

Una experiencia de aplicación de tecnología digital de imágenes en la preservación y difusión de patrimonio histórico

Fernando J. Boro
Juan José Santos
Sandra G. Sauro

Instituto de Historia Argentina y Americana "Dr. Emilio Ravignani"
Programa de recuperación y preservación de Patrimonio Histórico
Facultad de Filosofía y Letras - Universidad de Buenos Aires

25 de mayo 217, 2° piso - Capital Federal - (1002)
T.E.: 334-7512 E-mail: rtravig@criba.edu.ar

La presente ponencia tiene por objeto destacar las posibilidades que las nuevas técnicas de procesamiento digital de imágenes ofrecen en el área de la conservación y divulgación del patrimonio histórico y cultural. Se apoya en la experiencia recogida como integrantes del Programa de recuperación y preservación de Patrimonio Histórico, actividad patrocinada por la Fundación Antorchas y la empresa IBM Argentina, con sede en el Instituto de Historia Argentina y Americana "Dr. Emilio Ravignani" de la Facultad de Filosofía y Letras de la UBA, bajo la dirección del Profesor José Carlos Chiaramonte.

Una breve mirada al problema de la preservación de los documentos históricos

La preservación de los documentos históricos, ya sean manuscritos o impresos de época, ha sido siempre una preocupación de las instituciones vinculadas al quehacer histórico y cultural. Tanto los institutos de investigación como las bibliotecas y archivos, oficiales o privados, que conservan este importante patrimonio, se han enfrentado a lo largo de su existencia con la necesidad de encontrar maneras de asegurar que las fuentes originales, a partir de las cuales se elabora la investigación, enseñanza y difusión de la historia, se conserven en el mejor estado posible. No es éste el lugar para un análisis detallado de los diversos factores de riesgo que afectan el estado de los archivos históricos, pero valga señalar que su conservación depende de la aplicación simultánea de varios saberes especializados, aunando a los archivistas, especialistas en conservación de papel, historiadores, bibliotecarios, especialistas en fotografía antigua, etc.

No obstante lo señalado, existe una situación particular que merece ser destacada. Dado que esos originales son, en su gran mayoría, papeles muy antiguos, y que por otro lado, la elaboración de una investigación histórica implica inevitablemente una fase de consulta de los mismos, existe desde luego una tensión entre la situación ideal de conservación -evitar su manipulación- y la labor investigativa que necesita manipularlos. Al mismo tiempo es obvio que no tendría sentido impedir lisa y llanamente su consulta so pretexto de evitar su degradación, ya que uno de los motivos primordiales de garantizar una adecuada conservación no es otro que permitir la utilización de los archivos para la investigación histórica presente y futura, la que luego se plasmará en artículos y libros que hacen avanzar el conocimiento de nuestro pasado. Cualquier persona que haya visitado un archivo histórico sabe que estas instituciones imponen una serie de cuidados para la consulta de los materiales originales; por dar un ejemplo, en algunos casos se prohíbe fotocopiar la documentación. Estas medidas se incrementan en forma proporcional al mal estado de los papeles, llegando en ciertos casos extremos a prohibirse directamente su consulta.

Por estas razones, una de las formas más eficaces para garantizar al mismo tiempo la conservación de las fuentes -más allá de las condiciones químicas y ambientales adecuadas de almacenamiento- y su consulta por los investigadores consiste en la reproducción del documento. De manera esquemática, pueden reconocerse tres etapas en la evolución de las técnicas empleadas para la reproducción documental, que a su vez se distinguen en cuanto al objeto que se busca reproducir, bien sea el contenido o la imagen del documento:

a) en primer lugar, y caracterizado por utilizar al papel como soporte, la reproducción del contenido textual del documento, como es el caso de la copia manuscrita, mecanografiada o bien de

la edición impresa. Los orígenes de este método son muy antiguos, y en lo referente a la copia manuscrita se siguió utilizando aproximadamente hasta el primer tercio de nuestro siglo;

b) en segundo lugar, la microfilmación, que recupera la imagen fotográfica y es el método más extendido hasta nuestros días;

c) por último, el método más reciente y aún no muy difundido que consiste en la captura digital de la imagen y su inclusión en una base de datos.

Los dos últimos ya no utilizan al papel como soporte y producen una copia fiel del original, lo que si bien es una gran ventaja plantea a su vez el problema de la conservación y la durabilidad de sus respectivos soportes: la película revelada y el disco óptico -u otro medio de almacenamiento apto para información digital.

Históricamente, el recurso de la reproducción documental se aplicó de diversas formas, asociadas con las posibilidades tecnológicas de cada época, y no siempre estuvo vinculado a un objetivo de preservación en sentido estricto. Durante el tiempo en que sólo era posible la copia en soporte papel, la reproducción no podía emplearse de manera masiva, ya que en un archivo con cientos de miles o incluso millones de originales, reproducir toda la documentación duplicando la masa de papeles parecía una tarea sin sentido. Como veremos, las tecnologías de reproducción de la imagen ofrecieron soluciones más satisfactorias a este problema.

El Instituto Ravnani y la protección de archivos históricos

El Instituto Ravnani, desde sus inicios en 1921, se abocó a la recolección, catalogación y edición de colecciones documentales valiosas para la historia argentina y americana. En su larga trayectoria en el manejo de archivos pasó por cada una de las etapas ya mencionadas, desde las copias manuscritas, que permitieron acceder a documentación de valor para nuestra historia que se conservaba en repositorios europeos y americanos, hasta la edición de fuentes en forma de libro. Más adelante, y siguiendo una tendencia generalizada, se adoptó la microfilmación como medio de conservación y consulta de las fuentes históricas. Las ventajas que esta última presentaba respecto de los métodos anteriores consistían, por un lado, en la obtención de una reproducción fiel de la imagen documental -no sólo de su contenido-; por otro, en la facilidad de almacenamiento en un espacio reducido por comparación con los volúmenes en papel. Asimismo, los microfílm podían ser reproducidos a bajo costo, lo que permitía obtener copias de archivos de otro modo difícilmente accesibles. Por dar sólo un ejemplo de lo dicho podemos mencionar la recuperación de las imágenes del Archivo Pedro De Angelis -una colección documental valiosa para la historia del Río de la Plata en los siglos XVIII y XIX-, cuyos originales se conservan en la Biblioteca Nacional de Río de Janeiro, por razones que no es del caso exponer aquí, gracias a la obtención de sus copias en soporte microfilm que ahora están disponibles para la consulta en el Archivo Documental del Instituto Ravnani.

Asimismo, el Instituto Ravnani tuvo a su cargo las tareas de ordenamiento y catalogación del Archivo del Brigadier General Don Juan Facundo Quiroga y del Archivo del Doctor Rufino de Elizalde (quien fuera Ministro de RREE de la Argentina en la presidencia de Bartolomé Mitre), que permitieron la microfilmación de ambos fondos en el marco de un convenio con la Biblioteca del Honorable Congreso de la Nación y del Archivo General de la Nación. El buen resultado de esta labor hizo posible la apertura de esos archivos a la consulta pública.

El Programa de recuperación y preservación de Patrimonio Histórico

Aprovechando las experiencias realizadas en la conservación de archivos, en el año 1993 se proyectó la creación de este Programa, orientado a la preservación de documentación en riesgo. Desde ese momento, y con la inquietud de actualizar nuestros métodos de trabajo, nos abocamos a evaluar la posibilidad de incorporar la nueva tecnología de digitalización de imágenes. En particular, estábamos interesados en agilizar el proceso de consulta e impresión de las fuentes históricas, ya que como usuarios de bases de datos informáticas conocíamos la potencia de esos programas en operaciones de búsqueda con criterios simples o combinados, algo que los microfílm no permitían hacer. En éstos, la búsqueda de un documento específico en un rollo debe hacerse de manera

secuencial, vale decir, para llegar al fotograma buscado es imprescindible recorrer todos los anteriores, lo que implica una demora en la consulta del mismo. Además, la "hoja de ruta" de la microfilmación sigue el orden real de los documentos tal como aparecen en las carpetas, generalmente basado en un criterio cronológico, de modo que la realización de una sola consulta habitualmente implica recorrer varios rollos. Por último, si bien la mayoría de las bibliotecas y repositorios pueden contar con un lector de microfilms, no ocurre lo mismo con los lectoimpresores, por lo cual no es tan sencillo producir una impresión de las imágenes contenidas en los microfilms.

El problema de esa evaluación era que no había en nuestro medio una experiencia similar que permitiera comparar y analizar las ventajas y dificultades que podrían presentarse. No ignorábamos la existencia de un proyecto de digitalización de documentación histórica en el Archivo de Indias de Sevilla (España), pero éste, por la magnitud de los recursos puestos a su disposición, contaba con la posibilidad de programar sus propias aplicaciones y formatos gráficos, de modo que no podíamos tomarlo como un punto de referencia en sentido estricto, aunque lo reconocemos como el primer antecedente en cuanto a la aplicación de esta tecnología en archivos históricos en gran escala. Debe tenerse en cuenta que no bastaba con saber que la tecnología de captura digital de imágenes ya se encontraba en franco desarrollo, sino que había que considerar que debía procesarse documentación antigua en condiciones irregulares de legibilidad, a veces con tintas desvaídas y con grados de contraste muy diversos. Dado que los programas de base de datos orientada a imágenes estaban diseñados para capturar la documentación producida por empresas en fecha reciente, la preocupación por definir estos puntos no era un simple prurito de especialistas en historia. Además, si la consulta por pantalla iba a reemplazar al lector de microfilm, debíamos asegurarnos de que la visualización de la imagen digital tuviera la suficiente calidad como para permitir la lectura cómoda del documento. Asimismo, para el caso de originales con deficiente grado de legibilidad, era necesario que el sistema permitiera la ampliación de la imagen sin pérdida de nitidez, esto es, sin la aparición del efecto de "pixelado". Otro punto a evaluar era la velocidad real de presentación de los documentos en pantalla, considerando el tamaño de los archivos de imágenes. Finalmente, dada la importancia que tiene en nuestra actividad la posibilidad de intercambio entre instituciones de los archivos procesados, se requería que el sistema utilizara un formato gráfico que no impidiera su reconocimiento y lectura por otras aplicaciones informáticas. De este modo se dejaba abierta la posibilidad de exportar imágenes a soporte CD-Rom u otro medio de almacenamiento apto para archivos de gran tamaño.

Finalmente, luego de muchas consultas y demostraciones, se decidió montar una red de computadoras IBM configurada para la captura, catalogación, consulta e impresión de imágenes de documentos históricos. La misma está compuesta por una estación de captura provista de un scanner por tonos de grises que permite procesar originales de hasta 46 por 32 cm., a una resolución de hasta 400 dpi, equipada con un monitor de 19" con resolución de 1600 por 1300 puntos; una de consulta con monitor de 21" color en alta resolución y otra estación para impresión en láser. Todas las estaciones permiten indexar la documentación, hacer búsquedas, visualizar en pantalla y solicitar impresiones. La red cuenta con un equipo lectograbador de discos magneto ópticos -con 650 Mb de capacidad por disco- para el almacenamiento de las imágenes. El software está integrado por el sistema operativo OS2 en las estaciones de trabajo, y OS2 Lan Server para los servicios de la red. La base de datos orientada a imágenes es VisualInfo apoyada por DB2. Esto fue posible gracias al apoyo financiero brindado por la Fundación Antorchas, que permitió adquirir dos de las tres estaciones de trabajo y todos los equipos periféricos. Asimismo, IBM Argentina instaló la red gratuitamente y donó una estación de trabajo más todo el software en uso. Hemos podido comprobar que se han cumplido los requisitos que imponía una actividad tan específica: la búsqueda de las imágenes puede hacerse por criterios selectivos, sin importar el orden de ingreso en la base de datos, y dependiendo de los campos predefinidos para la búsqueda. La visualización es excelente, pudiendo ampliarse varias veces las imágenes que así lo requieren. Desde luego, la calidad de la imagen depende de la resolución con que se la ha capturado, que en nuestro caso promedia los 200 dpi en blanco y negro. En promedio, el peso de cada archivo de imagen, que es automáticamente comprimido por el VisualInfo, oscila alrededor de 1 megabyte, considerando que en un mismo archivo es común que se graben varias imágenes -una carta, por ejemplo, puede tener varios folios-. Frente a originales con problemas de legibilidad capturamos en 16 tonos de grises, lo que suele ser suficiente para una correcta visualización. Trabajando con estos valores se obtienen muy buenas impresiones láser. Respecto de la velocidad del sistema, aunque muy superior por comparación al microfilm, aún no es la óptima. Sin embargo IBM ha donado recientemente un nuevo server muy potente que solucionará este punto. La reducción del tiempo de acceso a las imágenes también será

posible gracias a una migración de todo el software a sus versiones más recientes, que los técnicos de IBM están por comenzar.

El empleo de esta tecnología nos llevó a revisar los criterios clásicos del ordenamiento archivístico. Por ejemplo, es una práctica muy extendida en los archivos históricos catalogar los documentos según el archivo del cual proceden, comúnmente llamado archivo de origen. Pero a poco de andar, advertimos que este método producía una innecesaria complicación en la estructura de la base de datos, lo que llevaba a una pérdida de eficacia en las operaciones de búsqueda. Por ese motivo decidimos indexar la documentación de acuerdo con un criterio formal, definido en base a la clase de documento. Así, definimos varias clases o tipos de documentos, tales como correspondencia, impresos, mapas, etc. Cada clase, a su vez, tiene definidos campos específicos de acuerdo con su tipología y se comporta, de hecho, como una base de datos diferenciada. Por otro lado, la presencia de campos de búsqueda compartidos por cada una de las clases, permite al usuario efectuar consultas que involucren a todas las clases simultáneamente. Este modo de clasificar la información se reveló mucho más eficaz en el aprovechamiento de las posibilidades de búsqueda dentro de la base de datos, y no significó perder la vinculación con el campo archivo de origen, con lo cual es posible recuperar, en caso necesario, la unidad originaria del corpus documental.

Así, los resultados de búsqueda ofrecidos por la base de datos responden tanto a las necesidades más generales como a las más específicas de cada investigador, obteniéndose distintos listados de documentos según se hayan solicitado total o parcialmente los campos previstos. Por ejemplo, un investigador podría buscar correspondencia emitida desde Buenos a Salta, desde el año 1809 hasta 1816, en la que figure Cornelio Saavedra como remitente o destinatario, o bien pedir todos los documentos en donde se registre el nombre de Cornelio Saavedra a través de todas las clases definidas en la base de datos. De este modo el usuario puede recuperar y visualizar en pantalla los documentos que resulten de su consulta, así como pedir una impresión desde su estación de trabajo a la impresora láser a través de la red.

El proceso de trabajo

Previo a cualquier procesamiento digital, la documentación debe ser ordenada y catalogada según criterios cronológicos y temáticos, tarea que culmina con la edición impresa de su catálogo. La digitalización representa la primera etapa del trabajo informático. De acuerdo con el estado del original, se captura la imagen del documento mediante un escaner y, si es necesario editarla, se vuelve a escanear con la corrección adecuada en los valores de contraste, brillo, resolución, etc. El programa VisualInfo, por estar diseñado para la captura de la documentación legal de empresas, se basa en el supuesto de que el usuario no desea modificar la imagen original, por lo que no permite manipular las imágenes como lo hacen las aplicaciones de edición gráfica. Una vez obtenida una imagen apropiada, se la graba en una bandeja de entrada para su posterior indexación. Finalizada la indexación, las imágenes migran automáticamente al disco magneto óptico para su almacenamiento.

Los archivos bajo tratamiento

En una primera etapa estamos digitalizando una selección de los fondos documentales originales que se conservan en el Archivo Documental del Instituto Ravignani. Esta labor inicial nos permite familiarizarnos con el funcionamiento del sistema y poner a prueba la validez de la estructura de la base de datos tal como fue definida. Estos documentos conforman un corpus de valor para la historia argentina, y son los siguientes:

- Carpetas con documentación diversa -correspondencia de gobernadores y figuras políticas, impresos, proclamas, etc.- correspondiente a los siglos XVIII y XIX
- Archivo de Vicente Anastasio Echevarría (1749-1877) □
- Documentos relativos a la Batalla de Pavón
- Archivo de Miguel de Otero (1846-1848)

En una segunda etapa se trabajará con fondos documentales en manos privadas, en condiciones que no aseguren una adecuada conservación, y cuyo valor se estime de interés para la investigación. Esto significa desarrollar una labor de búsqueda en el largo plazo. En este sentido estamos tomando

contacto con historiadores del interior del país para reunir información sobre la existencia y localización de archivos que respondan a esas características.

Dado que no todos los poseedores de documentación accederán a cederla temporariamente para su procesamiento en sede, desde el inicio tuvimos la inquietud de implementar algún sistema portátil para digitalizar documentos. Desechadas, luego de varias pruebas, las alternativas de captura por señal de videocámara así como por fotografía digital, por la insuficiente resolución con que producen unas imágenes destinadas a ser leídas en pantalla, estamos evaluando la factibilidad de utilizar una notebook de última generación, un escaner de página completa y una unidad de discos ópticos, con 1 Gb de capacidad por disco, para el almacenamiento de imágenes. En ese caso, las imágenes se grabarían en formato TIFF, y una vez en sede se aplicaría la utilidad de importación, con lo cual los documentos ingresarían al sistema en red para su posterior indexación.

Finalmente, quisiéramos mencionar que por haber dado el paso de adoptar la tecnología digital para el tratamiento de archivos, están dadas las condiciones para producir, en el futuro, colecciones documentales digitales en soporte CD-Rom, lo que permitiría hacer circular ese patrimonio entre instituciones e investigadores de la Argentina y del exterior. Por ahora, razones de orden técnico aconsejan esperar que las nuevas versiones de los programas que permiten grabar bases de datos en CD-Rom sean capaces de automatizar la exportación de los archivos de imágenes con sus índices asociados. Hasta el momento, nuestra exploración al respecto arrojó resultados negativos.

La imagen que aparece al final del trabajo procede de este fondo documental. Se trata de una carta de Don Juan de la Cruz Monje y Ortega dirigida a Vicente A. Echeverría, fechada en Salta el 20 de agosto de 1810.