi/TIFF sin comprimir | Escala de grises, a 8 bits |
| PUBLICACIONES PERIODICAS (formatos A3 o superior) | 200 dpi/TIFF sin comprimir | Escala de grises, a 8 bits |
| TEXTO IMPRESO (formato hasta A4) | 300 dpi/TIFF sin comprimir | Escala de grises, a 8 bits |
| TEXTO IMPRESO (formatos A3 o superior) | 200 dpi/TIFF sin comprimir | Escala de grises, a 8 bits |
| MANUSCRITO (diferentes formatos) | 300 dpi/TIFF sin comprimir | Color, a 24 bits |
| FOTOGRAFIA / OPACOS B/N (diferentes formatos) | 600 dpi/TIFF sin comprimir | Color, a 24 bits |
| FOTOGRAFIA/OPACOS COLOR (diferentes formatos) | 600 dpi/TIFF sin comprimir | Color, a 24 bits |
| FOTOGRAFIA/NEGATIVOS Y DIAPOSITIVAS B/N | 800 / 2.400 dpi/TIFF sin comprimir (según tamaño del original) | Color, a 24 bits |
| FOTOGRAFIA/ NEGATIVOS Y DIAPOSITIVAS COLOR | 800 / 2.400 dpi/TIFF sin comprimir (según tamaño del original) | Color, a 24 bits |
| FOTOGRAFIA/ PLACAS | 800 / 2.400 dpi/TIFF sin comprimir (según tamaño del original) | Color, a 24 bits |
| MAPAS, PLANOS Y PERGAMINOS (formatos hasta A4) | 600 dpi/TIFF sin comprimir | Color, a 24 bits |
| MAPAS, PLANOS Y PERGAMINOS (formatos A3 ó superior) | 300 dpi. /TIFF sin comprimir | Color, a 24 bits |
| AUDIO | (formato .wav) | 24 bits. 96000 khz |
| IMAGEN EN MOVIMIENTO | H264 (formato .avi) | |

2.2. Especificaciones para la creación de copias derivadas para la web

Estas diferirán de acuerdo a las características del material y al tipo de proyecto. En líneas generales, contarán con:

- menor resolución (para las imágenes fijas, se recomienda 72 dpi)
- menor en tamaño
- un archivo comprimido (por lo general, JPG o PDF)

Es requisito que estas copias sean nombradas de la misma manera que la copia máster, siempre que la edición de las copias derivativas no produzca varios objetos digitales diferenciados (ej. la digitalización de una cinta de audio / video y los tracks que de ella se desprendan).

2.3. Metadatos

Dado que los metadatos son datos estructurados que describen el contexto, el contenido y la estructura de los documentos y su gestión a lo largo del tiempo, en cualquier proyecto de digitalización, las organizaciones deberán contemplar un tratamiento integral en el que se incluya la gestión de los metadatos y, si corresponde, la firma electrónica.

La presente **Guía** establece los metadatos mínimos y obligatorios que deben utilizarse y gestionarse en la totalidad de los proyectos de digitalización que se desarrollen en la órbita del SNDH.

De acuerdo con eventuales necesidades específicas, se podrán sumar campos de metadatos a los listados a continuación, pero en ningún caso reemplazarlos.

Se recomienda establecer con claridad quién, cuándo y cómo consigna los distintos metadatos sobre el proyecto de digitalización.

Metadatos mínimos y obligatorios:

| Denominación | Descripción |
|----------------------|--|
| Archivo | Nombre del Archivo que custodia la unidad de descripción. |
| Código de referencia | Código que identifica la unidad de descripción. Debe coincidir con la estructura del fichero digital |
| Nivel de descripción | Identificar el nivel de la unidad de descripción |
| Serie documental | Nombre de la serie documental a la que pertenece. |
| Título | Campo de texto que da nombre al objeto. Se trata de una descripción breve del documento, expediente, etc. |
| Fechas del original | En caso de tratarse de una unidad documental simple, se reseñará la fecha de producción de la misma. En caso de tratarse de un expediente, se reseñará la fecha de apertura y de cierre del expediente. |
| Volumen del original | Se consignará el volumen de la unidad de descripción que se ha digitalizado. |



| Denominación | Descripción |
|---|---|
| Soporte del original | Se consignará el soporte de la unidad de descripción que se ha digitalizado. |
| Nombre del productor | Nombre del productor del fondo |
| Condiciones de acceso | Identificación de los diferentes niveles de acceso al objeto según el tipo de información que contenga. |
| Responsable de la digitalización / migración | Identifica operador de la digitalización o de la conversión de formato. |
| Fecha de la digitalización / migración | Se consignará la fecha en la que se ha realizado la digitalización o conversión de la unidad de descripción. |
| Formato digital | Extensión del archivo obtenido (p.ej.: TIF) |
| Resolución espacial (dpi) | Indicar la resolución en dpi de las imágenes digitales |
| Tamaño (Megabytes) | Tamaño del archivo obtenido expresado en Megabytes |
| Tipo de imagen | Master o Derivada (es decir, de consulta) |
| Profundidad de bits | 1bit (bitonal); 8bits (grises); 24bits (color); 48bits (color) |
| Compresión | Con o sin (en caso afirmativo indicar tipo, ej.: TIF LZW) |
| Espacio de color | Gray Gamma 2.2; sRGB; Adobe RGB (1998). Indicar lo que corresponda (no aplica con imágenes de 1bit). |
| Software | Utilizado para la captura y/o proceso, nombre y versión |
| Dispositivo o equipo | Utilizado para la captura y/o proceso, marca y modelo |
| Observaciones | (No debe utilizarse el campo para otros metadatos específicos, los cuales deben consignarse en el proyecto particular que los requiere) |

2.4. Nombramiento de archivos

El nombre de los objetos digitales obtenidos guardará vinculación con el fondo digitalizado y se recomienda definir un patrón codificado, homogéneo y preciso para los distintos documentos, evitando la utilización de nombres descriptivos. Es recomendable no utilizar símbolos extraños que puedan ocasionar problemas en la búsqueda, migración o trabajo en lote con los archivos. Algunos de los símbolos a evitar son:

/ o \
tildes varios
(o)
¿? o i !
.
:
*
“
< o >
|
,
ñ

Es importante, además, evitar los nombres excesivamente largos, ya que pueden dificultar la edición, el renombrado en lote u otras actividades concomitantes al trabajo con los objetos digitales.

2.4. Tratamiento de las imágenes

El tratamiento de las imágenes digitalizadas debe balancear adecuadamente tanto las evidentes posibilidades de mejorar el acceso a las mismas, como los recursos que se requieren para su implementación y, no por último menos importante, la alteración en la imagen digital de las características originales del documento.

En el caso de que se realice la edición de imágenes, ya sea para mejorar el acceso, o por cuestiones ligadas a la protección de datos sensibles, derechos de imagen, etc. ese tratamiento debe hacerse exclusivamente sobre las copias derivadas, debe ser aprobado por la máxima autoridad (que es la que debe autorizar la asignación de recursos humanos y materiales a la tarea) y debe trabajarse en estrecha colaboración con las áreas responsables de la documentación original.

Asimismo, cualquier tratamiento debe consignarse en los metadatos, en el campo “Observaciones” en general, pero las características específicas del tratamiento deben estar disponibles para los usuarios que las requieran.

2.5. Control de la calidad del proceso

Es importante verificar la calidad del proceso para corroborar el cumplimiento de los objetivos.

La valoración de las digitalizaciones debe realizarse en tres niveles (visualización/sonido, codificación y datos asociados), sobre la base de un muestreo del 20% de los objetos digitales. En caso de detección de irregularidades, el muestreo debe aumentar hasta deducir el alcance de las dificultades.

La verificación abarcará la evaluación del resultado de la digitalización (nivel de fidelidad, ausencia de defectos), el análisis de la codificación binaria de los archivos digitales, la organización lógica dentro de los medios de almacenamiento (comprobar que se asignaron los nombres correctos a los documentos y se los estructuró adecuadamente en los discos de computadora, externos u otros medios digitales) y el chequeo de los datos asociados (correcta asignación de metadatos a cada archivo digital).

2.6. Recepción y devolución del material original y sus instrumentos de clasificación y descripción

En instituciones que cuenten con estructura organizativa suficiente, el área encargada de la digitalización deberá recibir (junto con la autorización de la dirección de la institución y el informe técnico de conservación) el material debidamente organizado, tanto física como conceptualmente, con los instrumentos de descripción correspondientes.

Al finalizar el proceso de digitalización, debe devolverse la documentación, respetando el orden original en el que se encontrara, sin alterar los instrumentos de descripción y sin empeorar las condiciones de conservación.

2.7. Reconocimiento Óptico de Caracteres (OCR)

El OCR (*Optical Character Recognizer*), es una herramienta más que no reemplaza a los procesos archivísticos previos, éste debe determinarse y justificarse con anterioridad al desarrollo del escaneo.

Al respecto, es necesario aclarar que muchas veces la recuperación de la información digitalizada es delegada al mencionado OCR. Pero el hecho de poseer documentos digitales en formato de texto y no de imagen resultado de OCR no asegura que la información pueda ser fácilmente recuperada.

Luego del OCR deben efectuarse controles de calidad manuales que contrasten el resultado de este proceso con la imagen digital o bien con el documento original. Asimismo, el hecho de poseer un documento cuyo contenido sea un texto en lugar de una imagen no implica que posea metadatos, ni que estos estén completos y/o correctamente asignados.

2.8. Marca de agua

En ningún caso se insertarán marcas de agua en las copias máster.

Respecto de la posibilidad de insertar una marca de agua en las copias de consulta, debe analizarse su conveniencia en relación con los recursos y el tiempo que debe emplearse y el sentido de aplicarla.

En ese sentido, y respecto de la identificación de la procedencia, por ejemplo, esto debe consignarse en los metadatos, lo mismo que el régimen de acceso y, eventualmente de (re)utilización, por lo que no parece una adecuada relación costo / beneficio, si sumamos la tecnología actual que permite removerla.

El cambio de paradigma respecto del acceso, así como un análisis pormenorizado respecto de la aplicación de restricciones por derechos de uso, intelectual, etc... son los vectores que, mientras se implementa una política institucional por parte de la Dirección General, deben guiar el análisis de cada proyecto de digitalización, en lo que hace a la incorporación de marcas de agua.

2.9. Organización de los archivos digitalizados

Las digitalizaciones estarán organizadas en una estructura jerárquica de carpetas, con forma de árbol, que refleje el esquema de organización de los fondos, dentro de ella se instalarán los ficheros con los documentos digitalizados.

Esta estructura aparecerá duplicada, una de ellas destinada a las imágenes de conservación y otra para las de consulta.

De manera periódica los responsables de la custodia de las copias comprobarán la legibilidad de todos los soportes en los que se han alojado las copias digitales. El objetivo es que se mantenga el acceso continuo a los datos digitales.

En orden a prevenir la obsolescencia tecnológica, y con una periodicidad que se establecerá en cada caso, deberá llevarse a cabo la migración de los soportes de almacenamiento que contengan tanto las copias de conservación como las copias de consulta. Lo mismo podría aplicarse a la migración de formatos.

Asimismo, debe implementarse un sistema de back up en servidores distintos a aquellos en los cuales en los que se conservan estos máster y copias derivadas.

3. TERCERA PARTE.

Estrategias de Preservación de los Objetos Digitales

Uno de los desafíos más importantes en esta materia es mantener la autenticidad, fiabilidad, integridad y usabilidad en el tiempo. En este sentido, resulta esencial instalar en todo proyecto de digitalización una adecuada política de preservación antes de comenzar las tareas, para asegurarse que la inversión realizada en digitalización esté respaldada en el mediano y largo plazo.

La preservación del material digital no puede ser esporádica. Es imprescindible una estrategia de preservación digital proactiva desde el estadio más inicial de la vida del material digital. La inacción, el olvido y el abandono del material digital lo harán inaccesible e inutilizable en un corto período de tiempo.

Las amenazas contra la pervivencia del material digital giran en torno a tres conceptos fundamentales: la inestabilidad de los dispositivos de almacenamiento, la obsolescencia de las tecnologías de acceso y la problemática de la integridad y autenticidad del material digital.

Los artefactos digitales son elementos (aún) de gran fragilidad que se deterioran rápidamente y pueden fallar en cualquier momento por agentes externos y ambientales como el calor, la humedad, los contaminantes, el uso intensivo o una manipulación inadecuada.

La preservación y cuidado de estos dispositivos de almacenamiento es necesaria pero no suficiente para la correcta preservación de la información digital. Se ha observado que la vida útil de un dispositivo de almacenamiento es sorprendentemente corta. Sin embargo, aún excede al tiempo por lo que puede ser leído e interpretado por el software o hardware adecuado.

Los objetos digitales no sólo deben conservarse interpretables sino que también deben ser fiables, es decir, auténticos e íntegros.

Estos requerimientos deben ser tenidos en cuenta a la hora de diseñar los protocolos de conservación a largo plazo con estrategias como la migración de los datos que sin duda altera la cadena de bits.

ANEXO I - Glosario

Accesibilidad: ver usabilidad.

Autenticidad: Consiste en la acreditación de un documento de archivo de ser lo que pretende ser sin alteraciones o corrupciones. Los documentos auténticos son los que han mantenido su identidad e integridad al paso del tiempo gracias a la evidencia de su carácter, requisitos o circunstancias inherentes.

Clasificación: Operación intelectual que consiste en el establecimiento de las categorías y grupos que reflejan la estructura jerárquica del fondo. Constituye la información esencial del contexto en el que han sido producidos y de las relaciones existentes entre documentos.

Captura: acción mediante la cual se obtiene la imagen digitalizada a partir de la representación original.

Descripción archivística: Elaboración de una representación fidedigna de la unidad de descripción y, en su caso, de las partes que la componen mediante la recopilación, análisis, organización y registro de la información que sirve para identificar, gestionar, localizar y explicar los documentos de archivo, así como su contexto de producción.

Dispositivos de captura: Dispositivo a través del cual se obtiene la imagen digitalizada desde la representación original (analógica). Son dispositivos de captura las cámaras digitales y los scanners.

Documento de archivo: Información registrada, producida o recibida durante la iniciación, desarrollo o terminación de una actividad personal o institucional y que incluye contenido, contexto y estructura suficientes para servir como testimonio de esa actividad.

Se caracteriza, en líneas generales, por ser único y estar interrelacionado a otros documentos por insertarse en un sistema más amplio (el fondo de archivo). Este amplio concepto cubre todas las tipologías documentales. Los documentos pueden adoptar varias formas y representaciones. Se suelen representar como elementos de información delimitados de forma lógica, por ejemplo, como documentos diferenciados. Pero cada vez con mayor frecuencia encontramos documentos en forma de elementos de información distribuidos, como las bases de datos relacionales y los documentos compuestos.

Fiabilidad: implica que el contenido de un documento sea cierto. La fiabilidad es, por lo tanto, una noción relativa, y por esto se habla del grado de fiabilidad de un documento. 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).

Fondo de archivo: La totalidad de los documentos producidos o recibidos por una persona física o jurídica, pública o privada en razón de su función o actividad.

Identificación archivística: Primera etapa de la labor archivística, que consiste en indagar, analizar y sistematizar las categorías administrativas y archivísticas que sustentan la estructura de un fondo. Mantiene relación directa e inmediata con la clasificación.

Integridad: Propiedad o cualidad de un documento cuando se encuentre completo e inalterado. Manejabilidad: ver usabilidad.

Marca de agua: técnica que consiste en insertar un mensaje (oculto o no) en el interior de un objeto digital (imágenes, audio, vídeo, texto).

Metadatos: En el ámbito de la gestión de documentos, datos que describen el contexto, el contenido y la estructura de los documentos y su gestión a lo largo del tiempo.

Objeto digital: Un objeto que puede ser representado por una computadora, tal como un tipo de documento informático generado por un sistema particular o una aplicación de software, por ejemplo, un documento de procesador de textos, una hoja de cálculo o una imagen. Un documento digital puede comprender uno o más objetos digitales. (Archivo Nacional de Australia, traducción propia).

OCR (Optical Character Recognition): se trata de un proceso de reconocimiento óptico de caracteres que por medio de un software “lee” una imagen escaneada de texto y la traduce a texto para su posterior uso (ASCII, u otros formatos de texto).

Ordenación: Operación archivística realizada dentro del proceso de organización, que consiste en establecer secuencias dentro de las categorías y grupos, de acuerdo con el principio de procedencia y orden original // La operación de reponer el orden original de las unidades dentro de las series o grupos constituídos.

Preservación: Los procesos y operaciones involucradas en asegurar la supervivencia técnica e intelectual de documentos auténticos a lo largo del tiempo. La preservación incluye el control ambiental, la seguridad, la creación, el almacenamiento, la manipulación y los planes de acción en caso de desastre para documentos en todos los formatos, incluyendo los documentos digitales.

Preservación digital: La serie de actividades gestionadas necesarias para garantizar el acceso continuo a los materiales digitales por el tiempo que sea necesario. La preservación digital se define de manera muy amplia como todas las acciones necesarias para mantener el acceso a materiales digitales más allá de los límites de del soporte o de cambio tecnológico.

Proceso específico para mantener los documentos de archivo digitales durante y a través de las diferentes generaciones de la tecnología y a través del tiempo con independencia de donde residan.

Trazabilidad: creación, incorporación y conservación de información sobre el movimiento y uso de los documentos.

Usabilidad: Posibilidad de consulta de los documentos de archivo, determinada por la normativa vigente, su control archivístico y su estado de conservación. En ámbito informático: la capacidad o facilidad futura de la información de poder ser reproducida y por tanto utilizada. Véase además accesibilidad, manejabilidad.