

Recomendaciones técnicas

01. Recomendaciones Técnicas para la digitalización de documentos

Versión 1.0 de 02/07/2010



1. OBJETO Y ALCANCE

La Consejería de Cultura de la Junta de Andalucía, siguiendo la *Recomendación de la Comisión Europea de 24 de agosto de 2006 sobre la digitalización y la accesibilidad en línea del material cultural y la conservación digital*, ha puesto en marcha el Banco de Documentos de Andalucía (BAiDA), como repositorio digital común para el patrimonio documental conservado en los archivos del Sistema Andaluz de Archivos, que permitirá la consulta de las imágenes de los documentos a través de un portal web.

Para la adecuada integración de las imágenes en este repositorio digital, así como para la salvaguarda del Patrimonio Documental Andaluz mediante la constitución de un archivo de seguridad del mismo, se redactan las presentes recomendaciones, con una serie de pautas comunes que deberán ser cumplidas durante la ejecución de los distintos proyectos de digitalización que se lleven a cabo en los archivos andaluces.

Asimismo, estas directrices técnicas podrán ser tenidas en cuenta en cualquiera de los procesos de digitalización emprendidos en los archivos de Andalucía, consistentes en la conversión a formato digital de documentos originales no digitales, generalmente manuscritos.

Estos procesos, junto a una manifiesta finalidad de incrementar el acceso a los documentos de archivo mediante el uso de internet, comparten otros objetivos: reducir la reiterada manipulación y la consulta directa de los documentos originales, frágiles y en muchos casos deteriorados, y crear una copia de seguridad de unos materiales únicos y valiosos.

2. DOCUMENTACIÓN SOBRE EL PROCESO DE DIGITALIZACIÓN

Debe tenerse en cuenta que es importante que el propio proyecto de digitalización de documentos se encuentre bien documentado. Así pues, antes de desarrollar la digitalización en sí deberá crearse un documento en que se refleje claramente cuales serán los criterios de digitalización a emplear. Ello dará continuidad al trabajo, algo de utilidad en caso de que cambien los equipos que lo desarrollen, y servirá de apoyo para futuras planificaciones o replanteo de los trabajos de reproducción realizados.

3. CAPTURA DE LA IMAGEN

En el proceso de digitalización se tendrán en cuenta las siguientes consideraciones:

1. La manipulación de los documentos se realizará según los criterios de conservación preventiva comúnmente aceptados para los materiales de archivo. El personal técnico del archivo podrá



establecer pautas específicas para la manipulación de determinados materiales, de acuerdo a sus características, con el fin de garantizar su conservación e integridad. Se examinará cuidadosamente el conjunto de documentos a reproducir, para determinar si es preciso establecer criterios de manipulación específicos.

2. Se utilizarán guantes y batas y se evitará la excesiva manipulación.
3. De modo general, se optará por capturar los documentos dentro de medidas fijas, de modo que se facilite la operación de recortado de la imagen, automatizando ésta si es posible.
4. No se desencuadernarán los documentos ni se adoptarán medidas que vayan contra su integridad física, ni aunque la finalidad sea obtener una mejor reproducción.
5. Los originales con problemas de conservación se abrirán con un ángulo nunca superior a 120°.
6. La mesa portalibros no debe permitir el uso de prensa o cristal externo alguno sobre el original.
7. La mesa portalibros o el balancín no debe emplear sistemas hidráulicos; tiene que utilizar sistemas mecánicos o electrónicos.
8. La digitalización se realizará preferentemente con un escáner aéreo o cenital. Otros dispositivos de captura de la imagen posibles son la cámara de fotos digital y el escáner de sobremesa.
9. En ambos casos, sería aconsejable el uso de cartas de escala de grises y/o color, de cara a obtener un perfil de color que sea fiel a los originales capturados.

En caso de optar por el uso del escáner, deberá tenerse en cuenta lo siguiente:

1. El escáner de sobremesa solamente deberá utilizarse para digitalizar documentos que no estén encuadernados.
2. El escáner no mantendrá la iluminación directa sobre el original o carecerá de ella al finalizar la captura digital.
3. El escáner deberá utilizar una iluminación de débil potencia o podrá trabajar con luz ambiente. En el caso de que el escáner tenga iluminación, ésta tendrá una potencia de 1 a 126 lux/hora, y no podrá liberar una temperatura superior a 0,5° centígrados en relación a la temperatura ambiente.
4. El escáner deberá tener capacidad de capturar originales desde formatos DIN A1; igualmente tendrá capacidad de capturar volúmenes que alcancen hasta 50 cm. de grosor.
5. La iluminación del escáner será variable, permitiendo modificar el ángulo del rayo de luz para posibilitar la captura de originales con relieve (sellos, grabados, etc.) y originales con brillo.

En caso de elegir la cámara de fotos digital para la captura de imágenes, habrá que tomar en consideración lo siguiente:



1. La distancia focal recomendada para la captura será de 50 mm. En la práctica, y teniendo en cuenta que el factor de conversión de las focales digitales es de 1,6 aprox., el objetivo a emplear deberá tener una distancia focal de 35 mm.
2. La columna que sujete a la cámara digital deberá estar nivelada y en posición cenital respecto al documento a digitalizar.
3. Deberá evitarse el uso de flash, eligiendo en su lugar una iluminación directa sobre el original, sea natural o artificial, y en cualquier caso de débil potencia de luxes.

4. RESOLUCIÓN Y CONTROL DE CALIDAD

1. Durante todo el proceso de digitalización es necesario efectuar controles de calidad a fin de comprobar el adecuado cumplimiento de los criterios establecidos y que los productos obtenidos se adecuan a los parámetros definidos en lo relativo a resolución, profundidad de color en bits, formato y compresión, si la hay.
2. El proceso de digitalización estará supervisado por un técnico en imagen digital, con la finalidad de controlar la calidad del producto resultante y hacer las recomendaciones y modificaciones convenientes, en su caso.
3. La supervisión se realizará a partir de la visualización de la imagen en el monitor o bien desde la operación matemática del *índice de calidad (QI) para la escala de grises y color*, fórmula comúnmente aceptada que pone en relación la calidad de la reproducción con el tamaño del carácter de los documentos y la resolución de captura escogida en la digitalización. Este Índice se recoge en el Anexo I.
4. El tamaño del archivo digital, el tipo de archivo y la resolución de captura dependerán de las características de los originales a reproducir. Se adoptará como patrón base el cuadro comparativo del Anexo II. Las características del original, en lo que se refiere a tipo de letra, nivel de contraste, etc., condicionarán que la imagen digital pueda ser modificada respecto al original, en orden a facilitar su legibilidad.
5. Se efectuarán ensayos de acceso a los ficheros con el fin de establecer un equilibrio entre calidad visual aceptable para el usuario y el tamaño del fichero, de manera que el tiempo de captura sea amable. De modo general, la captura digital se realizará en color (RGB, a 24 bits), con un índice de calidad igual a Bueno ($QI = 5.0$). Los técnicos del Archivo podrán decidir el cambio a escala de grises (a 8 bits) cuando el original así lo requiera.
6. La captura digital de originales de gran formato (mapas, planos, pergaminos, etc.) se realizará para la copia de conservación en formato TIFF, utilizando un sistema de cuadrícula. Para la difusión de estos grandes formatos se realizará además una captura general con una



resolución óptima en sistema de compresión JPEG o PNG, y se dividirá en coordenadas, con objeto de obtener una malla cuadriculada. Cada cuadrícula de la malla se capturará de manera separada con un formato no superior a DIN A3, en alta resolución, y con referencia a la captura general mediante el uso de enlaces o hipervínculos.

7. Cuando se realicen capturas con un sistema de compresión, se optará por el grado de compresión que garantice que la calidad de la imagen sea óptima y la pérdida de información mínima. El índice de calidad (QI) mínimo exigido será de 5.0.
8. En las imágenes digitales obtenidas se incorporarán los metadatos adecuados de control que se indican en la estipulación quinta, en especial el Archivo de procedencia, la signatura y la fecha de digitalización.
9. La capacidad mínima de captura del sensor, sea CCD o CMOS, del escáner o cámara digital será de 6.000 píxeles en escala de grises y color RGB.
10. Podrán realizarse todos aquellos procesos encaminados a la mejora y optimización de la imagen captada, cambiando el color, el brillo y el contraste, bajo las directrices del responsable técnico que supervise los trabajos. Dichos procesos podrán documentarse.

5. FORMATOS Y SOPORTE DE LAS COPIAS

Como resultado de la digitalización de documentos se deben obtener dos tipos de copias:

5. 1. Copia de conservación

Son reproducciones digitales de alta calidad realizadas con fines de preservación, en formatos RAW o TIFF sin compresión agresiva que afecte a la calidad de la imagen. Tendrán un índice de calidad (QI) igual a Bueno (QI=5.0) (véase Anexo I) y serán realizadas de acuerdo a las especificaciones técnicas recogidas en el Anexo II.

5. 2. Copia de consulta

Se trata de reproducciones digitales destinadas a la consulta o difusión a través de redes de comunicación, deberán tener un formato comprimido con pérdida, con una ratio de compresión alta que optimice la velocidad de transmisión y descarga en los servidores y, a la vez, permita la legibilidad adecuada de la copia. Contarán con una resolución de 100 dpi. En el caso de imágenes simples se optará por el formato JPG, y en el caso de imágenes múltiples se optará por el formato PDF.



5. 3. Marcado de imágenes

Las copias digitales obtenidas deberán ir convenientemente provistas de una marca de agua visible, que identifique el Archivo de procedencia/la institución que custodia los documentos originales.

En las imágenes de documentos reproducidas con cargo a proyectos de digitalización realizados por la Consejería de Cultura de documentos custodiados en sus archivos, se incorporará el sello de la Junta de Andalucía asociado al nombre de ésta y del Centro Directivo de que se trate, de acuerdo al vigente Manual de Identidad Corporativa de la Junta de Andalucía, consultable en la página web de la Consejería de la Presidencia (Comunicación Social, apartado 2.1.28).

Dicho sello, que funcionará como una marca de agua, tendrá en su tamaño mayor unas dimensiones aproximadas de 350 mm de diámetro y en cualquier caso nunca será superior al 10% del ancho de la imagen; además, presentará una opacidad no inferior al 40%, en escala de grises. E irá colocada en un lugar que no altere la apariencia original del documento, preferentemente en el ángulo inferior izquierdo de la imagen.

En las imágenes de documentos realizadas con cargo a subvenciones de reproducción concedidas por la Consejería de Cultura, deberá hacerse constar, tal como determina la legislación vigente en materia de subvenciones¹, de forma expresa el patrocinio de la Junta de Andalucía, indicando que ha sido concedida por la Consejería de Cultura de la Junta de Andalucía, incluyéndose su logotipo donde figure la Dirección General implicada.

5. 4. Almacenamiento y soportes de conservación de los trabajos

Con carácter general, se elegirán como soportes de almacenamiento de las imágenes digitales aquellos que ofrezcan una mayor garantía para la conservación y preservación inalterable de la información.

Las imágenes estarán archivadas en una estructura jerárquica de carpetas, con forma de árbol, que refleje el esquema de organización de los fondos dentro del Archivo de procedencia, siendo la última carpeta la correspondiente a la signatura de la unidad de instalación (caja, legajo, libro, etc.), y dentro de ella se instalarán los ficheros con las imágenes de dicha unidad de instalación.

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

¹ Orden de 6 de febrero de 2006, por la que se aprueban las bases reguladoras para la concesión de subvenciones a las Entidades Locales de la Comunidad Autónoma de Andalucía, para la realización de proyectos de conservación, tratamiento y difusión del Patrimonio Documental Andalúz conservado en archivos de titularidad local y se efectúa su convocatoria para el año 2006, (BOJA nº 39, de 27 de febrero), art. 15, aptdo. i.



Las copias digitales y sus respectivos metadatos se almacenarán por duplicado y de la siguiente manera:

- Una copia de conservación y una copia de consulta, almacenadas en discos magnéticos de al menos 500 GB de capacidad, con formato NTFS y conectables mediante puerto USB 2.0 ó Firewire I ó II; se conservarán en la sede del Archivo o de la institución titular de los documentos.
- Una copia de conservación y una copia de consulta, almacenadas en DVD, que se conservarán en unas dependencias distintas a la sede del Archivo y de la institución titular de los documentos.

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.

En cualquier caso, respecto a la protección de la información almacenada por medios electrónicos se estará a lo dispuesto en el *Real Decreto 3/2010, de 8 de enero, por el que se regula el Esquema Nacional de Seguridad en el ámbito de la Administración Electrónica* (BOE núm. 25, de 29 de enero, corr. errores 11 de marzo), y para el copiado de documentos con formatos en riesgo de obsolescencia se estará a lo dispuesto en el *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* (BOE núm. 25, de 29 de enero).

Los soportes de almacenamiento irán identificados con los datos de la institución responsable de la digitalización. En el ámbito institucional de la Junta de Andalucía esta identificación se elaborará de acuerdo con las recomendaciones establecidas en materia de Identidad Corporativa y consistirá en la marca genérica asociada al nombre de la Consejería u Organismo Autónomo, a la que se podrá incorporar el nombre del Archivo o Unidad Administrativa.

En el caso de reproducciones realizadas por otras administraciones, los soportes deberán ir identificados con los datos institucionales necesarios para constatar su procedencia y, en su caso, la titularidad de las imágenes.



Todos los soportes contarán con la información necesaria para conocer su contenido: número, título, contenido y fecha de creación.

6. METADATOS

Las imágenes digitales deberán contar con los metadatos descriptivos y técnicos necesarios para categorizarlas dentro de un contexto que facilite su comprensión y recuperación.

Por un lado, los metadatos descriptivos identifican el fichero con la información descriptiva del documento de archivo que ha sido digitalizado en él. Para cumplimentar estos metadatos descriptivos del activo digital, se utilizarán, como fuente principal, los instrumentos de descripción existentes en el Archivo de procedencia.

Se utilizará el estándar de metadatos DUBLIN CORE D.C. (<http://dublincore.org>). Los metadatos se cumplimentarán en la base de datos dispuesta al efecto, en la que se incorporarán, al menos, los Elementos D.C. considerados obligatorios en el Anexo III.

Por otro lado, las imágenes o ficheros llevarán incrustados metadatos técnicos con información relativa al proceso de digitalización, de acuerdo al estándar EXIF 2.x, siempre que sea posible.

7. FUENTES

Para la elaboración de estas directrices técnicas se han tenido en cuenta diferentes publicaciones y documentos técnicos, siendo a destacar los siguientes:

Asociación Española de Documentación e Información (2007). "Curso de digitalización de documentos"

Campuzano Juan, Francisco; Jiménez Chornet, Vicent (2005). "Recomendaciones técnicas para la captura y digitalización de fondos de Archivos del Reino de Valencia" [en línea]. 2005, mayo. <<http://dglab.cult.gva.es/ArxiuRegne/documents/Recdigitalizacion.pdf>> [consulta 2010/04/22]

Centre de Recerca i Difusió de la Imatge (CRDI), Girona, (2010). "Recursos en línea" [en línea] <http://www.girona.cat/sgdap/cat/crdi_portada.php> [consulta 2010/04/22]

Consejo Internacional de Archivos, (2010). "Recursos en línea" [en línea]. <<http://www.ica.org>> [consulta 2010/04/22]



Corujo, Luis; Sant´Ana, Mario (2010). “Recomendações para o produção de planos de preservação digital” [en línea]. Direcção Geral de Arquivos de Portugal, 2010, abril. <http://dgarq.gov.pt/files/2010/04/planopreservacaodigital_v2-0b.pdf> [consulta 2010/05/03]

Torres Freixinet, Luis (2007). “Recomendaciones sobre digitalización de documentos”. Curso Asociación de Archiveros de Andalucía.

Torres Freixinet, Luis (2009). “Digitalización del patrimonio documental”, en *TRIA, Revista Archivística de la Asociación de Archiveros de Andalucía*, núm. 15, 2009, pp. 241-298.

Trujillo Jiménez, Javier (2009). “Archivo y gestión de imágenes digitales”, en *TRIA, Revista Archivística de la Asociación de Archiveros de Andalucía*, núm. 15, 2009, pp. 299-334.

UNESCO (2002). “Directrices para proyectos de digitalización de colecciones y fondos de dominio público”. [en línea]. http://www.mcu.es/archivos/docs/pautas_digitalizacion.pdf [consulta 2010/04/26]



ANEXO I

TABLA DE ÍNDICE DE CALIDAD APLICABLE PARA LA CAPTURA DIGITAL SOBRE LA BASE DE UNA ESCALA 1:1.

ÍNDICE DE CALIDAD (QI)

Para los modos escala de grises y color RGB

ÍNDICE DE CALIDAD	QI
CASI ILEGIBLE	3.0
MÍNIMO	3.6
BUENO	5.0
SUPERIOR	8.0

RESOLUCIÓN DE DIGITALIZACIÓN

En relación a tamaño de carácter e Índice de Calidad

Tamaño del carácter	Dpi 100		Dpi 200		Dpi 250		Dpi 300		Dpi 400	
	Texto manuscrito	Texto impreso								
m/m										
1.0	1.3	1.7	2.6	3.8	3.2	4.8	3.9	5.8	5.2	7.8
2.0	2.6	3.8	5.2	7.8	6.5	9.7	7.8	11.7	10.4	15.6
3.0	3.9	5.8	7.8	11.7	9.7	14.6	11.7	17.5	15.6	23
4.0	5.2	7.7	10.4	15.6	13	19	15.6	23	20.8	31
5.0	6.5	9.7	13	19.5	16	24	19	29	26	39

Tomado de Campuzano y Jiménez (2005)



ANEXO II

CUADRO COMPARATIVO

Características materiales de originales	Resolución de captura/tipos de archivo a obtener	Modos de Color (Escala de grises/color RGB)
PUBLICACIONES PERIODICAS (Sólo texto, DIN A4)	200 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
PUBLICACIONES PERIODICAS (Texto e imagen b/n, DIN A4)	200 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
PUBLICACIONES PERIODICAS (Sólo texto, DIN A3)	300 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
PUBLICACIONES PERIODICAS (Texto e imagen b/n, DIN A3)	300 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
TEXTO IMPRESO, DIN A4	200 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
TEXTO IMPRESO E IMAGEN, DIN A4	250 dpi/TIFF sin comprimir	Color, a 24 bits
MANUSCRITO MONOCROMÁTICO, DIN A4	200 dpi/TIFF sin comprimir	Color, a 24 bits
MANUSCRITO COLOR, DIN A4	300 dpi/TIFF sin comprimir	Color, a 24 bits
FOTOGRAFIA/OPACOS B/N (10x15 cm.; 20x24 cm.; 24x30 cm.; 24x30 cm.)	300 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
FOTOGRAFIA/OPACOS COLOR (10x15 cm.; 20x24 cm.; 24x30 cm.; 24x30 cm.)	300 dpi/TIFF sin comprimir	Color, a 24 bits
FOTOGRAFIA/NEGATIVOS Y DIAPOSITIVAS B/N (paso universal de 35 mm.)	2.600 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
FOTOGRAFIA/ NEGATIVOS Y DIAPOSITIVAS B/N (formato medio, 4x5.5 cm.)	1.800 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
FOTOGRAFIA/ NEGATIVOS Y DIAPOSITIVAS B/N (formato medio, 6x7cm.)	1.200 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
FOTOGRAFIA/ NEGATIVOS Y DIAPOSITIVAS COLOR (paso universal de 35 mm.)	2.600 dpi/TIFF sin comprimir	Color, a 24 bits
FOTOGRAFIA/ NEGATIVOS Y DIAPOSITIVAS COLOR (formato medio, 4x5.5 cm.)	1.800 dpi/TIFF sin comprimir	Color, a 24 bits



Características materiales de originales	Resolución de captura/tipos de archivo a obtener	Modos de Color (Escala de grises/color RGB)
FOTOGRAFIA/ NEGATIVOS Y DIAPOSITIVAS COLOR (formato medio, 6x7cm.)	1.200 dpi/TIFF sin comprimir	Color, a 24 bits
FOTOGRAFIA/ PLACAS (10x15 cm.)	800 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
FOTOGRAFIA/ PLACAS (15x20 cm.)	600 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
FOTOGRAFIA/ PLACAS (20x25 cm.)	500 dpi/TIFF sin comprimir	Escala de grises, a 8 bits
MAPAS, PLANOS Y PERGAMINOS (formatos hasta DIN A4)	250 dpi/TIFF sin comprimir	Color, a 24 bits
MAPAS, PLANOS Y PERGAMINOS (formatos DIN A3 ó superior)	300 dpi. Captura general en JPEG, y malla cuadrículada en formato DIN A4 en TIFF sin comprimir.	Color, a 24 bits

-Tomado de Torres Freixinet (2007)



ANEXO III

METADATOS

Metadatos Dublin Core:

Metadatos del contenido del recurso		
Elemento Dublin Core (DC)	Información del elemento	Carácter del elemento
DC Source	Código de referencia. Signatura	Obligatorio
DC Title	Título. Nombre del contenido del recurso.	Obligatorio
DC Creator	Nombre del productor.	Obligatorio
DC Subject	Materia y palabras clave. Descriptores.	Recomendado
DC Coverage_Temporal	Fecha(s) del documento.	Obligatorio
DC Coverage_Spatial	Localización espacial. Ámbito geográfico. (Vocabulario controlado)	Recomendado
DC Rights	Derechos de la propiedad intelectual. Condiciones de reproducción.	Obligatorio
DC Language	Lengua. Idioma.	Obligatorio
DC Provenance	Archivo que custodia el documento original analógico.	Obligatorio
DC Type	Naturaleza o género del contenido del recurso. (Vocabulario controlado)	Obligatorio
DC Description	Descripción del contenido del recurso.	Recomendado
DC Relation	Referencia a un recurso relacionado.	Recomendado
Metadatos técnicos del recurso		
Elemento Dublin Core (DC)	Información del elemento	Carácter del elemento
DC Identifier	Identificador del recurso. Referencia unívoca e inequívoca.	Obligatorio
DC Format	Formato de la copia digital.	Obligatorio
DC Format	Resolución.	Recomendado
DC Format	Extensión.	Recomendado
DC Date	Fecha de la digitalización.	Obligatorio
DC Publisher	Editor. Entidad responsable de que el recurso esté disponible.	Obligatorio



Otros metadatos técnicos:

Nombre del metadato	Información del metadato	Carácter del metadato
OTR.Responsable_digitalización	Responsable de la digitalización	Recomendado
OTR.Otra_Información	Otros datos sobre la digitalización: dispositivo empleado en captura, distancia focal, etc.	Recomendado

