

Información de control del documento

Título	<u>IMADARA: Recomendaciones para la digitalización del patrimonio documental</u>	Fecha
Versión	1.0.	4-07-2017

TABLA DE CONTENIDOS

<u>1.- INTRODUCCIÓN</u>	<u>5</u>
<u>2.- GLOSARIO</u>	<u>8</u>
<u>3.- REQUISITOS DE LA IMAGEN ELECTRÓNICA</u>	<u>12</u>
<u>3.1.- Formatos admitidos para imágenes electrónicas</u>	<u>12</u>
<u>3.1.1. - Ficheros maestros</u>	<u>12</u>
<u>3.1.2.- Ficheros de difusión</u>	<u>12</u>
<u>3.1.3.- Miniaturas para documentos gráficos y textuales:</u>	<u>13</u>
<u>3.2.- Nivel de resolución</u>	<u>13</u>
<u>3.3.- Garantía de imagen fiel</u>	<u>15</u>
<u>4.- PROCESO TÉCNICO DE DIGITALIZACIÓN DE DOCUMENTOS</u>	<u>17</u>
<u>4.1. Identificación y preparación de los originales.</u>	<u>17</u>
<u>4.1.2. Estado de conservación</u>	<u>18</u>
<u>4.1.3. Tratamientos previos</u>	<u>18</u>
<u>4.1.4. Ordenación y foliación</u>	<u>19</u>
<u>4.1.4.1. Documentos sueltos</u>	<u>19</u>
<u>4.1.4.2. Documentos unidos en origen</u>	<u>19</u>
<u>4.2. Proceso de digitalización</u>	<u>20</u>
<u>4.2.1. Pautas para la manipulación de originales por parte del operador</u>	<u>20</u>
<u>4.2.1.1. Para todo tipo de documentación:</u>	<u>20</u>
<u>4.2.1.2. Documentos unidos en origen</u>	<u>21</u>
<u>4.2.1.3. Pergaminos</u>	<u>21</u>
<u>4.2.1.4. Documentos de gran tamaño</u>	
<u>4.2.1.5. Documentos gráficos</u>	
<u>4.2.2 Captura</u>	<u>22</u>
<u>4.2.3. Ajuste de parámetros</u>	<u>24</u>
<u>4.2.4. Codificación de ficheros</u>	<u>26</u>
<u>4.2.5. Optimización de la imagen</u>	<u>28</u>
<u>4.2.6. Validación</u>	<u>28</u>
<u>4.2.7. Requisitos de entrega de los soportes</u>	
<u>5.- CONTROL DE CALIDAD</u>	<u>31</u>
<u>5.1 Principios básicos de un programa de control de calidad</u>	<u>31</u>
<u>5.1.1 Objetivo</u>	<u>31</u>
<u>5.1.2 Métodos</u>	<u>32</u>
<u>5.2 Control de calidad del escáner</u>	<u>32</u>
<u>5.3 Recomendaciones para el control de calidad</u>	<u>32</u>
<u>6.- REQUISITOS OPCIONALES</u>	<u>32</u>
<u>6.1 Metadatos</u>	<u>33</u>
<u>6.2 Alto OCR</u>	<u>33</u>
<u>7.- PROPIEDAD INTELECTUAL</u>	<u>35</u>

8.- MARCO NORMATIVO Y DOCUMENTOS DE INTERÉS 36

ANEXOS

ANEXO I: PLIEGO DE PRESCRIPCIONES TÉCNICAS PARA DIGITALIZACIÓN DE FONDOS DE ARCHIVO 38

ANEXO II: NIVEL DE RESOLUCIÓN 35

ANEXO III: CONTROL DE CALIDAD 37

ANEXO IV: HOJA DE OPERADOR 39

1.- INTRODUCCIÓN

El objeto de este documento es establecer un protocolo de actuación y unas pautas normalizadas para todos los Archivos del sistema DARA que sean fácilmente asumibles por todos los centros, cuenten o no con técnicos especializados en digitalización, unificando así criterios y facilitando la difusión y conservación de los documentos históricos.

Si bien DARA permite la incorporación de documentos de todo tipo, como imágenes en movimiento o documentos sonoros, las presentes recomendaciones se centrarán sólo en imágenes fijas de documentos textuales. Se incluyen algunas orientaciones para documentos gráficos (fotográficos, grabados, dibujos, cartografía, etc.).

Hay que tener en cuenta que estas recomendaciones van dirigidas a la digitalización de documentos de carácter histórico y/o declarados de conservación permanente, así como a documentos de carácter excepcional. Quedan excluidos los documentos que no tengan estas características.

Por lo tanto, a través de IMADARA se pretende:

- incrementar el acceso de los ciudadanos a los documentos de archivo mediante el uso de internet
- divulgar el patrimonio histórico
- salvaguardar los documentos originales evitando su manipulación
- crear copias de seguridad de materiales únicos y valiosos
- crear un repositorio común de difusión y preservación del patrimonio histórico (DARA)

Estas recomendaciones están destinadas a todos los Archivos aragoneses que trabajan en la digitalización de imágenes proporcionando unas pautas comunes de tratamiento, preservación y difusión de los documentos digitalizados. Por extensión, pueden ser aplicadas por cualquier centro que lleve a cabo tareas de digitalización de documentos.

Los destinatarios del contenido de IMADARA son:

- Responsables de la digitalización de documentos
- Responsables de gestión, conservación y archivo de documentos

2.- GLOSARIO

Para la elaboración de este glosario se han tenido en cuenta las definiciones incluidas en la *Guía de aplicación de la Norma Técnica de Interoperabilidad de Digitalización de Documentos (G-NTI-D)*. Asimismo, se han consultado las *Directrices para proyectos de digitalización de colecciones y fondos de dominio público, en particular para aquellos custodiados en bibliotecas y archivos (UNESCO)* de la UNESCO (versión actualizada de 2014) y las *Recomendaciones para la digitalización de documentos en los archivos (JCyL)* de la Junta de Castilla y León.

En el caso de tomar literalmente una definición de una de las fuentes mencionadas se consigna entre paréntesis la fuente de la que se ha extraído. Si no se especifica nada es el grupo IMADARA el responsable de la redacción o adaptación del texto.

Archivo: en el contexto de IMADARA, institución que conserva los documentos objeto de digitalización.

Captura: en el contexto de IMADARA, se entiende por captura el proceso de recolectar la imagen digital de los documentos originales que más tarde será evaluada y procesada.

Carpeta: en el contexto de IMADARA, se entiende por carpeta el contenedor virtual en el que se almacenan archivos informáticos y/o otros subdirectorios, atendiendo a su contenido, a su propósito o a cualquier criterio que decida el usuario.

Compresión (de datos): consiste en la reducción del volumen de información tratable (procesar, transmitir o grabar). En principio, con la compresión se pretende transportar la misma información, pero empleando una menor cantidad de espacio. No obstante, la reducción de datos puede afectar o no a la calidad de la información. Se habla de “compresión sin pérdida” cuando los datos antes y después de comprimirlos son exactos. Por el contrario, la compresión con pérdida supone que el algoritmo de compresión elimina determinados datos para reducir aún más el tamaño, con lo que se suele producir también una pérdida de calidad. (JCyL)

Conservación: conjunto de procesos y operaciones dedicados a asegurar la permanencia intelectual y técnica de los documentos a lo largo del tiempo. (G-NTI-D)

Conversión: proceso de transformación de un documento u otro objeto digital de un formato, o versión de formato, a otro. (G-NTI-D)

Copia: duplicado de un objeto, resultante de un proceso de reproducción. (G-NTI-D)

Dato: una representación de hechos, conceptos o instrucciones de un modo formalizado, y adecuado para su comunicación, interpretación o procesamiento por medios automáticos o humanos. (G-NTI-D)

Derecho de propiedad intelectual (de autor): conjunto de normas y principios que regulan los derechos morales y patrimoniales que la ley concede a los autores, por el solo hecho de la creación de una obra literaria, artística, científica o didáctica, esté publicada o inédita. Una obra pasa al dominio público cuando los derechos patrimoniales han expirado. (G-NTI-D)

Digitalización: proceso tecnológico que permite la obtención de uno o varios ficheros electrónicos que contienen la imagen codificada, fiel e íntegra de un documento *papel* a través de técnicas fotoeléctrica. (G-NTI-D)

Disponibilidad: propiedad o característica de un documento que permite que pueda ser consultado, localizado, recuperado presentado o interpretado.

Documento: información de cualquier naturaleza archivada en un soporte y susceptible de identificación y tratamiento diferenciado. (G-NTI-D)

Documento fotográfico: documento cuyo soporte es la fotografía en cualquiera de sus aspectos técnicos: negativo, positivo, diapositiva... sea cual sea el soporte y procedimiento técnico empleados. Se diferencia entre opacos y transparentes por su distinto comportamiento ante la digitalización.

Documento gráfico: en el ámbito de IMADARA, se entiende por documento gráfico aquel que contiene fundamentalmente imagen como mapas, planos, grabados,... sea cual sea la técnica y soporte empleados. No se incluyen los documentos fotográficos.

Documento textual: en el ámbito de IMADARA, se entiende por documento textual aquel que contiene fundamentalmente texto, sea manuscrito o impreso, independientemente del soporte empleado.

Dominio público: en el ámbito de los derechos de autor, se entiende por dominio público la situación en que quedan las obras literarias, artísticas o científicas al expirar el plazo de protección de los derechos patrimoniales exclusivos que las leyes de derecho de autor reconocen en favor del derechohabiente y que implica que pueden ser explotadas por cualquier persona o corporación, pero siempre respetando los derechos morales (básicamente la "paternidad" de la obra). Esto sucede habitualmente transcurrido un término contado desde la muerte del autor. (JCyL)

Especificación técnica: conjunto de requisitos que figuran en un documento en el que se definen las características requeridas de un producto, tales como los niveles de calidad, el uso específico, la seguridad o las dimensiones, incluidas las prescripciones aplicables al producto en lo referente a la denominación de venta, la terminología, los símbolos, los ensayos y métodos de ensayo, el envasado, el marcado y el etiquetado, así como los procedimientos de evaluación de la conformidad. (G-NTI-D)

Esquema de metadatos: instrumento que define la incorporación y gestión de los metadatos de contenido, contexto y estructura de los documentos electrónicos a lo largo de su ciclo de vida.

Estándar: especificación técnica aprobada por un organismo de normalización reconocido para una aplicación repetida o continuada cuyo cumplimiento no sea obligatorio y que esté incluida en una de las categorías siguientes:

I. Norma internacional: norma adoptada por una organización internacional de normalización y puesta a disposición del público.

II. Norma europea: norma adoptada por un organismo europeo de normalización y puesta a disposición del público.

III. Norma nacional: norma adoptada por un organismo nacional de normalización y puesta a disposición del público. (G-NTI-D)

IV. Norma autonómica: norma adoptada por un organismo autonómico de normalización y puesta a disposición del público. (IMADARA)

Fichero maestro (o "master"): primer fichero resultante de digitalizar el original.

Formato: conjunto de reglas que define la manera de almacenar o intercambiar datos en memoria.

Formato de difusión: el utilizado para la difusión por Internet y para proporcionar copias de calidad a quien las solicite.

Imagen electrónica: resultado de aplicar un proceso de digitalización a un documento. (G-NTI-D)

Integridad: referido a un documento, propiedad o característica que indica su carácter de completo, sin alteración de ningún aspecto esencial. La integridad es un componente de la autenticidad junto a la identidad. (G-NTI-D)

Interoperabilidad: capacidad de los sistemas de información, y por ende de los procedimientos a los que estos dan soporte, de compartir datos y posibilitar el intercambio de información y conocimiento entre ellos. (G-NTI-D)

Marca de agua: en términos digitales, una marca de agua visible o invisible que se crea mediante la inserción de un modelo de bits en una imagen digital para identificar que el propietario de la imagen es conocido.(UNESCO)

Metadatos: datos que definen y describen otros datos. Existen diferentes tipos de metadatos según su aplicación.

Original: referido a un documento, que posee la cualidad de genuino, que se remonta directamente a su autor y que no ha sido copiado ni imitado de otro y, por tanto, es eficaz (produce efectos).

Profundidad del bit: la profundidad del bit de una imagen se refiere al número de bits utilizado para describir el color de cada píxel. (UNESCO)

Rango dinámico: rango de diferencia tonal entre la parte más clara y la más oscura de una imagen.

Resolución: es la densidad de píxeles por pulgada utilizada en el muestreo de un documento en relación al área del fichero digital. Es la característica que determina la calidad de una imagen.

Soporte: objeto sobre el cual o en el cual es posible grabar y recuperar datos.

3.- REQUISITOS DE LA IMAGEN ELECTRÓNICA

La imagen electrónica obtenida a través de un proceso de digitalización, aunque no pueda ser completamente idéntica al documento en el soporte original, sí debería ser suficientemente fiel al contenido original y garantizar su integridad.

Existen diferentes formatos de ficheros para imágenes, cada uno con sus características específicas, sus ventajas e inconvenientes en relación a su capacidad para soportar diversas especificaciones técnicas. La utilización de un determinado formato conlleva en ocasiones la aplicación de determinadas técnicas de compresión. Además, la técnica seleccionada para la obtención de la imagen electrónica, ya sea con o sin compresión, podría provocar en mayor o menor grado algún tipo de pérdida en relación al original, como puede ser pérdida de nitidez, marcas de agua u otras características del soporte original.

Por norma general siempre se trabajará con formatos sin compresión para la generación del fichero maestro. En caso de que, por problemas de capacidad de almacenamiento del Archivo que custodia los originales, sea necesario optimizar el espacio de los ficheros resultantes se utilizarán sistemas de compresión sin pérdidas.

3.1.- Formatos admitidos para imágenes electrónicas

3.1.1.- Ficheros maestros

Ficheros maestros para documentos textuales:

- TIFF revisión 6.0 sin compresión.

- Espacio de color RGB a 24 bits (8 bits por canal).
- Perfil de color Adobe RGB (1998), ProPhoto RGB o eciRGBv2.
- En caso de necesitar aplicar técnicas de compresión porque así se solicite, sólo se podrá utilizar compresión Lossless LZW.

Ficheros maestros para documentos gráficos:

- TIFF revisión 6.0 sin compresión.

- Espacio de color RGB a 24 bits o escala de grises (8 bits por canal).
- Perfil de color Adobe RGB (1998), ProPhoto RGB o eciRGBv2 para color RGB o perfil basado en "gamma" o en "ganancia de punto" (respetando los valores por defecto del software usado) para escala de grises.
- En caso de necesitar aplicar técnicas de compresión porque así se solicite, sólo se podrá utilizar compresión Lossless LZW.

3.1.2.- Ficheros de difusión

La difusión de los contenidos supone permitir la accesibilidad a los contenidos digitalizados por medio de copias del master o bien descargando directamente a través de internet. Todos los contenidos digitalizados deberán facilitarse en formatos estándar y gratuitos que permitan la reproducción fiel y secuenciada del documento original y faciliten la lectura e impresión del documento.

Los archivos que se generen en PDF deberán intentar representar fielmente el original físico digitalizado.

La difusión de archivos en JPEG deberá garantizar una visualización de calidad de la imagen.

Ficheros de difusión para documentos textuales:

- JPEG

- Espacio de color RGB a 24 bits (8 bits por canal).
- Perfil de color igual al fichero maestro o sRGB.
- Compresión Lossy, submuestreo YUV 4:2:0 (2x2,1x1,1x1) y codificación de Huffman.
- Nivel de compresión máximo de un 15% respecto al permitido por el software usado (calidad mínima de un 85%).

- PDF

- Subconjunto PDF/A según normativa ISO 19005-1:2005 (PDF/A-1) o ISO 19005-2:2011 (PDF/A-2).
- Espacio de color RGB a 24 bits (8 bits por canal).
- Perfil de color RGB calibrado o basado en ICC.
- Optimización para internet (PDF linearizado).
- Se realizará OCR en capa oculta para los documentos con letra impresa.
- Compresión MRC (Mixed Raster Content) según normativa ISO 16485:2000.
- Nivel de compresión máximo de un 20% respecto al permitido por el software usado (calidad mínima de un 80%).
- Multipágina (un fichero por cada unidad documental).
- En caso de que el Archivo lo solicite, se añadirá una **marca de agua** en el margen inferior derecho de cada página, con un tamaño proporcional al documento para que no destaque y evitando tapar información o imágenes relevantes. Dicha marca de agua la proporcionará el archivo que custodia la documentación. Se respetarán los márgenes inferior y lateral derecho y tendrá una opacidad del 25% (nivel de transparencia al 75%).
- Se procurará que cada fichero PDF no supere los 50 MB de tamaño. En caso de ser necesario, se fraccionará el fichero PDF que supere dicho tamaño en las partes que se requieran, nombrando cada fichero según se especifica en el epígrafe 4.2.4. *Codificación de ficheros*.

Ficheros de difusión para documentos gráficos:

- JPEG

- Espacio de color RGB a 24 bits o escala de grises (8 bits por canal).
- Perfil de color igual al fichero maestro (tanto para RGB como para escala de grises) o sRGB.
- Compresión Lossy, submuestreo YUV 4:2:0 (2x2,1x1,1x1) y codificación de Huffman.
- Nivel de compresión máximo de un 15% respecto al permitido por el software usado (calidad mínima de un 85%).

3.1.3.- Miniaturas para documentos gráficos y textuales

Para los documentos textuales se realizará una miniatura de la primera página (o una página representativa) de cada libro o unidad documental. Para los documentos gráficos se realizará una miniatura para cada documento.

Siguiendo las indicaciones de Europea "La imagen o imágenes proporcionadas para las miniaturas puede ser cualquiera que elija que represente mejor al objeto digital. En el caso de un libro, la portada puede ser adecuada, o en el caso de que no figure el título en la portada, podría ser una fotografía de la página del título. Elija algo que sea interesante y fácil de encontrar para los usuarios".

- JPEG

- Espacio de color RGB a 24 bits o escala de grises (8 bits por canal).
- Perfil de color sRGB o igual al fichero maestro.
- Compresión Lossy, submuestreo YUV 4:2:0 (2x2,1x1,1x1) y codificación de Huffman.
- Nivel de compresión máximo de un 15% respecto al permitido por el software usado (calidad mínima de un 85%).
- Se recomienda que el tamaño de las miniaturas sea de 400 píxeles por el lado más largo.
- Como norma general, la miniatura mantendrá la proporción de aspecto del documento original.

3.2.- Nivel de resolución

Para documentos textuales y documentos gráficos que contengan texto relevante

Para decidir el nivel de resolución adecuado a cada imagen no sólo hay que tener en cuenta la tasa de muestreo del dispositivo de captura (ppi) sino también la escala de captura y el mínimo tamaño de letra del documento original.

Por ello, para decidir la resolución mínima necesaria para cada conjunto de documentos usaremos las fórmulas desarrolladas por el Departamento de Preservación y Conservación de la Biblioteca de la Universidad de Cornell para la digitalización de textos y manuscritos. Dichas fórmulas proporcionan la resolución de escaneado mínima necesaria para una correcta lectura en función de la calidad que deseemos y del tamaño mínimo del tipo de letra más significativo. Las fórmulas se basan en unos índices de calidad (QI) para las digitalizaciones en escala de grises y color RGB, donde un índice QI de 3,6 es baja legibilidad, un índice QI de 5.0 es una legibilidad media y un índice QI de 8.0 es una alta legibilidad. Dichos índices de calidad son un estándar desarrollado para la captura en microfilm que se refieren a la legibilidad del texto (ANSI/AIIM MS23-1991).

(Para una explicación más detallada sobre las fórmulas utilizadas consultar el Anexo II).

Se recomienda que cada Archivo fije una resolución mínima para la misma serie documental y, en caso de ser necesario, se corrija al alza según la siguiente tabla:

Tamaño de letra	Resolución aproximada
2 mm	150 ppi
1,5 mm	200 ppi
1 mm	300 ppi
0,75 mm	400 ppi
0,6 mm	500 ppi
0,5 mm	600 ppi

Para documentos gráficos (opacos y transparentes)

En el caso de los documentos gráficos, además de tener en cuenta la tasa de muestreo del dispositivo de captura (ppi) habrá que considerar el soporte del original. Al digitalizar un documento gráfico realmente estamos realizando una fotografía de una fotografía.

Para los documentos opacos el tipo de soporte en el que ha sido positivado (papel baritado, papel plástico, tela, soportes sólidos, etc...) o si tiene trama o no, puede delimitarnos el nivel de resolución máximo de escaneo ya que a partir de cierto nivel de resolución nos encontraremos con que aumentamos el tamaño del fichero resultante pero no su calidad efectiva.

Para los documentos gráficos transparentes, la sensibilidad de la película o del soporte en cristal puede hacer visible el "grano" de la emulsión. A partir de cierta resolución podemos capturar sus imperfecciones o un "grano" más grande (aumentado el tamaño del fichero final) sin capturar más calidad en la fotografía.

Como referencia consideraremos que una resolución adecuada para los soportes gráficos opacos será de entre 300 ppi y 600 ppi. En el caso de los soportes gráficos transparentes habrá que digitalizarlos con la resolución suficiente para mantener el nivel de nitidez conseguido en el soporte original. Para los soportes transparentes con sensibilidad alta será recomendable no superar los 2.400 ppi y para los soportes transparentes con sensibilidad baja podremos llegar hasta los 4.800 ppi (si se aprecia el grano de la emulsión a simple vista o tiene un tamaño considerable estaremos ante un soporte con sensibilidad alta).

En la siguiente tabla se exponen los niveles de resolución mínimos adecuados para cada soporte, teniendo en cuenta lo dicho anteriormente:

SOPORTE	CARACTERÍSTICAS	RESOLUCIÓN
Opaco	Hasta un tamaño de 18x24 cm	600 ppi
	A partir de 18x24 cm	300 ppi
Transparente de celuloide (Negativos y positivos)	35 mm y 16 mm	Desde 2.600 ppi hasta 4.200 ppi
Transparente de vidrio o celuloide	4x5,5 cm 4,5x6 cm	1.800 ppi
	6x6 cm 6x7 cm 6,5x9 cm	1.200 ppi
	9x12 cm 10x15 cm	800 ppi
	15x20 cm y tamaños superiores	600 ppi

Al seleccionar la tasa de muestreo adecuada en el dispositivo de captura no se podrá superar la resolución óptica nativa.

3.3.- Garantía de imagen fiel

Durante el proceso de digitalización se debe seguir en todo momento un plan de gestión de calidad que garantice que las imágenes producidas son fieles al documento en su soporte original.

Para ello, como mínimo, en cada proyecto de digitalización y siempre y cuando no cambien las condiciones o dispositivos de captura, se incluirá una imagen de muestra que incluya una carta de color para poder garantizar la idoneidad del balance tonal de la imagen. En caso de que se utilicen varios dispositivos de captura o se cambien las condiciones de digitalización, se incluirá una imagen con su correspondiente carta de color para cada dispositivo. Dichas imágenes estarán correctamente identificadas para poder relacionarlas con sus ficheros de captura correspondientes.

Opcionalmente, se podrá incluir también en la imagen de muestra una escala que se relacione con el tamaño del original a digitalizar. En caso de que el dispositivo de captura sea aéreo y permita distintas escalas, es obligatorio que la misma unidad documental esté reproducida a la misma escala, independientemente del tamaño de sus páginas.

Para garantizar la integridad de los ficheros generados en el proceso de digitalización, se generarán sumas de verificación o “checksum” para detectar cambios accidentales en la secuencia de datos. Dicho proceso se llevará a cabo siempre que se repliquen los mismos ficheros a distintos dispositivos de almacenamiento. Para realizar la comprobación de la integridad de los ficheros será válido cualquier algoritmo de los reconocidos como estándar actualmente por el mercado, si bien se recomienda el uso del algoritmo SHA-256.

El nombre del fichero que contendrá las sumas de verificación será “<ALGORITMO>SUMS” y contendrá el valor de todos los documentos de la carpeta en la que se encuentre.

Un ejemplo de validación, sería el fichero SHA256SUMS con contenido:

```
e091a875b29fdf1494509c406b36f5996bbbfb6fbac646dd13bae81675c28b51  
AHPHU_J_00002.djvu
```

```
144928e82cc82f12e242e2cb8da2cde2450b407bd77a8da01f328f5325a9548e  
AHPHU_J_00002_0001.djvu
```

```
d9a8e3c78fc167bfbfd808a07f0be0899887677f87721d433cb1a518c82ca9b46  
AHPHU_J_00002_0002.djvu
```

```
acb9d9043dca6a8ee26093edd78491f7bb7b2413dc7df2ca7f381b401a21fe86  
AHPHU_J_00002_0003.djvu
```

4.- PROCESO TÉCNICO DE DIGITALIZACIÓN DE DOCUMENTOS

4.1. Identificación y preparación de los originales.

La digitalización de originales de archivo tiene la finalidad de conservar y difundir la información que nos proporcionan los documentos ante la obligación de su conservación.

Sin embargo, cuando se aborda este tipo de proyectos no sólo es importante el contenido informativo, sino que se debe entender al documento como un *todo*, de forma que su estructura, soporte, formato y estado de conservación, nos lleven a reconocer el documento digitalizado como una copia casi exacta del original. Es por esto que el proceso se debe realizar con el máximo cuidado y respeto.

La digitalización, por lo tanto, no sólo hay que entenderla como una forma de difusión de los documentos, sino también como una forma de conservación y conocimiento de los originales.

Como complemento al trabajo de digitalización y para que el resultado final sea lo más exhaustivo y completo posible, se hace necesaria una preparación previa que después ayudará a que la mecánica de digitalización sea lo más rápida, eficiente y exhaustiva posible.

En esta labor previa habría que abordar diferentes puntos:

4.1.1. Tipo de documento

4.1.2. Estado de conservación

4.1.3. Tratamientos previos

4.1.4. Ordenación y foliación

4.1.1. Tipo de documento

Los documentos de archivo tienen una tipología muy variada y es necesario que esta circunstancia se tenga presente a la hora de abordar el trabajo de digitalización.

■ Libros encuadernados

En este tipo de documentos la estructura, los materiales y los posibles adornos exteriores son importantes para la historia del propio documento y para que los posibles estudiosos de este campo puedan investigar sobre el tema.

■ Pergaminos

Los documentos con este soporte presentan algunas dificultades (tamaño, sequedad, arrugas, etc.) que se irán abordando en los puntos siguientes. Normalmente se presentan escritos por una cara, pero en el reverso es frecuente que aparezcan fechas, notas o pequeños resúmenes del contenido.

Los pergaminos doblados sólo se desdoblarán si su manipulación lo permite, nunca se deben forzar, ya que existe riesgo de rotura.

■ Documentos de gran tamaño

En este apartado podemos encontrar desde libros de contabilidad a planos y mapas.

■ Documentos grapados o cosidos

Se deberá valorar en cada caso la necesidad de desgrapar o descoser el original para que la reproducción sea fiel y no se pierda información en la zona de la costura.

En el caso de documentos grapados, la eliminación de las grapas se realizará levantando por la parte trasera del documento, las dos partes de la grapa para después extraerla por la parte delantera.

En caso de documentos cosidos, este trabajo deberá realizarlo un restaurador/a ya que después habrá que volver a coser el original para mantener su estructura.

Sobre todo en estos casos es necesaria la foliación del documento antes del desmontaje para conservar el orden original.

4.1.2. Estado de conservación

Los documentos no siempre presentan un buen estado de conservación y es frecuente que, ante un proyecto de digitalización, esta circunstancia sea una de las que decidan los documentos a reproducir. Siendo esta una buena elección no está exenta de riesgos por la manipulación a la que van a ser sometidos documentos que ya presentan problemas, por eso habrá casos en los que será necesario posponer la digitalización hasta su tratamiento o incluso renunciar a ella.

Por lo tanto:

- Sería muy recomendable, incluso indispensable en muchos casos, hacer una valoración previa a la digitalización de los problemas concretos que se puedan presentar como consecuencia del estado de los documentos.

- Sería recomendable la buena comunicación entre el encargado de la digitalización y las personas responsables de la documentación para decidir protocolos de actuación en caso de que surjan dificultades.

- Ante cualquier duda, siempre primará la integridad del documento original.

Alteraciones y problemas que se pueden presentar:

Microorganismos

- El soporte papel degradado por microorganismos es muy frágil, llegando a perder su consistencia. Con documentación que presente este problema hay que actuar con mucha precaución. La manipulación, siempre que sea posible, debe ser muy cuidadosa ya que son frecuentes las roturas y pérdidas. Es habitual que las hojas con este problema estén pegadas entre sí, en este caso es necesaria la intervención del restaurador para separar los soportes y colocar hojas intermedias que permitan la manipulación en el momento de la digitalización.

- La manipulación de documentos con problemas de microorganismos se realizará con guantes, para evitar contagios entre documentos y una posible sensibilización del operador.

Tintas

- Los documentos con graves problemas de tintas metaloácidas no se manipularán, ya que existe un gran riesgo de pérdida de soporte. Si en el momento de la digitalización el deterioro ya está produciendo roturas y pérdidas, se debe evitar toda manipulación hasta que un profesional restaurador realice labores de consolidación en el documento. Además sería necesario colocar hojas intermedias que ayuden al digitalizador a manipular el documento sin que se produzcan más pérdidas ó degradación del original.

Pérdidas y roturas

- Los soportes que presenten zonas con roturas deberán tratarse con cuidado para no producir más daños en el original. Además sería aconsejable una consolidación de estas zonas.

4.1.3. Tratamientos previos

Una vez que el archivo ha decidido qué documentos se van a digitalizar y, teniendo en cuenta lo expuesto sobre el estado de conservación, en muchos casos son necesarias unas pequeñas intervenciones para que la digitalización sea exitosa en los documentos que no presentan buenas condiciones físicas.

Para la realización de estas tareas es recomendable la preparación de la documentación por parte de un restaurador. Por ejemplo:

- Unión de grietas, roturas y fragmentos separados del original.
- Consolidación del soporte y separación de hojas pegadas en caso de documentos afectados por microorganismos.
- Instalación de papeles intermedios que permitan la manipulación cuidadosa del original, cuando el soporte está muy deteriorado por microorganismos o tintas.
- En el caso de pergaminos que no se puedan desplegar, y necesitan un proceso completo de humectación y estirado.

4.1.4. Ordenación y foliación

4.1.4.1. Documentos sueltos

Antes de proceder a la digitalización, la documentación tiene que estar ordenada de acuerdo con los principios que establezca el Archivo. Un sistema habitual de controlar la ordenación es la foliación.

4.1.4.2. Documentos unidos en origen

4.1.4.2.a Foliación original

- Respetar al máximo la foliación existente.
- Los errores de foliación se mantienen. En la descripción aparecerá explicada esta circunstancia.
Foliación original 1-91, 86-230.
- Si al comienzo del libro/expediente hay hojas sin foliar, pero más adelante empieza la foliación original, la parte que no está foliada se comienza con: A, B, C...
- Si el original está foliado pero no llega hasta el final, seguir la foliación con el número que corresponda.
- Las hojas intermedias o cuadernillos que estén sin foliar, se foliarán de la siguiente manera:
 - 1 hoja con bis 7, 7bis, 8, 9...
 - Varias hojas con a,b,c... 59, 60, 61, 61a, 61b, 61c, 62, 63 ...
- Cuando en el original hay muchos cuadernillos sin foliar hay que valorar si se vuelve a foliar entero, aunque es mejor respetar el original.
- Si el original está paginado, seguiremos paginando.
- Si a mitad del libro hay un trozo paginado, lo anterior lo foliaremos, lo posterior lo seguiremos paginando.

Papeles sueltos

- Como norma se foliarán con el número que les corresponda.
- Si el documento ya está foliado, se le pondrá el número del folio anterior con un bis.
71, 71bis, 72, 73, 74, 75, 75, 77, 77bis, 78...
- En el caso de que haya más de un papel se añadirá a, b, c...
71, 71a, 71b, 71c, 72, 73, 74, 75, 76...

- Los documentos se conservarán en la posición en la que aparezcan.
- Si son más grandes que el original se mantendrá la imagen original.
- Papeles en las tapas:
 - Primero se intenta buscar relación con el original (si falta algún día, faltan fechas, partes de un documento, etc.), y en caso positivo el documento se inserta en su sitio correspondiente.
 - Si no encontramos relación con el original, se foliarán al final del documento.

Hojas en blanco

- Si existe, se respeta la foliación antigua. Se foliarán las páginas en blanco que estén opuestas a páginas con texto o algún tipo de contenido.

4.1.4.2.b Foliación nueva

- Se foliará con números arábigos consecutivos, incluyendo la portada y los índices.

Papeles sueltos

- Su ordenación se decidirá en cada caso.
- Los documentos sueltos u hojas de libro que por tamaño aparezcan doblados, se desdoblarán y se foliarán por el anverso en el sentido de la lectura.

Hojas en blanco

- No se folian salvo que estén opuestas a páginas con texto o algún tipo de contenido.
- Si se encuentran al final del documento, sólo se foliará la primera. Se contarán todas para que conste su número en la descripción.
- Si se encuentran al final del documento, pero la última o últimas hojas tienen cualquier texto, marca o sello, entonces se deben foliar todas, y en la descripción aparecerá el número de folios en blanco.
- Si hay folios en blanco entre dos grupos de folios escritos: se folia todo.
 - 1-4 , 5-50 (en blanco), 51-60
- Las hojas en blanco no se paginan.

4.2. Proceso de digitalización

4.2.1. Pautas para la manipulación de originales por parte del operador

La digitalización no puede poner en riesgo la integridad del documento. Al existir distinta tipología documental y diversos dispositivos de captura no es posible hacer un listado exhaustivo del protocolo de actuación. Por norma general el sentido común y la experiencia son las mejores guías.

4.2.1.1. Para todo tipo de documentación

- Nunca podrá prevalecer un criterio de producción sobre un criterio de conservación. A cada documento habrá que dedicarle el tiempo adecuado a su estado de conservación.
- Se comenzará por llevar a cabo una inspección visual general de la documentación para comprobar su estado de conservación, tipología, tamaño, etc. y de este modo escoger el mejor dispositivo de captura.
- El operador deberá tener en cuenta que la integridad del documento no es igual para todos los tipos de soporte. Se realizará el manipulado con cuidado y especialmente en el caso de que pueda suponer un riesgo para su integridad en función de su estado de conservación.

- En el caso de que la documentación presente hongos o bacterias se tendrá un especial cuidado en limpiar la superficie de captura que esté en contacto con el original para evitar infectar otros originales que se capturen posteriormente.
- En el caso de que la documentación presente roturas o degradación en su soporte se tendrá en cuenta su estado de conservación para decidir si es recomendable proceder con su digitalización o no. En cualquier caso, no se procederá con la digitalización si no ha habido una preparación previa de la documentación y se realizará, preferentemente, con un dispositivo cenital.
- Si durante el proceso de captura el operario se encontrara insectos en la documentación se comunicará inmediatamente al Archivo para que actúe según considere oportuno.
- No se debe utilizar ningún elemento punzante, metálico o de cualquier otro tipo, que pueda producir daños en los documentos originales.
- Es recomendable que el operario disponga de una mesa auxiliar para poder tener siempre controlada la documentación y evitar desorden y posibles caídas en su manejo.
- Al finalizar su jornada de trabajo el operario no dejará ningún documento dentro del dispositivo de captura si no que se guardarán en los lugares habilitados para tal fin.
- En caso de necesitar dejar un testigo en la documentación como referencia, se utilizará papel libre de ácido o papel barrera.

4.2.1.2. Documentos unidos en origen

En este tipo de documentos, la estructura, los materiales y los posibles adornos exteriores son importantes para la historia del propio documento y para que los posibles estudiosos de este campo puedan investigar sobre el tema. Con estas premisas se recomienda:

- En caso de usar dispositivos de captura cenital con mesa prensa libros se seleccionará para cada documento la presión adecuada, variándola todas las veces que sea necesario; así mismo la distancia entre los lados de la balanza del soporte se adecuará al grosor de cada documento. El objetivo es que el lomo del volumen no sufra demasiada presión ni rozaduras al capturarlo.
- En caso de usar dispositivos de captura cenital sin mesa prensa libros o escáneres, se asegurará que la superficie a digitalizar se encuentre a la misma distancia del sensor de captura en toda su superficie. En cualquier caso, es primordial que ninguna parte del documento quede fuera de la profundidad de campo que permita el dispositivo.
- Digitalizar las tapas, guardas y, en su caso, el lomo del documento.
- No forzar nunca la encuadernación.
- En los documentos gruesos, no se forzará su apertura.
- No utilizar prensa libros en caso de hojas con sello de placa.

4.2.1.3. Pergaminos

- Siempre se digitalizará anverso y reverso cuando ambos contengan información. Si el reverso no tiene ninguna información la decisión de su reproducción quedará a criterio del Archivo.
- En caso de existir sellos pendientes completos, restos de sellos o sólo los cordones de los mismos, estos elementos formarán parte de la imagen del documento.
- Se recomienda además, realizar también imágenes de los sellos, anverso y reverso, en las que se puedan apreciar bien sus características.

4.2.1.4. Documentos de gran tamaño

- Se extremará la precaución en el despliegue del original para no producir roturas o arrugas.

- Para los documentos grandes que estén doblados se evitará, en la medida de lo posible, desplegar el documento al completo. Es preferible que se trabaje por partes.
- Es necesario tener elementos de peso suave para sujetar el documento en caso necesario, como pequeñas pesas con protección, tiras de Mylar etc. y mesas auxiliares para un buen apoyo, evitando en todo momento que el documento no tenga soporte.
- Se conservará el documento en su posición, tal y como aparece, después se desplegará para que el operador pueda trabajar; se preparará la documentación para que, en caso de ser necesario, su secuencia de trabajo sea:

1) doblado por el anverso

2) desplegado

3) doblado, por el reverso

4.2.1.5. Documentos gráficos

- En una primera inspección previa a la digitalización se separará la documentación según su tipología (opacos/transparentes, papel/celuloide/vidrio) para seleccionar el dispositivo de captura adecuado en función de sus condiciones.
- El operador deberá procurar mantener limpia la superficie de captura y los originales, especialmente para los soportes transparentes. Para ello se recomienda el uso de guantes de algodón.
- No se permitirá la utilización de spray anti-newton.

4.2.2. Captura

El objetivo principal de un proyecto de digitalización es la conservación y la difusión. Ambos puntos se respetarán escrupulosamente a la hora de digitalizar los originales. La conservación porque una reproducción digital de calidad puede sustituir en la mayoría de los casos a la consulta del original. Y la difusión porque dicha reproducción puede ser consultada por medios telemáticos desde cualquier lugar del mundo.

Para mantener un equilibrio entre ambos criterios se partirá de la premisa de que la calidad de la captura será la mejor posible siempre que no haya riesgo para la integridad de la documentación. En caso de duda, se primará la integridad del original antes que la calidad de la captura.

Se recomienda usar el dispositivo de captura más adecuado para cada tipo de documentación; por ello, ante grandes proyectos de digitalización, sería recomendable agrupar la documentación por tipología, tamaño, grosor... para poder dar el tratamiento adecuado a cada grupo. Recomendamos, por orden de preferencia, el uso de:

- Dispositivos cenitales (cámaras planetarias, escáneres aéreos o cámaras fotográficas profesionales) combinados con una mesa prensa libros que permita la selección de presión y distancia de la balanza.
- Dispositivos cenitales (cámaras planetarias, escáneres aéreos o cámaras fotográficas profesionales) combinados con una balanza para el documento o una mesa de reproducción.
- Escáneres planos de sobremesa de tamaño adecuado a la documentación.
- Escáneres planos de gran formato.
- Escáneres de gran formato (de rodillo) para documentación en buen estado de conservación. Para soportes delicados y valiosos se deberá proteger el original con tereftalato de polietileno (PET) apto para conservación de documentos.
- Escáneres de negativos.
- Cámara fotográfica profesional para documentos de gran tamaño, realizando las capturas por secciones cuadrículadas y uniéndolas posteriormente por software.

Para fondos delicados o valiosos no se recomienda el uso de escáneres documentales, multifunción o con alimentador (ADF).

De este modo, según el tipo de documento, recomendamos, aunque no exclusivamente, el uso de unos dispositivos preferentemente sobre otros:

	Dispositivos idóneos	Dispositivos permitidos	
Documentos unidos en origen (encuadernados o cosidos)	Dispositivos cenitales + Prensa libros	Dispositivos cenitales + Mesa de reproducción	Escáner plano (de tamaño adecuado a la documentación)
Documentos y hojas sueltas	Escáner plano (de tamaño adecuado a la documentación)	Dispositivos cenitales (con o sin prensa libros)	
Pergaminos	Dispositivos cenitales (con o sin prensa libros)	Escáner plano (de tamaño adecuado a la documentación)	
Documentos de gran tamaño	Escáner plano (de tamaño adecuado a la documentación)	Dispositivos cenitales (con o sin prensa libros)	Escáner de gran formato (rodillos)
Documentos gráficos opacos	Escáner plano (de tamaño adecuado a la documentación)	Dispositivos cenitales + Mesa de reproducción	
Documentos gráficos transparentes	Escáner de negativos (de tamaño adecuado a la documentación)	Escáner plano (mediante adaptadores)	Dispositivos cenitales + negatoscopio

4.2.3. Ajuste de parámetros

Hemos de tener en cuenta que cada fabricante puede usar distinta terminología en el software que acompaña al dispositivo de captura. Para dispositivos domésticos o con un software básico no encontraremos muchas opciones de configuración, sin embargo, los ajustes de parámetros recomendados son ampliamente aceptados como estándares, por lo que aunque no podamos seleccionar alguna opción tenemos una alta probabilidad de que el fichero resultante sí vaya a cumplir con dichos criterios:

- **Profundidad de color:** Se seleccionará "True color", "Color verdadero", color 24 bits o 8 bits por canal (todas ellas denominaciones para el mismo concepto) para color RGB o escala de grises para blanco y negro.
- **Perfil de color:** Se seleccionará Adobe RGB (1998), ProPhoto RGB o eciRGBv2 para color RGB o Dot Gain o Grey Gamma para escala de grises. Si no se puede seleccionar alguno de los anteriores, se respetará el propietario incluido por el fabricante.
- **Temperatura de color:** Para escáneres y dispositivos con iluminación incorporada no podremos variar la temperatura de color; para dispositivos con luz externa, procuraremos que tenga entre 4.500K y 6.500K, prefiriendo el punto intermedio de 5.200K.
- **Resolución:** Seleccionaremos una tasa de muestreo expresada en ppi según se indica en el punto 3.2. *Nivel de resolución*.

- **Formato y compresión:** Seleccionaremos el formato siguiendo las indicaciones del punto 3.1. *Formatos admitidos para imágenes electrónicas.*

Para cumplir con unos mínimos criterios de calidad y homogeneidad se tendrán en cuenta los siguientes puntos:

- Se eliminarán todas las variables posibles durante la captura de tal manera que lo único que cambie sea el documento.
- Se calibrará cada dispositivo de captura. Siempre que cambie la iluminación, se varíe de software de captura o de equipo informático se volverá a calibrar.
- Se evitará la entrada de luz parásita en el dispositivo de captura.
- Si la iluminación es externa, será constante y uniforme.
- Para dispositivos cenitales no podrá haber variaciones de distancia en el plano de captura.

Como norma general, se seguirá el siguiente procedimiento de trabajo:

- Se capturará una imagen que contenga una carta de color por cada dispositivo o siempre que haya variaciones en la iluminación.
- Se seleccionará el nombre de la carpeta /ficheros del documento a digitalizar según lo indicado en el epígrafe 4.2.4. *Codificación de ficheros.*
- Si la documentación a digitalizar está foliada/paginada, la secuencia de captura será la del orden de foliación/paginación; si no es el caso, la secuencia de captura será la del orden que presenten los documentos según los proporcione el Archivo.
- En el caso de documentación unida en origen se recomienda una captura a doble página; así respetaremos no sólo el contenido del documento sino también la estructura de su composición. Se recomienda comenzar la digitalización por las tapas e incluir las guardas. Queda a criterio del Archivo la digitalización de los lomos y, en caso de ser necesario, los cortes del libro.
- En el caso de la documentación unida en origen se digitalizarán las páginas en blanco que estén opuestas a páginas con texto o algún tipo de contenido. En caso de aparecer algún trozo de papel unido a la encuadernación o alguna página rota que no esté considerada como un folio a la hora de preparar la documentación pero que impida la correcta captura de la totalidad del texto de algunas de las páginas entre las que se encuentra, se capturará como si realmente fuera un folio.
- En el caso de folios sueltos, legajos, fichas, pergaminos o cualquier documentación no unida en origen, se realizará la captura secuencialmente de cada folio en una imagen (anverso/reverso). Se respetará la estructura documental de tal manera que todos los folios del mismo documento tendrán la misma nomenclatura, cambiando sólo el número de página. En este caso no se digitalizarán las páginas que no presenten ningún tipo de contenido (texto, filigranas, dibujos, etc.) tanto en su anverso como reverso. Si se digitaliza el anverso de un folio, es recomendable digitalizar su reverso aunque éste aparezca en blanco, siendo potestad del Archivo variar dicho criterio.
- Para pergaminos o documentos que presenten sellos, restos de sellos, cordones o cintas, estos se procurarán incluir dentro de la captura siempre y cuando se pueda respetar el nivel de resolución exigido en este documento. Si no se puede respetar el nivel de resolución, se permitirá digitalizar el anverso/reverso en dos imágenes aparte.
- En dispositivos cenitales de captura se recomienda dejar un margen de seguridad respecto al documento al realizar el encuadre.
- En dispositivos de captura que incluyan mejoras en el proceso de digitalización, se permitirá utilizar las que estén implementadas por hardware (por ejemplo el sistema ICE -Image Correction & Enhancement-) pero no se recomienda el uso de aquellas que se realicen por software (filtros de reducción de grano, interpretaciones de niveles y curvas, etc...).
- La captura de la imagen estará horizontal respecto al sentido de visualización; si es necesario capturar la imagen con alguna desviación por cuestiones técnicas, posteriormente se enderezará dicha imagen por software hasta conseguir su horizontalidad correcta. En el caso de

los documentos textuales encuadernados o cosidos que no se puedan abrir simétricamente, se buscará que la unión de ambas páginas esté correctamente vertical.

- Para documentos gráficos se procurará respetar el rango dinámico del original, especialmente en los originales transparentes.
- Para los documentos gráficos en blanco y negro se preferirá realizar el escaneo en color RGB y después de optimizar la imagen convertirla a escala de grises.
- En los documentos gráficos en negativo se procederá a invertirlos para que el fichero resultante muestre su positivo correspondiente.
- En los documentos gráficos se procurará respetar los niveles de exposición de los originales aunque estos estén claramente sobreexpuestos/subexpuestos. Sólo si el Archivo así lo solicita se ajustarán los niveles para intentar recuperar la mayor información del artefacto fotográfico.
- En las imágenes finales se deberá eliminar el margen de seguridad dejando tan solo un marco (preferiblemente de tonalidad oscura). Se considerará que el marco tiene un tamaño aceptable siempre que sea mayor o igual a 5 píxeles y menor o igual a 60 píxeles.
- Se deberá corregir la inclinación de la imagen buscando la horizontalidad correcta. Se procurará que las páginas no tengan inclinación hacia ningún lado, pero se considerará aceptable una desviación que no sea mayor de 0,5° hacia cualquiera de los lados. En el caso de documentación a la que posteriormente se le aplicará un proceso de OCR, las líneas de texto deberán tener la menor desviación posible (aunque suponga una ligera desviación en la página).
- Se consignará cualquier incidencia que aparezca en una “hoja de operador” (*Anexo IV*). Se procurará registrar al menos la documentación que presente hongos o bacterias, la que esté en mal estado de conservación y en general toda aquella que presente particularidades dignas de mención. Dicha hoja estará a disposición del Archivo que custodia de la documentación si así lo desea.
- Se pueden añadir metadatos adicionales con posterioridad a la captura, pero es imprescindible que los ficheros resultantes respeten como mínimo:
 - Identificación del dispositivo de captura
 - Empresa o entidad digitalizadora
 - Fecha de creación
 - Medidas en píxeles
 - Espacio de color
 - Perfil de color
 - Profundidad de color
- Se vigilará a la hora de elegir el software de edición que no elimine los metadatos incrustados en la imagen.
- Si el software de captura lo permite se añadirá el copyright.

En general se aplicarán todas aquellas operaciones de mantenimiento preventivo y comprobaciones rutinarias que permiten garantizar mediante su cumplimiento que, en todo momento, el estado de la aplicación de digitalización y los dispositivos asociados producen imágenes fieles e íntegras tal y como se ha definido. Estas operaciones incluyen el mantenimiento, limpieza y actualizaciones de los dispositivos de captura, así como la comprobación periódica de que el software utilizado cumple con los estándares en los algoritmos de generación de imágenes. Se deberá también comprobar la vigencia de la presente normativa y sus posibles ampliaciones posteriores.

4.2.4. Codificación de ficheros

Las imágenes se ofrecerán con una estructura lógica en carpetas. Los ficheros se agruparán en carpetas denominadas con la signatura del documento digitalizado según la relación de signaturas que facilitará el Archivo.

En este apartado se describe la nomenclatura de las imágenes digitalizadas. Los campos descriptivos se separarán por “_” (guión bajo).

Todo fichero deberá indicar en primer lugar el Archivo al que pertenece el documento según las siglas que utiliza DARA para identificarlos. Por ejemplo:

AHPHU = Archivo Histórico Provincial de Huesca

AHPZ = Archivo Histórico Provincial de Zaragoza

AHPTE = Archivo Histórico Provincial de Teruel

AMZ = Archivo Municipal de Zaragoza

Seguirá, si procede, la sección dentro de dicho archivo (etiqueta MARC 090\$g).

A continuación, se indicará el número de contenedor (caja, legajo o equivalente) dentro de la sección (etiqueta MARC 090\$f). Como norma general dicho indicativo se cuadrará a cinco cifras, completando con ceros por la izquierda cuando sea necesario. Cada Archivo podrá utilizar el número de cifras necesario garantizando la homogeneidad de las signaturas.

En caso de que dentro del contenedor haya un número de orden (expediente o equivalente) (etiqueta MARC 090\$c), figurará tras el número de contenedor, cuadrado a cuatro cifras.

Para las páginas del documento, cada imagen respetará la secuencia numérica cuadrada a cuatro cifras. En el caso de los pergaminos se añadirá al final, en lugar de la numeración, si es el anverso (A) o el reverso (R). En el caso de sellos hay que identificar también anverso (SA) y reverso (SR); si hay varios sellos se indicará el número de orden (de izquierda a derecha): S1A, S1R, S2A, S2R.

Habrá que respetar para todos los ficheros su extensión correspondiente. Hay que tener en cuenta que la denominación es unívoca.

Por ejemplo, la primera imagen del documento 2 del contenedor 12 de la sección J del Archivo Histórico Provincial de Huesca se denominará:

AHPHU_J_00012_0002_0001

La entidad es AHPHU, la sección de archivo es “J”, el contenedor es “00012”, el orden dentro del contenedor es “0002” y es la primera imagen capturada de ese documento “0001”.

Tendremos en cuenta que puede haber situaciones en las que no sean necesarios todos los campos. De la misma forma, si el Archivo tiene más subdivisiones en los contenedores, también se contemplará en la denominación de las imágenes digitales. En caso de duda con la denominación de los ficheros, se consultará con el Archivo.

Todos los ficheros pertenecientes al mismo documento estarán contenidos dentro de la misma carpeta que se denominará según el mismo criterio (sin tener en cuenta la secuencia de captura).

En el caso de los ficheros de consulta PDF que necesiten ser fraccionados para no superar el tamaño recomendado en el epígrafe 3.1.3. *Ficheros de consulta para documentos textuales*, cada fichero se denominará según lo indicado anteriormente, anexionando como sufijo un guión medio “-” y el número de fragmento cuadrado a 2 dígitos:

AHPHU_J_00012_0002-01.pdf

AHPHU_J_00012_0002-02.pdf

AHPHU_J_00012_0002-03.pdf

...

Los archivos que contengan las imágenes en miniatura se denominarán igual que el archivo al que corresponden, anteponiendo el sufijo “min”.

Todas las carpetas que contengan documentación con distintas finalidades de uso (maestro, copia de consulta), se ordenarán bajo una carpeta cuya denominación especificará dicho formato.

Ejemplo:

MAESTRO

AHPHU_J_00012_0002

AHPHU_J_00012_0002_0001.tif

AHPHU_J_00012_0002_0002.tif

...

CONSULTA

AHPHU_J_00012_0002

minAHPHU_J_00012_0002_0001.jpg

AHPHU_J_00012_0002_0001.jpg

minAHPHU_J_00012_0002_0001.jpg

AHPHU_J_00012_0002_0002.jpg

Por último indicar que, con el objetivo de evitar incompatibilidades entre diferentes sistemas, tanto el nombre de las carpetas empleadas como el nombre resultante de los archivos de imagen, contendrán exclusivamente caracteres alfabéticos (a-z y A-Z), dígitos (0-9) o símbolos “_” y “-”.

Se evitará el uso de caracteres acentuados (incluida la letra Ñ) y de cualquier otro símbolo que distinto a los mencionados “_” o “-”.

4.2.5. Optimización de la imagen

El fichero maestro deberá estar lo más ajustado posible a sus parámetros ideales (temperatura de color, brillo, contraste, curvas tonales...) desde el origen de la captura. Los dispositivos de captura deberán estar correctamente calibrados y todo el proceso de digitalización deberá estar normalizado. El objetivo es conseguir una reproducción digital “forense” del original; no debe ser “embellecer” el original o hacer una interpretación artística.

Por ello en el fichero maestro solo se permitirá realizar ajustes del tipo:

- Correcciones geométricas.
- Recortar el lienzo de la imagen, en caso de ser necesario, según se indica en el epígrafe 4.2.3. *Ajuste de parámetros*.
- Enderezar el lienzo de la imagen, en caso de ser necesario, según se indica en el epígrafe 4.2.3. *Ajuste de parámetros*.

Se evitará realizar cualquier ajuste del tipo:

- Restauraciones digitales.
- Eliminación de imperfecciones que aparezcan en el original.
- Manipulaciones cuya intención sea corregir defectos en el proceso de digitalización.
- Interpolaciones digitales.
- Uso de las "máscaras de enfoque".

Se recomienda, sin embargo, que para la generación de los ficheros de difusión se apliquen mejoras que puedan ayudar a interpretar el documento original, en especial los documentos textuales que puedan presentar dificultad en su lectura. Más concretamente, se recomienda si es necesario:

- Ajustes de brillo / contraste que faciliten la lectura en documentos textuales.

- Restauración del enfoque para dispositivos de captura que incorporen un filtro de paso bajo, mediante "máscaras de enfoque".

4.2.6. Validación

Tras la captura de los documentos se deberá verificar que:

- Las imágenes digitales tienen la resolución adecuada o el índice QI indicado para su correcta visibilidad/lectura.
- Las imágenes digitales tienen un enfoque suficiente para poder apreciar detalles según su resolución de captura, sin abusar de filtros de corrección.
- La iluminación es uniforme en todas las partes del documento y el color de la imagen digital representa lo más fielmente posible al color del original.
- El conjunto de imágenes pertenecientes al mismo documento sea una representación fiel e íntegra de la unidad documental.
- Las imágenes digitales estén correctamente alineadas.
- Las imágenes digitales no tengan márgenes añadidos, más que el indicado en IMADARA.
- En el caso de que no se cumplan estas premisas se deberán subsanar dichos defectos. Si la carencia de calidad pertenece al proceso de digitalización, se deberán corregir los puntos erróneos ya que afectarán a toda la producción de imágenes.

4.2.7. Requisitos de entrega de los soportes

Como mínimo la empresa o servicio de digitalización generará y entregará a la entidad responsable del proyecto una copia de las obras digitalizadas en formato TIFF 6.0/ISO 12639:2004 y otra en formato JPEG/ISO/IEC 10918:1994, como formatos de máxima calidad y formato comprimido y optimizado para su lectura en internet, respectivamente.

El soporte físico empleado para la entrega serán discos duros USB 3.0 (no se recomienda el uso de "llaves USB"), los cuales contendrán los archivos de imagen así como los metadatos asociados a las mismas. Se recomienda la entrega por duplicado de los soportes físicos empleados.

Una vez se hayan recibido los soportes físicos, se recomienda realizar una copia completa de sus contenidos a los sistemas de almacenamiento corporativos.

5.- CONTROL DE CALIDAD

El Archivo responsable del proyecto se reservará el derecho de inspección y revisión del desarrollo y ejecución de los servicios objeto de digitalización.

Durante la ejecución de los trabajos de digitalización, la empresa o servicio de digitalización se comprometerá, en todo momento, a facilitar a las personas designadas por el Archivo que gestione el proyecto, la información y documentación que éstas soliciten para disponer de un pleno conocimiento de las circunstancias en que se desarrollan los trabajos, así como de los eventuales problemas que puedan plantearse y de las tecnologías, métodos y herramientas utilizados para resolverlos. En particular deberán comunicarse las incidencias que puedan producirse a lo largo del proceso de digitalización, tales como falta de texto, páginas deterioradas, ilegibilidad del original, errores de numeración, etc., de manera que el Archivo responsable del proyecto de digitalización pueda indicar a la empresa o servicio de digitalización los procedimientos que se deben llevar a cabo para gestionar estas incidencias. *(Se podrá utilizar como ayuda la plantilla del Anexo IV).*

Tanto por parte de la empresa digitalizadora, como por parte del Archivo responsable del proyecto, ha de realizarse un seguimiento del trabajo durante la ejecución del mismo. La empresa o servicio de digitalización habrá de subsanar lo más rápidamente posible los errores detectados con este control de calidad.

La empresa o servicio de digitalización deberá garantizar el correcto funcionamiento de los trabajos realizados durante el plazo de un año a contar desde la fecha de finalización del servicio, obligándose a realizar gratuitamente durante dicho tiempo las correcciones y modificaciones necesarias para subsanar los errores que eventualmente pudieran descubrirse.

5.1. Principios básicos de un programa de control de calidad

El control de calidad es un elemento importante en cada una de las etapas de un proyecto de digitalización. Sin este trabajo no será posible garantizar la integridad y consistencia de los ficheros de imágenes.

Deben tomarse medidas para minimizar las variaciones entre los diferentes operadores. Es necesario un programa de control de la calidad tanto para proyectos propios como para aquellos que se contratan externamente.

Una diferencia importante es que en un proyecto contratado externamente de forma total o parcial los requisitos de calidad a menudo tienen que formularse antes de la firma del contrato, debido a su naturaleza legal vinculante. En los proyectos propios se pueden ir creando poco a poco los programas de control de calidad como parte de las tareas del proyecto.

Aunque el control de calidad es un factor crucial para asegurar los mejores resultados no existe un modo normalizado para asegurar una determinada calidad de la imagen durante su captura. Los diferentes documentos originales requieren diferentes procesos de escaneo, lo que debe tenerse en cuenta cuando se desarrollan programas de control de calidad.

5.1.1. Objetivo

Una cuestión importante en un programa de control de calidad es si incluirá ¿la colección completa de imágenes o una muestra? ¿todo tipo de ficheros (ficheros de conservación, ficheros de acceso, ficheros de miniaturas)?

La respuesta depende del objetivo del proyecto de digitalización, de los productos requeridos y de los niveles de calidad y puntos de referencia elegidos. Si el programa de digitalización es muy

limitado o los requisitos de calidad son extraordinariamente altos, tendrá más sentido examinar la colección completa imagen por imagen. Sin embargo, en la mayoría de los programas es suficiente con establecer un plan de muestreo que cubra por ejemplo el 10% de las imágenes producidas por cada dispositivo de escáner durante un determinado periodo de tiempo (un día, una semana, un mes). Si un porcentaje previamente especificado de las imágenes elegidas es incorrecto, entonces deberá revisarse el grupo completo de imágenes.

Un programa de control de calidad siempre incluye los ficheros de conservación que se producen y en la mayoría de los casos también tendrá en cuenta otros productos como los ficheros de acceso.

5.1.2. Métodos

Las herramientas automáticas de evaluación de la imagen de las que disponemos actualmente no son, por lo general, suficientes para materiales que se requieren para fines culturales y científicos. Por lo tanto, la evaluación de la calidad visual debe llevarse a cabo de forma manual.

Deben tenerse en cuenta las limitaciones técnicas que pueden afectar a la evaluación. Los métodos recomendados son para la evaluación en pantalla:

- visualización de las imágenes escaneadas a 1:1 (ampliados al 100%)
- uso de objetivos para evaluar la reproducción en escala de grises y color
- uso de objetivos de resolución e histogramas para evaluar la resolución espacial y la reproducción tonal.
- utilización de métodos para la medida del ruido y herramientas de detección de objetos.

5.2. Control de calidad del escáner

Cuando un proyecto de digitalización está en marcha, deben establecerse medidas de control de calidad del escaneado que permitan a los operadores estar seguros de que los dispositivos de escaneo están operando dentro de los parámetros previstos. Las cuestiones que más conciernen a los resultados son: resolución espacial, reproducción tonal, reproducción del color, ruido y detección de objetos. En los proyectos en los que se digitalizan materiales de gran formato, como mapas y planos, la precisión geométrica es también un factor importante.

5.3. Recomendaciones para el control de calidad

Las condiciones para la calidad de la imagen digital se pueden describir en tres etapas:

- Identificación de los productos finales deseados y de los propios fines de la digitalización. Decida qué es lo que quiere producir y cuándo.
- Establecimiento de normas. Defina niveles aceptables de calidad de las imágenes digitales basados tanto en las características de los documentos originales como en las prestaciones del sistema de digitalización de imágenes que se vaya a utilizar
- Decisión sobre los puntos de referencia. Decida con qué se van a comparar los productos de salida del proceso de digitalización

Dependiendo de estas decisiones, hay que seleccionar lo que deberían incluir los programas de control de calidad. *(Para una explicación más detallada consultar el Anexo III).*

6.- REQUISITOS OPCIONALES

6.1. Metadatos

Será decisión del Archivo valorar si se solicita o no la generación de metadatos durante los procesos de digitalización. En el caso de hacerlo se recomienda seguir los estándares que a continuación se indican.

Este apartado “Metadatos” está basado en el documento *Requisitos técnicos de los proyectos de digitalización de patrimonio bibliográfico y de prensa histórica de la SGCB* publicado por el Ministerio de Educación, Cultura y Deporte Subdirección General de Coordinación Bibliotecaria. En la dirección <http://hdl.handle.net/10421/8981> es posible encontrar el documento “Anexo I-Ejemplos-Mets-SGCB.pdf” que contiene documentos ejemplo que siguen las recomendaciones aquí indicadas.

DARA emplea EDARA (norma basada en MARC21) para la descripción de sus contenidos de archivo. En futuras versiones de DARA se emplearán archivos METS para describir y/o carga de los objetos digitales.

En principio, y salvo que se acordara previamente otra cosa, el Archivo responsable del proyecto de digitalización es el encargado de proporcionar a la empresa o servicio de digitalización los registros EDARA y la información de fondos de las obras que se van a digitalizar. La empresa o servicio de digitalización utilizará estos registros para embeberlos en los registros METS, teniendo buen cuidado en hacer coincidir los números de control y firmas de los registros EDARA y de los METS.

En cuanto a los registros METS, salvo que se acuerde previamente otra cosa, será la empresa o servicio de digitalización la encargada de generar y entregar para cada monografía (obra impresa, manuscrito, material fotográfico, partituras, mapas, etc.) un archivo METS. Estos METS deberán ceñirse al Perfil METS “Spanish Virtual Library of Bibliographical Heritage - Digital Resources Ingest and Preservation BVPB-METS profile”

En particular los METS deben respetar las siguientes especificaciones:

- Los METS deben incluir direcciones correctas a los namespaces y a los schemas utilizados, y deben estar bien formados y ser válidos.
- Deben incluir una sección de metadatos descriptivos (`dmdSec`) que embeba el registro bibliográfico en formato MARC21XML y, a continuación, el registro de fondos en formato MARC21XML de fondos.
- Deben incluir una sección de metadatos descriptivos (`dmdSec`) en la que se relacione la imagen que va servir como icono representativo de la obra (normalmente una portada o ilustración).
- Los METS deben incluir una sección de archivo (`fileSec`), que en un único `fileGrp` agrupe elementos `file` para cada imagen de las que conforman la obra. Además, en etiquetas `FLocat` para cada elemento `file` se incluirán las direcciones a las imágenes correspondientes.
- Las direcciones de los METS apuntarán a las imágenes de difusión que se encuentran en los mismos discos duros que estos METS.
- Se incluirá una sección de mapa estructural (`strucMap`) con atributo `TYPE="LOGICAL"` que deberá incluir la información de paginación, cuando el proceso de digitalización genere múltiples imágenes para un único objeto físico.

- Incluirán los metadatos que se precisan para la preservación de la copia digital. Estos metadatos se obtendrán del Diccionario de Datos de PREMIS, en su versión vigente en el momento de realizar el proceso de digitalización (v3.0 a la hora de redactar este documento).
- Asimismo, se incluirá en el METS información sobre derechos de propiedad intelectual según el esquema `METSRights` en el que se detallarán las características de derecho de uso de las publicaciones según determine el Archivo.

La ubicación del archivo XML con formato METS estará basado en el nombre del primer fichero de mencionado en la sección `fileGrp` y tendrá el siguiente formato:

```
<NOMBRE_FICHERO_IMAGEN_SIN_EXTENSIÓN>_mets.xml
```

6.2. ALTO OCR

Se recomienda, que siempre que sea viable, se realice procesado de OCR (Optical Character Recognition / Reconocimiento óptico de caracteres) para las páginas digitalizadas de publicaciones seriadas; la empresa o servicio de digitalización deberá entregar un archivo OCR en formato ALTO (Analyzed Layout and Text Object) por cada página digitalizada. Estos archivos OCR hacen posible la búsqueda en el texto de las páginas digitalizadas.

La ubicación del archivo XML con formato ALTO estará basado en el nombre del primer fichero de mencionado en la sección `fileGrp` y tendrá el siguiente formato:

```
<NOMBRE_FICHERO_IMAGEN_SIN_EXTENSIÓN>_alto.xml
```

Esta recomendación se realiza sin detrimento de la inclusión del procesado OCR en los documentos en formato PDF.

7. PROPIEDAD INTELECTUAL

Como norma general, en el marco de los proyectos de digitalización, sólo se digitalizarán obras que estén en el dominio público o cuyos derechos pertenezcan al Archivo responsable del proyecto de digitalización. Como excepción a esta norma, podrán digitalizarse obras sujetas a derechos de la propiedad intelectual cuando se hayan cedido previamente los derechos de reproducción y comunicación pública, por medio de un convenio o acuerdo, al Archivo responsable del proyecto de digitalización.

El titular de los derechos de propiedad intelectual de las imágenes resultantes del servicio de digitalización será el Archivo encargado del proyecto de digitalización, no pudiéndose reproducir las mismas sin el permiso previo de dicho Archivo.

De acuerdo con las recomendaciones de la Comisión Europea, los materiales que se encuentren en el dominio público se mantendrán en el dominio público después de su digitalización.

Los metadatos resultantes de la digitalización se licenciarán bajo CC0 1.0, y serán reutilizables sin restricción, incluidos los fines comerciales tal como se especifica en el Acuerdo de intercambio de datos con Europeana publicado el 1 de julio de 2012.

La empresa o servicio de digitalización se comprometerá a garantizar la confidencialidad de los datos recibidos en el marco del servicio de digitalización y a no hacer ningún uso de los mismos fuera de lo establecido en acuerdo o contrato previo, siendo responsable de las consecuencias que se deriven del incumplimiento de este apartado. Así mismo no podrá hacer ningún uso de los registros e imágenes generados mediante el servicio de digitalización ni de los registros o imágenes que le puedan ser entregados para la realización de dicho servicio por el Archivo conservador de los fondos a digitalizar o el Archivo que gestione el proyecto.

En los metadatos de las imágenes se añadirán las condiciones de licencia de uso de acuerdo con la codificación de <http://rightsstatements.org>. DARA promoverá el acceso a esa información mediante los enlaces a las páginas correspondientes de <http://rightsstatements.org>

8.- MARCO NORMATIVO Y DOCUMENTOS DE INTERÉS

Ley orgánica 15/1999, de 13 de diciembre, de Protección de Datos de Carácter Personal.

Real Decreto 1164/2002 de 8 de noviembre, por el que se regula la conservación del patrimonio documental con valor histórico, el control de la eliminación de otros documentos de la Administración General del Estado y sus organismos públicos y la conservación de documentos administrativos en soporte distinto al original

Ley 39/2015, de 1 de octubre, del Procedimiento Administrativo Común de las Administraciones Públicas

Real Decreto 1671/2009, de 6 de noviembre, por el que se desarrolla parcialmente la Ley 11/2007, de 22 de junio, de acceso electrónico de los ciudadanos a los servicios públicos. (Derogado parcialmente)

Real Decreto 4/2010, de 8 de enero, por el que se regula el Esquema Nacional de Interoperabilidad en el ámbito de la Administración Electrónica.

Resolución de 19 de julio de 2011, de la Secretaría de Estado para la Función Pública, por la que se aprueba la Norma Técnica de Interoperabilidad de Digitalización de Documentos.

Tutorial de digitalización de imágenes (Biblioteca de la Universidad de Cornell / Departamento de Preservación y Conservación)

Anexos

Anexo I:

Pliego de prescripciones técnicas para digitalización de fondos de archivo

1.- OBJETO DEL CONTRATO

Se consignará en este apartado la denominación general del proyecto, así como las principales tareas a llevar a cabo:

Ej. Proyecto de digitalización de los expedientes de Responsabilidades Políticas del Archivo Histórico Provincial de Huesca

Tareas a realizar:

- Preparación de los originales para su digitalización.
- Captura digital de las imágenes con un escáner de luz fría.
- Preparación de las imágenes para su conservación y consulta.

Asimismo, en la medida de lo posible, se incluirá toda aquella información de la documentación que ayude a la empresa adjudicataria a calcular y valorar adecuadamente el trabajo a realizar:

- Volumen de documentación y estimación del número de imágenes resultante

Ej. 14 cajas de archivo y 4 carpetas. 15.000 imágenes aproximadamente

- Tamaño de los documentos, al menos de los más frecuentes, con indicación de posibles excepciones; grosor de los volúmenes encuadernados; presencia de hojas desplegadas y hojas sueltas, etc.
- Foliación o no de la documentación.
- Estado de conservación de la misma, con expresión de los posibles afecciones (hongos, tintas degradadas, roturas, suciedad, etc.).
- Presencia de elementos que puedan dificultar la digitalización (cosidos, grapas, bisagras, clips, etc.).
- Sistemas de embalaje de la documentación (cajas, carpetas, sobres, rulos, plegados, etc.), indicando al menos, los más frecuentes.

Ej. Embalaje en: cajas de archivo (90%); legajos (10%)

2.- CARACTERÍSTICAS TÉCNICAS

2.1.- Requisitos de la imagen electrónica.

Se consignarán en este apartado las características específicas que habrán de tener las imágenes electrónicas obtenidas a través de la digitalización, de acuerdo con las necesidades que tengamos y en el marco de lo establecido o recomendado por IMADARA, en lo relativo a:

- Ficheros maestros (textuales y gráficos)

Ej. Formato TIFF, revisión 6.0 sin compresión

- Ficheros de difusión

Ej. Documentos textuales: formato PDF (fichero no superior a 50 MB)

Ej. Documentos gráficos: formato JPEG

- Miniaturas

Ej. Tamaño de la miniatura: 400 pixeles por el lado más largo

Ej. Espacio de color RGB a 24 bits

- Marcas de agua

Ej. Se añadirá en el margen inferior derecho de cada página, con un tamaño proporcional al documento para que no destaque

- Nivel de resolución (sin ir en detrimento de las exigencias propias de cada Archivo). Se sugiere fijar un mínimo de resolución y, en caso necesario, corregir al alza.

Ej. Tamaño de letra más común: 1 mm. Nivel de resolución mínima: 300 ppi

- Garantía de imagen fiel

Ej. Se incluirá una imagen de muestra que incluya una carta de color para poder garantizar la idoneidad del balance tonal de la imagen

2.2.- Proceso de digitalización

Si se considera oportuno, se consignará en este apartado, de acuerdo con lo establecido en IMADARA, ciertas pautas en relación con:

- La manipulación de originales por parte del operador

- Los dispositivos idóneos para la captura

- El ajuste de parámetros de los mismos

También se incluirá en este apartado, de acuerdo con lo que al respecto establece IMADARA, lo relativo a:

- Codificación de ficheros: estructura lógica que deben presentar los ficheros de imágenes agrupados en carpetas

Ej. AHPZ_J_00012_0002_0001

- Optimización de imagen

- Validación

- Requisitos de entrega de soportes

Ej. El soporte físico empleado para la entrega serán discos duros USB 3.0, que contendrán tanto los archivos de imagen como los metadatos asociados a los mismos.

Si los soportes están sujetos a normas de identidad corporativa, se harán constar en este apartado:

Ej. En el disco duro aparecerá el logo del Gobierno de Aragón sin departamento.

2.3.- Requisitos opcionales

Se especificarán en este apartado las características que habrán de tener los ficheros resultantes de la digitalización, de acuerdo con lo establecido por IMADARA, en lo relativo a.

- Metadatos

Ej. Cada título se entregará con un fichero XML, estructurado según el esquema METS.

Los METS deben incluir direcciones correctas a los namespaces y a los schemas utilizados y deben estar bien formados y ser válidos

- Reconocimiento óptico de caracteres (OCR)

Ej. Para las páginas digitalizadas de publicaciones seriadas, la empresa de digitalización deberá entregar un archivo OCR en formato ALTO

3.- CONTROL DE CALIDAD

Se incluirán en este apartado las cláusulas que establezca el centro directivo para la revisión y control de los trabajos de digitalización, de acuerdo con las recomendaciones que al respecto se hacen en IMADARA.

Ej. La empresa deberá garantizar el correcto funcionamiento de las imágenes digitalizadas durante el plazo de un año, a contar desde la fecha de finalización del servicio, obligándose a realizar gratuitamente las correcciones necesarias.

4.- PROPIEDAD INTELECTUAL

Se consignarán en este apartado las cláusulas que afecten a la propiedad intelectual de las imágenes resultantes, de acuerdo con lo establecido al respecto en IMADARA.

Ej. El titular de los derechos de propiedad intelectual de las imágenes resultantes será el Archivo responsable del proyecto de digitalización, no pudiéndose reproducir las mismas sin el permiso de dicho Archivo.

5.- LUGAR DE EJECUCIÓN DE LOS TRABAJOS

Se incluirá en este apartado una breve descripción del espacio de trabajo, indicando todos aquellos factores que puedan afectar positiva o negativamente en el tiempo de realización de las tareas de digitalización. Tales como:

- Horario de apertura y cierre del centro
- Mobiliario de la sala de trabajo
- Existencia de suficientes tomas de luz y conexión a Internet
- Existencia de un sistema de climatización
- Dinámica de trabajo en lo relativo a recogida y devolución de documentos
- Distancia de la sala de digitalización al depósito donde se encuentra la documentación
- Medios para el traslado de documentación

Ej. La digitalización a partir de los originales se realizará en las instalaciones del Archivo. El lugar concreto y el horario para la realización de los trabajos de captura digital será establecido por cada Archivo y la empresa adjudicataria de mutuo acuerdo.

La sala de trabajo cuenta con suficientes tomas de luz y acceso a Internet.

La documentación será servida por el personal del Archivo, previa petición del operador.

6.- MEMORIA

Se consignará en este apartado la necesidad de presentar una memoria del trabajo desarrollado, especificando las partes del mismo que obligatoriamente habrá de reflejar.

Ej. El día de terminación del contrato se deberá entregar una memoria de las tareas realizadas. En la memoria se especificarán, al menos, las condiciones técnicas, procedimiento, plazos, volumen, así como los errores detectados y la forma de subsanarlos.

7.- PRESUPUESTO DEL CONTRATO

Ej. El importe del presupuesto del contrato es de 6. 000 euros, IVA incluido, con cargo a la aplicación presupuestaria número 2345

8.- PLAZO DE EJECUCIÓN

Ej. El plazo de ejecución será de 3 meses, a partir de la firma del contrato.

9.- REQUISITOS DEL ADJUDICATARIO

Si procede, se establecerá en este apartado la acreditación de la competencia técnica del responsable

Ej. El responsable de la ejecución del contrato deberá estar en posesión de titulación requerida y haber realizado trabajos relacionados con la materia de contratación en los tres últimos años.

10.- DOCUMENTACIÓN A APORTAR POR LOS LICITADORES

Se consignará aquí la documentación a aportar por los licitadores, de acuerdo con el pliego de cláusulas administrativas particulares.

Anexo II:

Nivel de resolución

En el epígrafe 3.2. *Nivel de resolución* se hace referencia al índice de calidad basado en la fórmula desarrollada por el Departamento de Preservación y Conservación de la Biblioteca de la Universidad de Cornell. Dicha fórmula expresa que la resolución óptima para una imagen en escala de grises o color RGB se puede describir como:

$$\text{Resolución} = (1,5 * QI) / (0,039 * h)$$

1,5= Constante que corrige errores de muestreo.

QI= Índice de calidad que queremos aplicar.

0,039= Factor de conversión de pulgadas a milímetros (o multiplicar por 25,4).

h= Tamaño en milímetros del tipo de letra más pequeño.

Para su inclusión en DARA, las imágenes digitalizadas deberán tener un índice QI de 8.0, por lo que:

$$\text{Resolución} = (1,5 * 8) / (0,039 * h)$$

o lo que es lo mismo:

$$\text{Resolución} = (12 / h) * 25,4$$

Tomando como ejemplo un tamaño de letra de 1mm la resolución resultante sería de :

$$\text{Resolución} = (12 / 1) * 25,4 = 304,8 \text{ ppi}$$

Por ello, se propone como guía de referencia la tabla de resoluciones según el tamaño de letra más pequeño para el que se quiera conservar su legibilidad:

Fórmula	Tamaño de letra	Resolución aproximada
$\text{Resolución} = (12 / 2) * 25,4 = 152,4$	2 mm	150 ppi
$\text{Resolución} = (12 / 1,5) * 25,4 = 203,2$	1,5 mm	200 ppi
$\text{Resolución} = (12 / 1) * 25,4 = 304,8$	1 mm	300 ppi
$\text{Resolución} = (12 / 0,75) * 25,4 = 406,4$	0,75 mm	400 ppi
$\text{Resolución} = (12 / 0,6) * 25,4 = 508$	0,6 mm	500 ppi
$\text{Resolución} = (12 / 0,5) * 25,4 = 609,6$	0,5 mm	600 ppi

Para cualquier imagen digital, hay que entender que el valor numérico de resolución de la imagen expresado en ppi no significa nada por sí mismo si no conocemos el tamaño del documento original. Digitalizar un documento en un tamaño de imagen de 4960 x 3508 píxeles, solo indica que podríamos obtener un A5 a 600 ppi, un A4 a 400 ppi, un A3 a 300 ppi, un A2 a 200 ppi, un A1 a 150 ppi, etc. pero el tamaño de la imagen digital siempre es el mismo (4960 x 3508 píxeles).

Cuando hablamos de resolución nos referimos a la cantidad de píxeles que agrupamos en una misma superficie (normalmente pulgadas) en la imagen digital. Se expresa como píxeles por pulgada y explica de qué modo distribuimos los píxeles totales de una imagen cuando vamos a digitalizar o a imprimir.

Anexo III:

Control de calidad

Para llevar a cabo un control de calidad de la documentación digitalizada lo primero que hay que hacer es establecer unos parámetros mínimos que consideremos aceptables. Dichos parámetros pueden variar entre Archivos e, incluso, entre colecciones dentro del mismo Archivo.

Para una colección de fotografías nos interesa sobre todo que se aprecie el artefacto gráfico, rango dinámico del original, detalles, etc. Sin embargo para documentos textuales el interés principal suele ser que se pueda distinguir la letra del original con facilidad. Es por ello, que cada Archivo debe fijar sus propios criterios para cada tipo de documentación.

Como recomendación, sugerimos visualizar las imágenes con un tamaño del 100% y llevar a cabo una revisión sobre una muestra aleatoria no consecutiva (preferiblemente se escogerán imágenes realizadas en distintos días y con distintos dispositivos si es el caso) de cada uno de los formatos entregados. Se recomienda que dicha muestra sea de, al menos, un 10% del total de las imágenes.

Conteste a las preguntas planteadas en el siguiente cuadro, cuyas respuestas deben ser en todos los casos afirmativas. Si para la totalidad de la muestra hay alguna respuesta negativa (fuera del margen de los criterios establecidos), se realizará otra revisión aleatoria de distintas imágenes y en caso de que se repita dicha anomalía no se podrá dar por buenas las imágenes digitales.

	SÍ	NO
¿Corresponden los formatos entregados con los formatos solicitados?		
¿Los formatos digitales respetan las características técnicas requeridas?		
¿Se han digitalizado todas las páginas y documentos solicitados?		
¿El conjunto de imágenes pertenecientes al mismo documento es una representación fiel e íntegra de la unidad documental?		
¿Las imágenes digitales tienen la resolución adecuada o el índice QI indicado para su correcta visibilidad / lectura?		
¿Las imágenes digitales tienen un enfoque suficiente para poder apreciar detalles según su resolución de captura, sin abusar de filtros de corrección?		
¿La iluminación es uniforme en todas las partes del documento y el color de la imagen digital representa lo más fielmente posible al color del original?		
¿Las imágenes digitales están correctamente alineadas?		
¿Tienen las imágenes digitales el margen indicado?		
¿Respetan los ficheros de difusión el límite de tamaño sugerido?		
¿Están los ficheros de difusión optimizados para internet?		
¿Respetan los ficheros de difusión la ubicación y nivel de transparencia de la marca de agua en caso de llevarla?		

Como ayuda para comprobar si las imágenes “master” tienen la resolución adecuada o el índice QI indicado para su correcta lectura podemos consultar la siguiente tabla de tamaños estandarizados:

Resolución recomendada según tamaño del documento y tamaño de letra
--

Tamaño	2 mm	1,5 mm	1 mm	0,75 mm	0,6 mm	0,5 mm
A0	7016 x 4960	9933 x 7016	14043 x 9933	18724 x 13244	23406 x 16555	28087 x 19866
A1	4960 x 3508	7016 x 4960	9933 x 7016	14043 x 9933	16555 x 11693	18724 x 13244
A2	3508 x 2480	4960 x 3508	7016 x 4960	9933 x 7016	11693 x 8268	14043 x 9933
A3	2480 x 1748	3508 x 2480	4960 x 3508	7016 x 4960	8268 x 5846	9933 x 7016
A4	1748 x 1240	2480 x 1748	3508 x 2480	4960 x 3508	5846 x 4134	7016 x 4960
A5	1165 x 827	1748 x 1240	2480 x 1748	3508 x 2480	4134 x 2913	4960 x 3508
A6	874 x 620	1165 x 827	1748 x 1240	2480 x 1748	2913 x 2067	3508 x 2480

Si queremos saber qué tamaño en píxeles debe tener una imagen de un documento digitalizado a una resolución concreta, aplicamos la siguiente fórmula:

Tamaño en píxeles = Longitud del documento (en cm) * Resolución / 2,54

Por ejemplo, si queremos conocer el tamaño de la imagen de un documento A4 a una resolución de 300 ppi:

Tamaño en píxeles lado mayor = $29,7 * 300 / 2,54 = 3508$

Tamaño en píxeles lado menor = $21 * 300 / 2,54 = 2480$

Tamaño en píxeles = $3508 x 2480$

Anexo IV:

Hoja de operador

Signatura	Fecha de digitalización	Observaciones

