

Estandarización de procesos de captura digital en la Biblioteca Nacional de Chile

Roberto Aguirre Bello

RESUMEN

En los últimos años, muchas bibliotecas, archivos y centros de documentación han emprendido el desarrollo de importantes proyectos de digitalización de sus colecciones patrimoniales. Sin embargo, este proceso, que tiene como principal objetivo conservar y facilitar el acceso al patrimonio por parte de la comunidad, puede implicar una pérdida importante de recursos si no se considera en su ejecución una metodología estandarizada de trabajo que permita normalizar desde el punto de vista técnico los procesos de captura digital de los materiales.

En la Biblioteca Nacional de Chile, al iniciarse los proyectos de digitalización “Biblioteca Virtual Cervantes” y “Memoria Chilena”, se estableció a través de un estudio teórico-práctico una propuesta metodológica de trabajo estandarizado, orientado a establecer las condiciones adecuadas para la realización de los procesos de digitalización de colecciones, basado fundamentalmente en las características físicas del material análogo a digitalizar y en criterios de valor patrimonial, estado de conservación y demanda de uso, con el fin de minimizar los factores y/o causas de deterioro que pudieran poner en riesgo los materiales originales producto de su cambio de formato y al mismo tiempo maximizar los recursos disponibles permitiendo la recuperación de la información digital en el largo plazo. Se presentan la metodología aplicada y los resultados obtenidos.

Palabras clave: Digitalización, captura digital, estándares, bibliotecas, archivos, preservación.

ABSTRACT

Over the last years, many libraries, archives and documentation centers have undertaken major digitalization projects with their collections. Nevertheless this process is done for the conservation of the material and improving the access service for the community, it can also involve a significant loss of resources if standardized technical capture processes are not considered.

When the digitalization projects “Biblioteca virtual Cervantes” and “Memoria chilena” were launched at the Biblioteca Nacional de Chile, a theoretical and practical study was carried out in order to generate a standardized work methodology. Proper conditions for implementing digitalization processes with collections were addresses, based on the cultural value, physical features, conservation condition and demand for the use of the items. The goal was to minimize the factors and/or causes of deterioration that could endanger original materials as a result of the format change procedures, while simultaneously maximizing available resources allowing the recovery of digital information in the long term. The results are presented as a work experience.

Key words: digitalization, digital capture, standards, libraries, archives, preservation

Roberto Aguirre Bello, Licenciado en Arte mención Restauración, Pontificia Universidad Católica de Chile. Jefe Archivo Fotográfico y Digital Biblioteca Nacional de Chile.

INTRODUCCIÓN

Uno de los mayores retos para el desarrollo de proyectos que contemplan la puesta en línea de colecciones patrimoniales, es la generación de la propia información digital, debido a que en la mayoría de los casos ésta no se encuentra en tal formato, es decir, se hace necesario la realización de un proceso de transformación de la información que se encuentra en forma analógica a formato digital, proceso que llamamos comúnmente digitalización y que genera la necesidad de que se establezcan normas y procedimientos para su desarrollo de manera adecuada, controlada y sustentable¹.

Foto 1: Portal de la Biblioteca Nacional de Chile en la Biblioteca Virtual Cervantes. Proyecto de digitalización iniciado el año 1999. Primera experiencia de digitalización desarrollada en la Biblioteca Nacional.

Portal de contenidos culturales Memoria Chilena. Proyecto de digitalización iniciado el año 2000. Al año 2006, más de 700.000 imágenes digitalizadas.



Sin embargo, ésta no es una tarea sencilla, pues es necesario considerar varios aspectos como el estado físico y legal de la información a digitalizar, la comunidad y el uso a la que está dirigida, el tipo de formato analógico en que se encuentra (audio, video, microfilm, papel, etc.), además de otros factores propios de la conversión

1 Conway, P., 2000.

digital y sus problemas asociados como el tamaño de archivo (peso), resolución, formato y almacenamiento de la información generada.

Es importante mencionar que, debido a la rapidez de los cambios tecnológicos, al momento de pensar en digitalizar un documento, cualquiera que éste sea, lo primero que debe considerarse es la estrategia de preservación del archivo digital, es decir, la manera de asegurar la vigencia y transportabilidad de la información producida². Esto debido a que cuando se utiliza un formato que no es un estándar, lamentablemente tiende a desaparecer con el paso del tiempo y en consecuencia resulta difícil su conversión hacia otro formato que pudiera ser más estable. Por esta razón, la generación de políticas que permitan establecer el marco de acción para la captura, preservación, almacenamiento y publicación de los documentos digitales resulta imprescindible durante el proceso.

Teniendo en consideración estos conceptos, se ha intentado establecer una aproximación a la estandarización de los procesos de captura en la digitalización de documentos de diferentes soportes para la Biblioteca Nacional de Chile, aun teniendo en cuenta que diferentes organismos nacionales e internacionales coinciden en que no existen soluciones acabadas que resuelvan el problema en forma óptima y que sólo se puede contar con una plataforma mínima sobre la cual se debe orientar el trabajo y que por sobre ésta las soluciones definitivas se encuentran aún en etapa de investigación³.

La estandarización de los procesos en la Biblioteca Nacional se ha basado esencialmente en el formato análogo (tono continuo o trama) del material a digitalizar, considerando a su vez otros factores como su valor patrimonial, el estado de conservación y la demanda futura de utilización de los documentos convertidos a formato digital. A estos factores se suma la necesidad existente en la institución de optimizar el uso de los recursos tecnológicos y la capacidad de almacenamiento disponible. Todos estos aspectos determinan las condiciones y procesos que deben abarcarse desde el punto de vista técnico.

MÉTODOS

El formato análogo del material es el punto de partida para orientar el tipo de digitalización que se deberá realizar para cada material. En este caso nos hemos ocupado del problema que se presenta en los materiales más representativos de las colecciones de la Biblioteca Nacional como son los materiales registrados sobre soporte de papel. Estos se dividen fundamentalmente en:

- Materiales de tono continuo: (imágenes) fotografías, grabados, pinturas, dibujos, manuscritos, mapas o combinados (imágenes / textos).
- Materiales de trama: (imágenes y textos) periódicos, revistas, libros, etc.

2 Hazen, D., Horrell, J., Merrill-Oldham H., 2000.

3 Technical Guidelines for Digitizing Archival Materials for Electronic Access), 2004.

La condición de tono continuo o de trama de un determinado material nos presenta un primer problema para resolver desde el punto de vista de la captura digital, el que se traduce en la estrategia a utilizar para registrar la mayor cantidad de información relevante del material análogo disponible. Las imágenes de tono continuo, al no estar compuestas por una matriz de punto, presentan una mayor cantidad de información recuperable al ser digitalizadas que las imágenes de trama, por esta razón, aumentar la resolución o frecuencia de captura en este tipo de materiales nos permitirá obtener una mayor información del original. Por el contrario, la digitalización en una resolución alta de materiales de trama no nos permitirá obtener mayor información que la existente en la matriz de punto y nos producirá documentos electrónicos de gran tamaño de archivo, aumentando la problemática existente en los procesos de almacenamiento y respaldo de la información digital.



Foto 2. Ampliación de imagen de tono continuo. La Cruz de Reyes. Album Fotográfico "Vistas de Valparaíso" de Felix Leblanc. C. 1890. Colección Sala Medina Biblioteca Nacional. Presenta mayor cantidad de información recuperable.

Ampliación de imagen de trama. La Cruz de Reyes. Impresión portada libro "Historia de la Fotografía en Chile: Rescate de Huellas en la Luz. 2000. Presenta matriz de punto.



El valor del material en cuanto a su forma y contenido, al igual que su estado de conservación, nos indicarán la necesidad de la ejecución de procesos de digitalización orientados a la preservación de los materiales, labor que debe realizarse de manera estandarizada. Por su parte, las necesidades para su utilización dirigirán los procedimientos en cada caso ya que en este proceso existe la posibilidad de que un determinado material sea digitalizado exclusivamente para la realización de un catálogo automatizado que permita acceder al documento a través de una pantalla, o la realización de una reproducción lo más fiel del documento original para la formación de un archivo de respaldo con el fin de conservar y difundir el material a través de medios impresos en el futuro.

Por lo tanto, las condiciones de la digitalización se determinarán considerando los siguientes factores:

- Valor patrimonial del material análogo.
- Estado de conservación del material original.
- Demanda de uso del material análogo y del recurso digital.

Valor patrimonial del material análogo

La Biblioteca Nacional custodia materiales en diversos soportes y formatos. Algunos de ellos, dadas su originalidad, condición técnica/material o por su valor histórico o estético, han adquirido una especial relevancia que los destaca por sobre los demás y que nos obliga a agotar esfuerzos en función de transmitirlos a las futuras generaciones en las mejores condiciones de conservación que se encuentren a nuestro alcance. Estos son los materiales que denominamos "objetos de alto valor patrimonial". En estos materiales los procesos de digitalización y los peligros que éstos encierran (traslados, manipulación, exposición a la luz, etc.) deben reducirse al mínimo posible; siempre teniendo en consideración que, si bien existe un deterioro

inevitable en el proceso, el objeto reproducido en formato digital evitará una serie de riesgos mayores al material original en el futuro.

Con la finalidad de evitar el problema y el daño que se les pudiera producir a los **documentos de alto valor patrimonial** por el hecho de someterlos al proceso de digitalización en más de una oportunidad, se ha considerado que este tipo de materiales deberá ser digitalizado en alta resolución de manera de conformar un archivo de respaldo que pueda ser utilizado para satisfacer cualquier tipo de necesidad de utilización futura (publicación electrónica o impresa en cualquier soporte y formato). A partir de este archivo digital de alta resolución (*master*) se realizarán las copias necesarias para su utilización.

Por su parte, **los materiales de bajo valor patrimonial**, exceptuando aquellos que coinciden con alguno de los otros dos factores que podrían determinar otras condiciones para la digitalización, deberán ser digitalizados en baja resolución, ya que de ser necesario en el futuro podrán ser digitalizados nuevamente sin riesgo de pérdidas graves o irreparables para el patrimonio de la Biblioteca Nacional y evitando el alto costo de captura, procesamiento y almacenaje que involucran la digitalización en alta resolución.

Estado de conservación del material original

La digitalización es un proceso de cambio de formato orientado al respaldo de originales y a facilitar el acceso. No debe considerarse en ninguno de los casos como una solución definitiva a los problemas de conservación de los materiales originales presentes en las colecciones. Una buena digitalización puede evitar futuros deterioros causados por la manipulación excesiva de los originales, pero no garantizará su permanencia definitiva en el tiempo. Las colecciones de la Biblioteca Nacional poseen materiales que por su excesivo uso o malas condiciones de almacenamiento se encuentran en avanzado estado de deterioro. Estos documentos al ser expuestos a los procesos que involucran una conversión a formato electrónico (traslado, manipulación, luz, etc.) experimentarán en cierta medida un trastorno en su condición actual y un incremento en su estado de deterioro que deberá disminuirse al máximo posible.

Se ha considerado que los **materiales en mal estado de conservación**, con la finalidad de evitar el incremento de los deterioros que se presentan por el hecho de ser expuestos a factores y agentes de deterioro en forma repetida, deberán ser digitalizados a alta resolución de manera de conformar un archivo de respaldo que pueda ser utilizado para satisfacer cualquier tipo de necesidad de utilización futura (publicación electrónica o impresa en cualquier soporte y formato) y, a partir de este archivo digital de alta resolución (*master*) realizar las copias necesarias para su utilización.

Por su parte, los **materiales en buen estado de conservación**, exceptuando aquellos que presenten alguno de los otros dos factores (alto valor patrimonial o gran demanda de uso) que determinarán otras condiciones para la digitalización, deberán ser digitalizados en baja resolución, ya que de ser necesario en el futuro, podrán ser digitalizados nuevamente sin riesgo de incrementar deterioros graves o irreparables en los materiales de las colecciones y evitando los altos costos de captura, procesamiento y almacenaje que involucran la digitalización en alta resolución.

Demanda de uso del material análogo y del recurso digital

La Biblioteca Nacional es uno de los principales centros de acopio de materiales bibliográficos del país. Existe al interior de sus colecciones una gran cantidad de documentos de diferente naturaleza y contenido. Algunos de ellos se caracterizan por su alta demanda tanto por parte de usuarios externos como de la propia institución para la realización de investigaciones, publicaciones, exposiciones etc. Por el contrario, existen otros casos en que los documentos son de muy poca consulta y utilización. Este factor, de no ser considerado al momento de la digitalización, podría traducirse en el hecho de tener que repetir los procesos con las diferentes necesidades para cada caso y por lo tanto en una duplicación de trabajo, esfuerzo y recursos disponibles. Por otro lado, la digitalización con características de archivo *master* de documentos de baja demanda implicaría la inversión de un alto costo de captura y almacenaje de materiales que no serían de una utilidad mayor.

Para satisfacer las necesidades de los **materiales que presentan gran demanda** de uso se deberá plantear el proceso de digitalización orientado a la estrategia de conformar un archivo de respaldo de los materiales originales que pueda satisfacer cualquier tipo de necesidad de utilización futura (publicación electrónica o impresa en cualquier soporte y formato). A partir de este archivo digital de alta resolución (*master*) se realizarán las copias necesarias para su utilización.

Debido a los grandes tamaños de los archivos digitales y el costo de almacenaje que esto significa, se ha establecido que los objetos que presentan una baja demanda de utilización y que no cumplan con ninguno de los otros factores que podrían cambiar las condiciones de la digitalización se deberán digitalizar dependiendo de las necesidades puntuales que se presenten en cada caso, pudiendo ser eliminados después de haber cumplido el uso para el cual fueron digitalizados.

Considerando la condición análoga de los materiales y cada uno de los factores anteriormente descritos se ha podido establecer una política estandarizada de digitalización para las colecciones bibliográficas de la Biblioteca Nacional de Chile. Esta política será utilizada en todos los proyectos de digitalización emprendidos por la institución de manera de garantizar los estándares de calidad en los resultados y en los productos obtenidos.

RESULTADOS

Los resultados obtenidos a partir de la metodología empleada en los procesos de digitalización son los siguientes:

Materiales de tono continuo

- a) Materiales de alto valor patrimonial y/o mal estado de conservación y/o alta demanda de uso.

Digitalización con calidad de archivo (*master*): escala 1:1, resolución de captura 600 dpi, formato TIFF (Tagged Image File Format) sin compresión, color o escala de grises según sea la necesidad del caso.

- b) Materiales de bajo valor patrimonial, buen estado de conservación y baja demanda de uso.

Digitalización con calidad de copia de publicación se ajusta a las necesidades de cada caso.

Calidad de impresión: escala 1:1, resolución de captura 300 dpi, formato TIFF (Tagged Image File Format) sin compresión, color o escala de grises según sea la necesidad del caso.

Catálogo automatizado o web: resolución de 150 dpi con dimensión espacial de 750 x 550 píxeles en formato JPEG (Formato de protocolo estándar de compresión por interpolación) en color, escala de grises o bitonal según sean las necesidades de cada caso.



Foto 3. La Cruz de Reyes. Copia de papel albuminado. Album Fotográfico "Vistas de Valparaíso" de Felix Leblanc. C.1890. Colección Sala Medina Biblioteca Nacional. Fotografía digitalizada para copia de respaldo master.



Foto 4. Biblioteca Luis Montt, acceso calle San Ignacio. Fotografía en color. Fotografía Roberto Aguirre. 2003. Fotografía digitalizada para ser publicada en artículo, eliminando posteriormente el archivo digital.



Foto 5. Modo de Ganar el Jubileo Santo. Primer impreso chileno. 1776. Colección Biblioteca Nacional. Por su alto valor patrimonial, la obra fue digitalizada para la elaboración de copia de respaldo master.

Imágenes de trama

- a) Materiales de alto valor patrimonial y/o mal estado de conservación y/o alta demanda de uso.

Digitalización con calidad de archivo (*master*): escala 1:1, resolución de captura 300 dpi formato TIFF (Tagged Image File Format) sin compresión, color, escala de grises o bitonal según sea la necesidad del caso.

- b) Materiales de bajo valor patrimonial, buen estado de conservación y baja demanda de uso.

Digitalización con calidad de copia de publicación, se ajusta a las necesidades de cada caso.

Calidad de impresión: escala 1:1, resolución de captura 300 dpi, formato TIFF (Tagged Image File Format) sin compresión, color o escala de grises según sea la necesidad del caso.

Catálogo automatizado o web: resolución de 150 dpi con dimensión espacial de 750 x 550 pixeles en formato JPEG (Formato de protocolo estándar de compresión por interpolación) en color, escala de grises o bitonal según sean las necesidades de cada caso.

CUADRO

| Tipo de Imagen | Factores | Resolución | Tamaño | Formato |
|----------------|---|------------------------------------|--------------------------|-----------------------------|
| Tono continuo | Alto valor patrimonial. Mal estado de conservación. Alta demanda de uso. | 600 dpi | 1:1 | TIFF sin compresión |
| Tono continuo | Bajo valor patrimonial. Buen estado de conservación. Baja demanda de uso. | Impresión: 300 dpi Web: 150 dpi | 1:1 750 a 550 pixeles | TIFF sin compresión JPEG |
| Trama | Alto valor patrimonial. Mal estado de conservación. Alta demanda de uso. | 300 dpi | 1:1 | TIFF sin compresión |
| Trama | Bajo valor patrimonial. Buen estado de conservación. Baja demanda de uso | Impresión: 300 dpi Web: 150 dpi | 1:1 750 a 550 pixeles | TIFF sin compresión JPEG |

CONCLUSIONES

En la actualidad, ya no cabe duda que la irrupción de las nuevas tecnologías en la llamada sociedad de la información es una realidad cercana y creciente en el mundo de las bibliotecas, los archivos y los museos. Cada vez son más las instituciones que, movidas por el ánimo de preservar y difundir su patrimonio, están buscando en la captura digital una herramienta que les permita, por una parte, mantener la información a largo plazo y, por otra, posibilitar el acceso a ella desde lugares remotos. Hoy es frecuente encontrar en nuestras bibliotecas o archivos, además de los servicios tradicionales, catálogos en línea, salas de consulta de Internet, consulta de colecciones en redes locales, servicios de reproducción digital de colecciones y archivos de datos o información digital. Para cada uno de estos procesos la digitalización de documentos juega un papel preponderante.

La estandarización de los procesos de digitalización es una necesidad relevante en los flujos de trabajo planteados para la realización de proyectos o programas que contemplen la conversión digital de documentos en las instituciones de custodia de materiales patrimoniales, los que deben ser contemplados en cada uno de los pasos establecidos incluyendo la captura digital.

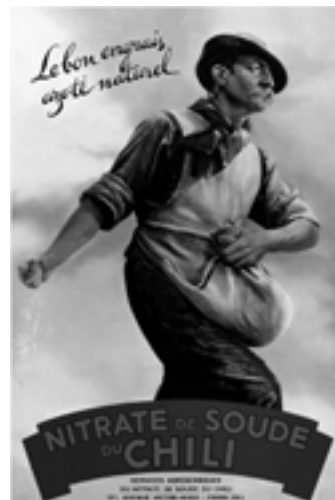
Muchos proyectos de digitalización que no han contemplado normalizar los procesos de captura han debido deshacer trabajos ya realizados con el consiguiente problema de duplicación de esfuerzos y recursos que ello implica. Por otra parte, el uso de formatos no estandarizados o de los llamados “propietarios” puede dificultar la recuperación de la información en el futuro producto de la obsolescencia de formatos. Siempre se debe estar atento a los movimientos del mercado tecnológico.

La Biblioteca Nacional de Chile, como una de las instituciones pioneras en el desarrollo de proyectos de digitalización de colecciones patrimoniales en el país, está llamada a tomar un liderazgo activo en el desarrollo de cada una de las tareas que se incluyen dentro de este proceso, formulando políticas y recomendaciones de normalización en los criterios de selección, captura, catalogación, almacenamiento y recuperación de documentos electrónicos orientados a su preservación a largo plazo.

BIBLIOGRAFÍA

CONWAY, P. *La Preservación en el Mundo Digital*. Santiago, Chile: Centro Nacional de Conservación y Restauración, 2000. 36 p.

HAZEN, D.; HORRELL, J. AND MERRILL-OLDHAM, H. *Cómo seleccionar colecciones de investigación para la digitalización*. Santiago, Chile: Centro Nacional de Conservación y Restauración, 2000. 31 p.



Fotos 6 y 7. Afiche de Francia. En *Una Experiencia fértil: la propaganda salitrera chilena en el exterior*. Colección Archivo Nacional. 2001. Imagen digitalizada con calidad de impresión.



Imagen digitalizada con calidad de publicación en web.

Technical Guidelines for Digitizing Archival Materials for Electronic Access: Creation of Production Master Files – Raster Images. U.S. National Archives and Records Administration (NARA), 2004. 87 p.
<http://www.archives.gov/research/arc/digitizing-archival-materials.pdf> (abril, 2006).

BIBLIOGRAFIA COMPLEMENTARIA

- Best Practices for Image Capture*. California Digital Library. 1999.
<http://www.gseis.ucla.edu/~howard/MOA2/bp90.doc> (abril, 2006)
- FLEISCHHAUER, C. *Digital Formats for Content Reproductions*. Washington D.C., U.S.A.: National Digital Library Program, Library of Congress, 1998.
<http://memory.loc.gov/ammem/formats.html> (abril, 2006)
- FREY, F. AND REILLY, J. *Digital Imaging for Photographic Collections : Foundations for Technical Standards*. Rochester, U.S.A.: Image Permanence Institute Rochester, 1999. 52 p.
- KLIJN, E. AND DE LUSENET, Y. *In the Picture Preservation and digitisation of European photographic collections*. Amsterdam, Holland: Commission on Preservation and Access, 2000. 69 p.
- KENNEY, A. AND RIEGER, O. *Moving Theory into Practice*. Society of American Archivists, 2001. 189 p.
- SERENSON et al. *Guides to Quality in Visual Resource Imaging*. Digital Library Federation, 2000.
<http://www.rlg.org/legacy/visguides/> (abril, 2006).
- Western States Digital Imaging Best Practices. Western States Digital Standards Group, Digital Imaging Working Group, Denver, Co. U:S:A: University of Denver and the Colorado Digitization Program, 2003. 41p.
http://www.cdphheritage.org/digital/scanning/documents/WSDIBP_v1.pdf (abril, 2006).