MANUAL PARA PROYECTOS DIGITALES:

Una herramienta de gestión para preservación y acceso

Primera Edición

Maxine K. Sitts, Editora

Northeast Document Conservation Center

Andover, Massachusetts, 2000

El Instituto de Servicios para Bibliotecas y Museos, agencia federal que promueve la innovación, la creación y el conocimiento, dio su apoyo al Northeast Document Conservation Center para la publicación de este libro, Manual para proyectos digitales: Una herramienta de trabajo para preservación y acceso.

El National Endowment for the Humanities, agencia del gobierno federal independiente para la concesión de fondos, proporciona financiamiento sustancial para el apoyo de actividades de terreno, entre las que se incluyen las publicaciones en el Northeast Document Conservation Center.

Las opiniones expresadas por los autores individuales en este libro son opiniones personales, y no representan las opiniones de institución u organización alguna, ni siquiera la del NEDCC. El NEDCC no garantiza que la información que contiene este volumen

está completa y es exacta, y no asume, y por este medio rechaza cualquier responsabilidad por pérdidas o daños causados por errores u omisiones en el *Manual para proyectos digitales: Una herramienta para preservación y acceso*, ya sea que dichos errores u omisiones sean resultado de alguna negligencia, accidente, u otra causa.

Número de Catalogación de la Biblioteca del Congreso
ISBN No. 0-9634685-4-5

Copyright © 2000 del Northeast Document Conservation Center. Todos los derechos reservados.

Ninguna parte de esta publicación se puede reproducir o transmitir en forma o soporte, ni almacenarse en soporte alguno o en sistema alguno de acceso sin previo permiso del Northeast Document Conservation Center, 100 Brickstone Square, Andover, MA 01810.

Esta publicación está impresa en un papel que cumple los requisitos de la Norma Nacional de Estados Unidos para la permanencia del papel y las ciencias de la información de materiales impresos para bibliotecas, ANSI Z39.48-1992 (ratificada en 1997).

Manual para proyectos digitales:
Una herramienta de gestión para preservación y acceso

Índice

PREFACIO

I. INTRODUCCIÓN

Las Conferencias Escuela de escaneo

Propósitos

Visión general del contenido

II. VISIÓN GENERAL: EL RAZONAMIENTO NECESARIO CUANDO SE PIENSA EN UN PROYECTO DE DIGITALIZACIÓN Y PRESERVACIÓN

Paul Conway, Yale University Library

Introducción

Las ventajas del acceso digital

Los riesgos de un proyecto de digitalización de imágenes

Lo que no es la digitalización de imágenes

La preservación en el mundo digital

Las relaciones que existen entre proceso, fuente y tecnología

La transformación de los principios de preservación

Conclusión

Resumen de los principios y los aspectos claves

Bibliografía

III. CONSIDERACIONES PARA LA GESTIÓN DE UN PROYECTO

Stephen Chapman, Harvard University Library

Trazarse objetivos

La planificación del proyecto: crearse un plan de trabajo y un presupuesto

La ejecución del proyecto: La gestión del flujo de trabajo y las fuentes

Bibliografía

IV. SELECCIÓN DE LOS MATERIALES QUE SE VAN A ESCANEAR

Diane Vogt-O'Connor

Introducción

¿Por qué hay que seleccionar los materiales para escanearlos?

El comité de selección

El proceso de selección

Cómo se unen todos los elementos: Cómo evaluar y priorizar las colecciones

Resumen de los aspectos claves

Bibliografía

V. VISIÓN GENERAL DE LOS ASPECTOS DEL DERECHO DE AUTOR

Melissa Smith Levine, National Digital Library Project

Introducción

Derecho de autor (Copyright)

¿Qué es el uso justo?

Otras consideraciones

Cómo actuar

Referencias para obtener información acerca de cuestiones legales en el Mundo en Línea

Bibliografía

VI. LAS PRIMERAS COSAS TÉCNICAS

Stephen Puglia, National Archives and Records Administration

Introducción

La imagen digital

Medidas básicas de la imagen

Procesamiento de la imagen digital

Comparación de la reconversión de formato

Bibliografía

VII. DESARROLLO DE LAS MEJORES PRÁCTICAS: LINEAMIENTOS A PARTIR DE ESTUDIOS DE CASOS

Introducción

1. Cómo trabajar con textos impresos y con manuscritos

Stephen Chapman, Harvard University Library

2. El trabajo con fotografías

Franziska Frey, Image Permanence Institute

3. Un estudio de caso de OCR

Eileen Gifford Fenton, JSTOR, University of Michigan

4. La digitalización de mapas y otros documentos de gran tamaño

Janet Gertz, Columbia University Libraries

5. El trabajo con microfilm

Paul Conway, Yale University Library

6. Cooperación en la digitalización de imágenes: Cómo escanear bien con otros colegas

Steven D. Smith, Amigos Library Services, Inc.

VIII. RELACIONES CON LOS PROVEEDORES DE SERVICIOS

Janet Gertz, Columbia University Libraries

Introducción

Por qué digitalizar en nuestra propia institución

Por qué utilizar a los proveedores

Cómo escoger los servicios y los proveedores

¿Cuáles son los objetivos del proyecto?

La SDI, Solicitud de información

La solicitud de propuesta (SDP)

Cómo evaluar la respuesta de los proveedores

El contrato

Cómo trabajar y comunicarse con los proveedores

El trabajo con los proveedores: Control de la calidad y correcciones en la manipulación

Bibliografía

IX. LONGEVIDAD DIGITAL

Howard Besser, University of California, Los Angeles School of Education & Information Studies

La corta vida de la información digital

El problema del visionaje

El problema de la interferencia

El problema de la inter-relación

El problema de la custodia

El problema de la traducción

Vías para alargar la longevidad digital

Bibliografía

X. COMENTARIO DE UN ESPECIALISTA: UN USUARIO EXPRESA SU OPINIÓN

Charles Rhyne, Reed College

La mayor parte de los materiales digitalizados no son válidos como evidencia

Muy pocos especialistas-profesores participan en el proceso

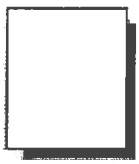
Los prototipos que se toman como patrón

Las imágenes fotográficas

Las normas tradicionales de las evidencias

Los supuestos defectos se pueden convertir en logros

Una sociedad que participa en pleno



Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

Manual para proyectos digitales: Una herramienta de gestión para preservación y acceso

Prefacio

El NEDCC se complace en presentar este manual a la comunidad de profesionales. Teniendo en cuenta que había muy poca literatura sobre este aspecto tan importante, el Northeast Document Conservation Center se dio a la tarea de desarrollar una guía para la gestión de proyectos de conversión a formato digital. El objetivo era producir un primer material de fácil uso que se concentrara en satisfacer las necesidades de información de bibliotecas, museos, archivos y otras instituciones que albergan valiosas colecciones. Este manual pretende servir como fuente y a la vez como respuesta a una necesidad. El mayor reto fue compilar y entregar la información más actualizada y útil lo más pronto posible. El proyecto tenía que realizarse dentro de un período de tiempo muy corto para poder brindar a nuestros lectores una información que fuera oportuna, aunque sabíamos que no iba a ser duradera.

La publicación se basa en la serie de exitosas conferencias patrocinadas por el NEDCC que se conocen con el nombre de "Escuela de Escaneo" ("School for Scanning"), que se han llevado a cabo en el territorio de Estados Unidos desde 1996 con el apoyo de la Fundación Mellon y del National Endowment for the Humanities (Fondo Nacional para las Humanísticas). El libro está dividido en diez capítulos cada uno de ellos escrito por uno o varios de los profesores de las conferencias y aborda temas paralelos a los temas discutidos en las mismas. Al igual que las conferencias, la publicación combina la tutoría sobre aspectos técnicos con una visión general sobre cuestiones más amplias, entre las que se incluye la necesidad de la preservación de los productos nacidos en formato digital. El libro comienza con la premisa de que la investigación acerca de la conversión a formato digital sólo tiene sentido si las instituciones están preparadas para proporcionar un acceso a largo plazo a las colecciones digitales.

El NEDCC espera que este manual sirva de ayuda a las instituciones para planificar proyectos basados en las consideraciones de calidad y acceso a través del tiempo. Quiero aprovechar la oportunidad para agradecer a Maxine Sitts, que ha sido una editora increíblemente eficiente; a Steve Dalton, Director de Servicios de Terreno del NEDCC, que sirvió como administrador del proyecto; y a Kim O'Leary que fungió como Webmaster. Quisiera agradecer en especial al Institute of Museum and Library Services (Instituto de Servicios a Bibliotecas y Museos), que apoyó la producción del manual a través de un fondo donado por el National Leadership Grant. Agradezco también el apoyo que la Oficina de Servicios de Terreno del NEDCC recibió del National Endowment for the Humanities.

Ann Russell
Directora Ejecutiva, NEDCC

Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

Manual para proyectos digitales: Una herramienta de gestión para preservación y acceso

I

Introducción

Este manual es el producto de cuatro años de trabajo de desarrollo y revisión del programa para las conferencias denominadas "Escuela de escaneo" ("School for Scanning") que el Northeast Document Conservation Center (NEDCC) ha llevado a cabo en todo el territorio de Estados Unidos. El manual ofrece un cúmulo de consejos, lineamientos y asesoramiento de instituciones que ya han realizado proyectos de digitalización de sus colecciones. En general, el manual recopila información acerca de las mejores prácticas en la digitalización y resume las lecciones aprendidas a partir de múltiples experiencias. El enfoque, aunque desde el punto de vista de la gestión, es también práctico y se basa en proyectos reales.

Las conferencias "Escuela de escaneo"

Las conferencias "Escuela de escaneo", financiadas por la Fundación Andrew W. Mellon y por el National Endowment for the Humanities (Fondo Nacional para las Humanísticas), han sido preparadas para administradores dentro de instituciones culturales, así como para bibliotecarios, archivistas, curadores y otros responsables y administradores de gestión de recursos culturales o naturales encargados de colecciones en soporte de papel (incluidas las fotografías). La conferencia, que se mantiene al día con los adelantos de la tecnología, ha preparado a los participantes para adoptar decisiones críticas con respecto a la gestión de proyectos digitales.

Las conferencias han sido todas un gran éxito. Cada una de ellas ha atraído a participantes muy capacitados, un promedio de 300 personas por evento. Un factor que ha propiciado la popularidad de las conferencias ha sido la continuidad de los profesores -- todos ellos profesionales activos de la

preservación y el acceso. Todos estos profesores, junto con el personal del NEDCC, han preparado este manual, a partir de sus propias presentaciones en las conferencias realizadas entre 1996 y el 2000. El diálogo continuo que se produce en estas conferencias ha proporcionado a los autores oportunidades únicas para actualizar y seleccionar la información en constante cambio acerca de la digitalización, basándose en las necesidades expresas de las instituciones cuyo primer objetivo es proteger, preservar y proporcionar acceso a los materiales y documentos que conforman nuestro patrimonio cultural e histórico.

Propósitos

Como sólo pocos cientos de individuos han podido asistir a cada una de las conferencias "Escuela de escaneo", cuestión que ha creado a su vez largas listas de espera para dichos eventos, el Institute of Museum and Library Services (Instituto de Servicios a Bibliotecas y Museos) accedió a apoyar al NEDCC en la producción de este libro para que llegue a un público más amplio.

Este manual sirve, por tanto, como herramienta de gestión para las instituciones interesadas en las cuestiones de preservación y acceso. Puede ayudar a los administradores y al personal en general a tomar mejores decisiones para:

- Determinar el momento y las circunstancias adecuadas para asumir un proyecto de digitalización
- Integrar las necesidades de preservación en proyectos de escaneo
- Seleccionar los materiales a escanear
- Trabajar con proveedores fuera de las instituciones
- Mantener el control de la calidad
- Desarrollar herramientas de indización y de navegación y preparar bases de datos
- Proporcionar acceso a redes.

El NEDCC espera que el manual resulte útil para una variedad de instituciones -- archivos, museos, sociedades históricas, y todo tipo de bibliotecas -- y que el mismo sirva para todos los niveles del personal. Esta versión impresa será particularmente útil a aquellos individuos en niveles introductorios e intermedios del aprendizaje, también para las sesiones de planificación en grupo y como referencia según se van planificando y

ejecutando los proyectos de digitalización. El manual también se encuentra en Internet, donde la información se actualizará con regularidad.

Los objetivos que se establecieron para el desarrollo de ambas versiones, la impresa y la electrónica son:

- Recopilar información sobre las mejores prácticas para la digitalización de colecciones retrospectivas según van surgiendo a partir del trabajo de varias grandes bibliotecas de investigación, y diseminar esta información a instituciones pequeñas y medianas
- Desarrollar una herramienta de enseñanza como apoyo a adiestramientos futuros
- Mantener una versión en Internet que sirva a un público aun más amplio y que sea de fácil búsqueda
- Balancear los intereses de acceso y los intereses de preservación.

Una visión general del contenido

Este manual . . .

* Interpreta la tecnología digital desde la perspectiva de las necesidades singulares de las instituciones encargadas de salvaguardar los tesoros culturales y de proporcionar acceso a los mismos
* Estudia cómo pueden justificar las instituciones la ejecución de proyectos de digitalización de imágenes
* Describe cómo hacer la gestión de proyectos de digitalización de manera que los mismos respalden los objetivos básicos de la institución y la misión de la misma
* Resalta la necesidad de tener en cuenta la preservación cuando se está digitalizando -- y explica cómo hacerlo
* Asesora la manera de enfrentar eficazmente a los proveedores
* Enfatiza la importancia de la evaluación de los proyectos
* Propicia que las instituciones compartan sus experiencias.

El manual comienza con un recuento de las razones que se tienen en cuenta para la digitalización y la preservación, lo que nos proporciona las bases para comprender cuáles son las implicaciones de preservación de los proyectos de reconversión a formato digital. El siguiente capítulo ofrece a los administradores una clara comprensión de las decisiones que

típicamente están bajo su control, de manera que puedan diseñar estrategias efectivas para diseñar, financiar y para la gestión de proyectos de digitalización.

El capítulo que trata acerca de la selección de los materiales, presenta un proceso en tres etapas para la selección de los materiales adecuados para el trabajo de digitalización y trata también acerca de cuáles son los criterios para el orden de prioridad de la digitalización. Hace énfasis en la importancia de buenas técnicas de selección para asegurar que se inviertan los recursos de manera inteligente. El capítulo sobre derecho de autor ofrece un breve panorama y una introducción a una serie de aspectos que deben tenerse en cuenta en cualquier proyecto de digitalización o de puesta en Internet. A continuación, un capítulo sobre aspectos técnicos que se concentra en los conceptos técnicos y la terminología que el administrador del proyecto debe conocer para tomar decisiones fundamentadas, bien si el proyecto se va a realizar en la propia institución o en otro lugar. Después se analiza una serie de estudios de casos con énfasis en los resultados prácticos de algunos proyectos de digitalización en los que se presenta la digitalización de textos impresos y manuscritos, fotografías, reconocimiento de caracteres ópticos (RCO), mapas y documentos de gran tamaño, microfilmes y algunos proyectos cooperativos. El capítulo sobre las Relaciones con los Proveedores analiza cómo se puede localizar y evaluar a los proveedores y como se puede monitorear el trabajo de los mismos, y ofrece algunas directivas para desarrollar solicitudes de información y solicitudes de propuestas. El siguiente capítulo se ocupa del futuro -- los aspectos de la preservación a largo plazo y los problemas de la longevidad del material digital. De interés especial es el análisis acerca de cómo puede contribuir nuestra comunidad a los esfuerzos por preservar la información digital. El capítulo final ofrece la visión desde la perspectiva del académico-investigador y del usuario-final de los materiales digitales.

Algunos autores han incluido cuestionarios o listas de aspectos a tener en cuenta para utilizarlos como herramientas para la planificación y la supervisión de los proyectos, así como también listas de los recursos.

Informaciones útiles

* Tómese todo el tiempo que sea necesario al comienzo del proyecto para definir con claridad cuáles son los objetivos y los resultados esperados

* Insista en que se debe realizar el trabajo con la mayor calidad técnica que pueda alcanzar la institución

* Integre los costos y las capacidades para llevar a cabo un mantenimiento a largo plazo de los materiales digitalizados

* Cultive un alto nivel de participación de todo el personal en la realización de los

proyectos de digitalización

* Coopere con otras instituciones siempre que sea posible para obtener mayores beneficios
* Comparta sus experiencias y sus resultados con otras instituciones.

Autores

Howard Besser, Profesor Asociado, Universidad de California, Los Angeles, Facultad de Educación y Ciencias de la Información

Stephen Chapman, Bibliotecario Especialista en Preservación de Proyectos Digitales, Biblioteca de la Universidad de Harvard

Paul Conway, Jefe del Departamento de Preservación, Biblioteca de la Universidad de Yale

Eileen Gifford Fenton, JSTOR Gerente de Producción, Universidad de Michigan

Franziska Frey, Especialista en Ciencias de la Imagen, Image Permanence Institute

Janet Gertz, Directora de Preservación, Bibliotecas de la Universidad de Columbia

Melissa Smith Levine, Esq., Proyecto de la Biblioteca Nacional Digital, Biblioteca del Congreso

Steven Puglia, Rama de Preservación de Medios Especiales, National Archives and Records Administration

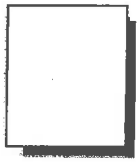
Charles Rhyne, Profesor e Historiador del Arte, Reed College

Steve Smith, Editor, *Microform and Imaging Review* y antiguo coordinador de Servicios de Digitalización de Imágenes, Amigos Library Services Inc.

Diane Vogt-O'Connor, Especialista en Archivos, Museum Management Program, National Park Service

Editora: Maxine Sitts, Consultora de Comunicación, San Leandro, CA

Colaboradores (Capítulo VI): Don Willis, Presidente, eGroup, LLC, and Connetex Consulting



Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

II

Visión general: El razonamiento necesario cuando se piensa en un proyecto de digitalización y preservación

Paul Conway
Biblioteca de la Universidad de Yale

Este capítulo trata sobre los fundamentos para poder comprender las implicaciones de preservación de los proyectos de reconversión a formato digital. Después de una breve descripción de las ventajas y desventajas de las tecnologías digitales, el autor define el concepto de preservación dentro del contexto digital y describe cómo los principios tradicionales de las prácticas de preservación se relacionan con la creación del producto digital. La clave para un buen programa de conversión a formato digital es la relación entre tres conceptos: (1) qué propósitos cumplirá el producto digital, (2) las características de la fuente documental, y (3) las capacidades tecnológicas que se utilizarán durante el proceso de conversión. El corazón de la empresa de conversión a formato digital es la afirmación del autor de que "la preservación es la creación de productos digitales que vale la pena mantener en el transcurso del tiempo". La preservación en el contexto digital es diferente de las acciones de preservación que se adoptan en los materiales originales, pero está estrechamente relacionada a las mismas. El capítulo concluye con la reiteración de la idea de la custodia *responsable*, una idea de gran importancia articulada hace más de cincuenta años para describir

el papel vital que desempeña la preservación en las instituciones culturales.

Introducción

En la obra *El motel de los misterios*, el ilustrador David Macaulay (1979) especula en torno a la idea de cómo las personas interpretarán dentro de 2,000 años la significación cultural de un motel barato en la carretera, **"Tutoentra"**, sepultado intacto bajo una montaña de propaganda barata y residuos contaminantes. Más que una sátira irónica de la arqueología como ciencia, el libro es un hábil recordatorio de los peligros que implica tratar de interpretar el pasado sin evidencias documentales. Entonces un letrero de *No molestar* se convertiría en el sello sagrado "que los funcionarios de la necrópolis colocan en la manilla del gran portón de entrada después de cerrar la tumba". Una tarjeta de cuenta deviene en "un altar móvil que debe llevarse con uno toda la vida y en la vida eterna en el más allá". Un televisor representa "la esencia de la comunicación religiosa". Los arqueólogos y los historiadores saben perfectamente que la ansiedad de registrar y mantener es, en la práctica, parte de la naturaleza humana. La verdad está incrustada en los símbolos y artefactos que creamos y que después guardamos por decisión propia o por casualidad. Y, sin embargo, en los albores del siglo XXI, nos enfrentamos al dilema de Howard Carson, el arqueólogo aficionado personaje principal de la novela de Macaulay: un enorme vacío de conocimientos que se llena con el mito y la especulación. La información en formato digital, la más novedosa moneda de nuestro mundo, es más frágil que los fragmentos de papiro que se encontraron enterrados con los Faraones.

La digitalización de imágenes es "el último grito de la moda". Los más importantes diarios dedican secciones enteras a las nuevas tecnologías digitales. A pesar de que los resultados de las más recientes encuestas demuestran que la mayor parte del contenido de Internet (83%) es de carácter comercial, y que sólo el seis por ciento es de carácter educacional o informativo, persiste el criterio de que todo lo que tiene valor se está digitalizando o se crea en formato digital.

Es cierto que en las bibliotecas y archivos las imágenes digitales son comunes. La calidad de estos productos de imágenes digitales puede ser espectacular. No hay dudas de que la calidad mejorará aun más según la técnica se perfeccione. Las organizaciones están reajustando sus presupuestos, buscan dinero y ya anticipan las nuevas tendencias para ejecutar proyectos de digitalización. Pero, ¿puede cualquier institución --biblioteca, archivo, sociedad histórica o museo --darse el lujo de dilapidar estos fondos? Sin embargo, si no se realizan serios esfuerzos para

garantizar el acceso a largo plazo a los ficheros digitales que creamos hoy, el riesgo de pérdida es extraordinario.

La preservación no es sólo para el mundo del papel. Sabemos que la tecnología de la digitalización de imágenes, en si misma, no ofrece respuestas fáciles a las cuestiones de preservación. En realidad, definir de manera simple qué significa preservar en el medio de la digitalización de imágenes es un verdadero reto. Y es todavía más difícil responder a la visión que nos puede dar dicha definición. El mundo digital nos presenta retos significativos, pero no elimina la necesidad de realizar una actividad de preservación eficaz y responsable (Waters & Garrett, 1996).

Las ventajas del acceso digital

La tecnología de digitalización de imágenes ofrece ventajas bien definidas a las instituciones con grandes colecciones de recursos académicos. El contenido de la información se puede entregar directamente al lector sin intervención humana. También un lector en un lugar remoto puede acceder a la información en formato digital, aunque ese tipo de entrega o de acceso puede poner a prueba las capacidades de hasta el más sofisticado equipo de proyección y las más perfectas redes. La calidad de la imagen digital es extraordinaria y se mejora constantemente. Ahora es posible representar casi cualquier tipo de material de investigación tradicional con una calidad visual tal que casi no es necesario remitirse a los materiales originales en la mayoría de los casos, si no en todos. El potencial de la búsqueda en texto total y de la sofisticada indización cruzada de las colecciones brinda a los lectores la posibilidad de usar los recursos de investigación tradicionales en formas novedosas. Los recién desarrollados sistemas de interfaces (la apariencia y la sensación de la pantalla de la computadora), combinados con otras formas de entrega de grandes ficheros de datos de imágenes, promete revolucionar las formas en que estos materiales de investigación se pueden utilizar para la enseñanza y el aprendizaje. No es de extrañar entonces el abrumador interés y la premura que existen por saltar e incorporarse al carro de la digitalización.

Los riesgos de los proyectos de digitalización de imágenes

Las presiones que se ejercen desde todos los frentes para digitalizar materiales de investigación tradicionales presentan riesgos muy definidos. Es enorme la inversión necesaria para la conversión a imágenes digitales -- posiblemente muchos dólares para todas y cada una de las páginas o cuadros que se convierten. Las tecnologías de

digitalización de imágenes requieren enormes inversiones de capital para establecer sistemas de soporte en medio de presupuestos cada vez más marginados o inexistentes. La conversión a imágenes digitales, en un ambiente operativo, exige por parte de la institución un compromiso profundo y permanente con la preservación tradicional, la total integración de la tecnología en procedimientos y procesos de gestión de la información, y una dirección significativa en el desarrollo de definiciones y normas adecuadas para la preservación digital.

El riesgo de pérdida es muy elevado -- mucho más elevado que en la mayoría de los otros programas y actividades que se llevan a cabo en una institución cultural. El remolino casi constante de desarrollo de productos que alimenta nuestras percepciones de cambio eleva aun más los riesgos. Cuando una biblioteca, un archivo, una sociedad histórica, un museo, o cualquier otra organización cultural con una política de preservación deja de experimentar con la tecnología digital y decide utilizarla para mejorar los servicios o transformar las operaciones, esa institución se ha adentrado en el camino de la preservación.

Lo que no es la digitalización de imágenes

En los últimos años, se han logrados avances significativos en la definición de los términos y en el desarrollo de un programa de investigación para preservar la información digital que o bien "nació en formato digital" o se ha transformado en formato digital a partir de fuentes tradicionales. "La preservación digital se refiere a los diversos métodos para conservar materiales en formato digital con vista al futuro", afirma una reciente declaración del Consejo de Recursos para Bibliotecas e Información (Waters, 1998). Entonces, la preservación digital se ocupa de la selección de medios interinos de almacenamiento, en la esperanza de vida de un sistema de digitalización de imágenes y en las expectativas de migración de los ficheros en formato digital a futuros sistemas a la vez que ambos mantienen plena funcionalidad y la integridad del sistema digital original. Recientemente la cadena de televisión PBS presentó el filme *Pensando en el futuro*, que presenta de manera muy gráfica el problema de la información digital y especula acerca de las consecuencias de la inacción, a la vez que ofrece algunas ideas muy valiosas con respecto a qué hacer ante el dilema.

Quizás sea prematuro que nosotros nos preocupemos por preservar objetos digitales antes de que logremos determinar cómo hacer productos digitales que valga la pena preservar. Las tecnologías de digitalización de imágenes crean una forma de información completamente nueva a partir de documentos tradicionales. La tecnología de digitalización de imágenes no es simplemente otra opción

de reconversión entre las herramientas de preservación. La digitalización de imágenes implica la transformación del concepto mismo de formato, no es simplemente crear una reproducción fiel de un libro, un documento, una fotografía, o un mapa en un soporte diferente. El potencial del mejoramiento digital, las posibilidades de la estructuración de índices, y la matemática de la compactación y la comunicación unidas alteran fundamentalmente el concepto de preservación en el mundo digital. Estas transformaciones, junto con las nuevas posibilidades que ofrecen a los profesionales de la información, nos obligan, en cambio, a transformar los servicios y programas de bibliotecas y archivos.

La preservación en el mundo digital

La esencia de la gestión de preservación tradicional es la asignación de recursos. Hay que buscar el personal, el dinero y los materiales, organizarlos y ponerlos a funcionar o trabajar para evitar el deterioro o renovar grupos seleccionados de materiales que se pueden utilizar. La preservación se preocupa, fundamentalmente, por las evidencias intrínsecas de una variada serie de formas y formatos. Los objetos se preservan para utilizarlos con muchos propósitos, académicos y de otro tipo.

Las personas responsables de tomar decisiones, han determinado que una pequeña porción del vasto océano de la información, estructurada en forma de colecciones de documentos, libros, colecciones y otros objetos, posee valor investigativo y prueba que han sobrepasado el tiempo y la intención de los que los crearon o los publicaron (Buckland, 1991). Esta diferenciación entre el valor de la evidencia intrínseca en el artefacto es el corazón del proceso de toma de decisiones que, en sí mismo, es neural para una gestión efectiva de los materiales de las bibliotecas, tanto las tradicionales como las digitales.

En el mundo digital, preservar es la creación de productos digitales que vale la pena mantener en el transcurso del tiempo.

Cada una de estas palabras tiene su propio peso.

- ES. La preservación es una realidad y no simplemente una metáfora para el acceso o un símbolo del mismo.
- CREACIÓN. El momento en el que debemos preocuparnos con respecto a la persistencia a largo plazo de los productos digitales es el momento en el que se diseña un sistema y antes de que comience la conversión a formato digital.

- PRODUCTOS. Un producto digital tiene su propia identidad y existe dentro de una economía de mercado. No es necesario vender u obtener una licencia para un producto digital para que el producto tenga identidad dentro de una comunidad de usuarios.

- QUE VALGA LA PENA. El trabajo de diseñar y crear un producto digital añade valor a la información que contienen los documentos originales. El valor añadido a un producto digital debe resultar, en última instancia, en un producto que sea un recurso esencial, vital y capital para la institución que en primera instancia ha decidido crear ese producto.

- MANTENER. La persistencia de los productos digitales exige que se conceda cuidadosa atención a mantener el contenido (los bits y los bytes) y su funcionalidad (cómo funcionan los bits en un sistema).

- EN EL TRANSCURSO DEL TIEMPO. La preservación en el mundo digital no es absoluta, sino que depende del continuado impacto transformativo del producto digital en el trabajo de información de los usuarios finales.

Es imposible llegar a un acuerdo entre las responsabilidades inherentes a la creación de productos digitales sin diferenciar entre la adquisición de las tecnologías de digitalización de imágenes para resolver un problema particular y la adopción de esas tecnologías como estrategia de gestión de la información. Adquirir un sistema de digitalización de imágenes para mejorar el acceso a los materiales de bibliotecas y archivos es tan simple como escoger la combinación de escáners, computadoras y monitores que satisfagan las especificaciones funcionales inmediatas. Cientos de organizaciones culturales ya han invertido o están planificando para adquirir sistemas de conversión a imagen digital y experimentar con sus capacidades. Incontables proyectos piloto han demostrado que digitalizar fuentes académicas constituye un reto mucho mayor que la digitalización de la correspondencia moderna de las oficinas y los archivos que fueron los que guiaron la tecnología hace dos décadas. Con el tiempo, la mayoría de estos proyectos piloto de menor envergadura se irán olvidando -- y se perderá la inversión inicial -- pues el costo de mantener estos sistemas es evidente, los proveedores van abandonando el negocio y los usuarios se acostumbran más a las bases de datos de imágenes de acceso remoto y a los últimos gritos de la moda digital.

Los administradores que tienen la responsabilidad de seleccionar los sistemas de conversión de materiales con valor a largo plazo también son responsables de preservar su inversión en el producto. Este compromiso es continuado -- no se puede diferir las soluciones relativas a la preservación con la esperanza de que surjan soluciones por arte de magia como si fuera un caballero de la época medieval con su reluciente

armadura. Una tasación del valor de un libro en un momento determinado, una colección de manuscritos o una serie de fotografías en su formato original es el punto de partida necesario para tomar una decisión o formarse un juicio con respecto a la preservación de la versión digital. Sólo el potencial del incremento del acceso a una colección digitalizada no añade valor a una colección subutilizada. De igual manera, las poderosas capacidades de un índice relacional no pueden compensar el hecho de que una colección de documentos cuya estructura, relaciones y contenido intelectual no se comprendan lo suficiente. El acceso aleatorio no es una poción mágica para una eficaz gestión de las colecciones.

Las relaciones entre los propósitos, la fuente y la tecnología

La clave para que un proyecto de conversión de formato o un programa ya en ejecución tenga éxito es una verdadera comprensión de las relaciones entre tres conceptos. Estos conceptos son (1) las características del material original que se convierte a otro formato, (2) las capacidades de la tecnología utilizada para hacer la conversión digital, y (3) los propósitos o usos que se le van a dar al producto digitalizado. La figura que aparece a continuación ilustra estas relaciones.

Los propósitos de preservación del producto digital

Existen tres aplicaciones diferentes aunque no excluyentes de las tecnologías digitales, en parte definidas por los posibles propósitos que para los usuarios tendrán los productos digitales.

Proteger los originales. La aplicación más común de las tecnologías digitales en un archivo o biblioteca es que las copias digitales se pueden utilizar como referencia inmediata en lugar de estar manipulando descuidadamente las fuentes originales. Se cumplen los objetivos de preservación, porque se limita el acceso físico a los documentos originales. Entre los ejemplos se incluyen los archivos de referencia de fotografías, recortes, o archivos verticales que permiten la identificación de objetos individuales que requieren un estudio más minucioso. El orden original de la colección, o de un libro, se puede congelar de manera muy parecida a las series de imágenes de un microfilme, en organización lineal. Este uso de la tecnología para la preservación se ha

convertido en una fuerza impulsora para motivar a archivos y bibliotecas a experimentar con las capacidades de hardware y software.

Representar los originales. Un sistema digital puede construirse para que represente el contenido de la información de las fuentes originales tan al detalle que el soporte digital satisfaga la mayoría, si no todo, el potencial de investigación y aprendizaje desde los documentos originales. Los sistemas de alta resolución que tratan de lograr un contenido abarcador y total y de capturar toda la información basándose en las normas en surgimiento y en las mejores prácticas, cumplen esta definición. Los sistemas de este nivel de calidad intermedio abren nuevas posibilidades de investigación y de uso y podrían tener un efecto transformador en las misiones de servicios de aquellos que crean los productos.

Trascender los originales. En un número muy pequeño aunque creciente de aplicaciones, la digitalización de imágenes promete generar un producto que se pueda utilizar para propósitos imposibles de alcanzar con las fuentes originales. Esta categoría incluye digitalización que utiliza iluminación especial para resaltar los detalles oscurecidos por el envejecimiento, el uso y deterioro causado por agentes ambientales; digitalización que utiliza intermediarios fotográficos especializados o digitalización de tan alta resolución que posibilita el estudio de las características del artefacto. Esta categoría también incluye productos de digitalización de imágenes que incorporan texto completo buscable (marcado o virgen). Además, los productos digitales que resaltan, organizan o mejoran el acceso a materiales de investigación muy dispersos pueden tener impacto trascendental en las personas que los utilizan.

Cada una de estas aplicaciones de preservación impone exigencias separadas, pero cada vez más rigurosas a las tecnologías digitales. En cada caso, el uso de una copia intermedia en película o papel puede facilitar el proceso de escaneo, o puede que esto no sea necesario o aconsejable. Por último, la disposición de las fuentes originales (incluidos los tratamientos de conservación antes o después de la conversión) es una cuestión completamente diferente de la decisión de emprender la conversión digital. En última instancia, el propósito de los productos de digitalización de imágenes está determinado por los usos que tendrán los mismos, mientras que la preservación de los documentos o fuentes originales estarán determinados por sus necesidades específicas de preservación.

Las características de los materiales originales que se convierten

Un importante reto cuando se decide convertir de formato analógico a formato digital es lograr una comprensión profunda de las características particulares de las colecciones o de los artículos individuales que se van a convertir (Robinson, 1993). Las características más importantes son:

- Formato de la fuente (incluido el tamaño del objeto, su estructura y su estado de conservación físico)
- El estado de conservación desde el punto de vista físico y su impacto en la capacidad del objeto para ser manipulado durante el proceso de conversión
- Características visuales (incluida la centralidad del texto versus ilustraciones)
- El color como transmisor esencial del contenido de la información
- El nivel de detalle (incluido el tamaño y el estilo de los tipos, el tipo de contenido ilustrativo y el rango general de los valores tonales).

Además de estas características específicas, el grado de similitud visual y física entre los objetos individuales en una determinada colección, que puede tener un impacto significativo en el costo, la calidad y la complejidad del proyecto de conversión.

Las capacidades de la tecnología para el escaneo

El tercer aspecto clave en la preparación de un producto de digitalización viable es la medición de las capacidades del sistema de hardware/software para digitalización de imágenes con relación a los documentos o fuentes originales y a los propósitos del producto. Las capacidades y el costo de los sistemas de conversión digital varían mucho. Una rigurosa ingeniería mecánica y eléctrica desempeñan un papel importantísimo en el diseño y la fabricación de herramientas especializadas para la conversión. Muchos productos se optimizan para la conversión de un solo tipo de documento. Todas las herramientas de conversión tienen limitaciones en términos del tamaño de los documentos originales que pueden manipular con un determinado nivel de resolución. Aunque el adagio "Uno tiene lo que uno paga" se aplica a la adquisición de hardware de conversión, no hay nada que sustituya la prueba cuidadosa y el marcaje de los sistemas de conversión (Besser y Trant, 1995).

Los usos que se espera hacer del producto pueden decidir la selección de las aplicaciones tecnológicas, pero lo contrario no es necesariamente válido. Es importante reconocer que las normas y las mejores prácticas que soportan el desarrollo del producto digital no deben estar guiados por las limitaciones presentes de la captura de la imagen digital, del

monitor y de la entrega. Cuestiones como una resolución limitada de las pantallas de los monitores y de los dispositivos de proyección que existen en la actualidad, el ancho de banda limitado de las redes locales o regionales, y las limitaciones de la resolución y el tono de la reproducción en las impresoras no deben determinar los umbrales de calidad del diseño del sistema de imagen.

Las relaciones entre las características de la fuente, las capacidades tecnológicas y los propósitos para los que se va a utilizar el producto final influyen en las definiciones de calidad, costo y acceso. En el área de calidad, por ejemplo, una fuente de entrada con características particulares, las limitaciones o costos de la tecnología de escaneo en un momento determinado y los usos esperados del producto, interactúan para establecer los requisitos del umbral de calidad de la imagen. De igual manera, los propósitos esperados del producto digital y las características del original pueden interactuar con las capacidades de la tecnología de digitalización para determinar el costo de creación del producto con un propósito intendido. Lo mismo se aplica al acceso, en el que la complejidad intelectual de los documentos originales y las especificaciones de las maneras en que se va a utilizar la imagen resultante se utilizarán para interactuar con la sofisticación (o falta de sofisticación) de las herramientas de software y hardware con las que se van a crear los archivos de metadatos y otros índices asociados.

Se transforman los principios de preservación

En las dos últimas décadas, ha surgido un consenso dentro de una comunidad de practicantes con respecto a una serie de principios fundamentales que deben regir la gestión de los recursos disponibles en un programa de preservación bien estructurado. Los principios de preservación en el mundo digital son los mismos que los del mundo analógico y, en esencia, definen las prioridades para extender la vida útil de los recursos de información. Estos conceptos son la longevidad, la selección, la calidad, la integridad y la accesibilidad.

La preservación en el mundo digital es una de las cuestiones centrales de nuestros días. Es una responsabilidad compartida entre muchas personas en muchas instituciones, que además tienen entre sus manos muchas tareas que realizar. Es crucial, por tanto, comprender de manera muy clara el impacto de la diferenciación de esta tarea en la acción de la preservación digital. La diferenciación de las tareas de preservación ayuda a los archivistas y bibliotecarios -- que son los que desarrollan el producto digital -- a conocer cuándo controlar el uso que hacen de las tecnologías digitales, cuándo necesitan ejercer su influencia en ciertas tendencias, y cuándo necesitan renunciar a sus expectativas de control o de influencia.

La transformación de la longevidad

La principal preocupación en la preservación tradicional es el soporte en el que se almacenará la información. La prioridad es extender la vida del papel, la película, y la cinta magnética estabilizando las estructuras y limitando las habilidades de los factores externos e internos para causar deterioro. El estudio de los factores externos ha conducido a especificaciones para controles ambientales adecuados y también a directivas de cuidado y manipulación adecuados, así como a procedimientos para la recuperación cuando se produce un desastre. Los avances que resultan de los esfuerzos por controlar o mitigar los factores internos de deterioro han conducido a su vez a normas para la producción de papel alcalino, microfilme de calidad para materiales de archivo, técnicas de desacidificación masiva, y soportes magnéticos más resistentes. Y, sin embargo, ahora que los archivistas y los bibliotecarios han definido las cuestiones relativas a la esperanza de vida de los soportes de almacenamiento, el concepto mismo de permanencia que dio impulso a la búsqueda de soportes "de archivo" está perdiendo su significado intelectual para la preservación (O'Toole, 1989).

La preservación en el contexto digital tiene poco interés en la longevidad de los discos ópticos y de los soportes de almacenamiento más modernos

y más frágiles. La viabilidad de los archivos de imágenes digitales depende mucho más de la esperanza de vida del sistema de acceso -- la cadena es tan fuerte como el eslabón más débil. Probablemente los soportes ópticos de hoy día durarán muchísimo más que la capacidad de los sistemas para acceder e interpretar los datos almacenados en ellos. Como a través del vendedor nunca se puede saber a ciencia cierta cuándo ya no se podrá mantener o utilizar un sistema, los que desarrollan los productos deberán anticipar que será necesario migrar (o sea trasladar) los datos de imágenes valiosos, los índices, y los softwares durante la vida profesional de los mismos a generaciones futuras de medios tecnológicos.

Los administradores de proyectos de digitalización pueden ejercer un gran control sobre la longevidad de los datos de imágenes digitales a través de una selección cuidadosa, una manipulación adecuada y con el almacenamiento en soportes resistentes y bien probados. Pueden influir además en la esperanza de vida de la información asegurándose de que los compromisos presupuestarios locales sean consecuentes en el nivel adecuado. En última instancia, ellos no tienen control sobre la evolución del mercado de la imagen, actividades corporativas especiales de investigación y desarrollo que tienen un impacto tremendo en la esperanza de vida de los sistemas digitales creados en la actualidad.

La transformación de la selección

Escoger es seleccionar. La preservación añade valor a través del proceso de selección. Escoger implica definir valores, reconocerlos en alguna cosa, y después decidir cómo enfrentarse a sus necesidades de preservación en la forma más adecuada para esos valores. En el transcurso del tiempo, el acto de preservar ha evolucionado desde salvar un material del olvido y colocarlo en edificaciones seguras, hasta formas más sofisticadas de evaluación de las condiciones y del valor de los materiales que ya constituyen colecciones. La selección de la preservación ha estado guiada en gran medida por la necesidad de estirar los recursos de los que se dispone de la manera más inteligente posible, lo que ha resultado en la máxima "ningún objeto deberá preservarse dos veces". El resultado neto es el surgimiento y crecimiento de una colección virtual especial de objetos preservados con una amplia variedad de técnicas, la más notable la de la reconversión al formato de microfilme. La selección es quizás el más difícil de todos los aspectos, precisamente porque es extático y lo conciben personas que o están completamente divorciadas del presente o están completamente dominadas por las exigencias de las necesidades.

La selección en el mundo digital no es una decisión que se toma para todos los materiales cuando están próximos al fin de su ciclo de vida, sino que es más bien un proceso continuo íntimamente relacionado con

el uso activo de los archivos digitales (Hazen, 1998). Los criterios aplicados cuando se toma una decisión de convertir documentos en papel o en película a imágenes digitales sólo son válidos dentro del contexto del sistema original. En realidad, una colección valiosa justifica el costo de una amplia estrategia de migración sin integrar todos los factores en el más amplio contexto intelectual de archivos digitalizados que se encuentran diseminados y combinar los usos de los mismos para enseñar y aprender.

Incluso cuando se reconoce que las decisiones en la selección no se pueden tomar de manera autónoma o en un vacío, los archivistas y los bibliotecarios pueden seleccionar qué libros, artículos, fotografías, filmes y otros materiales se convierten de soporte de papel o película al formato de imágenes digitalizadas. La influencia sobre el valor continuado de los archivos de imágenes digitales radica fundamentalmente en el derecho a decidir cuándo es el momento adecuado para migrar los datos en imágenes a sistemas futuros de almacenamiento y acceso, y cuándo un archivo digitalizado a sobrepasado su vida útil para la institución encargada de su preservación. Lo que los que desarrollan los productos no pueden controlar de ninguna manera es el impacto en el valor continuado sobre las capacidades de los lectores para encontrar y utilizar información en formato digital. Un producto digital que no se utiliza no tiene razón de existir, y por supuesto, no sobrevivirá mucho tiempo, pues es un simple artefacto en el proceso de conversión.

La transformación de la calidad

Los que desarrollan los productos digitales deben exigir la calidad de la imagen, que es el alma y el corazón de la preservación

Elevar al máximo la calidad de todos los trabajos realizados es un adagio tan importante en el campo de la preservación que muy pocas personas plantean este principio fundamental de manera directa. En lugar de hacerlo de manera directa, la literatura sobre preservación exige resultados de alta calidad en todas las opciones de tratamiento, los procesos de reconversión de formato, y las medidas preventivas. El compromiso con las normas de calidad -- hazlo una sola vez, pero hazlo bien -- permea todas las actividades de preservación, incluidas las normas de encuadernación en bibliotecas, los lineamientos de creación de microfilmes de archivo, los procedimientos de tratamientos de

conservación, la selección de los suministros, y la baja tolerancia de errores. La evolución de la microfilmación con propósitos de preservación como estrategia central a seguir con respecto a la mayoría de los materiales de biblioteca que están quebradizos, han colocado la calidad del soporte y la calidad de la imagen visual en el mismo plano. Tratando de obtener un microfilme de calidad, el compromiso entre la verdad visual y la estabilidad archivística se rige en gran medida por las características de los objetos que se seleccionan para preservar.

Por otra parte, la calidad en el mundo digital está condicionada de manera muy significativa por las limitaciones de capacidad de captura y tecnología de visualización. La conversión digital hace menos énfasis en la obtención de una reproducción fiel del original para favorecer la mejor representación del original con una tecnología determinada. Los mecanismos y las técnicas para juzgar la calidad de las reproducciones digitales son diferentes y más sofisticadas que las de evaluación de reproducciones en microfilme o en fotocopia (Kenney & Chapman, 1996). Además de eso, el principal objetivo de la calidad en la preservación es capturar la mayor cantidad de contenido intelectual y visual que sea posible y entonces presentar ese contenido a los usuarios en las formas más adecuadas según sus necesidades.

El mercado de la imagen ha incluido el principio de calidad máxima en la "solución" que sitúa el nivel mínimo aceptable de calidad en los usuarios de los sistemas actuales. Los que desarrollan los productos digitales deben exigir e insistir en la calidad de la imagen como el alma y el corazón de la preservación. Esto quiere decir elevar al máximo la cantidad de datos capturados en el proceso de escaneo, en las técnicas de mejoramiento de la imagen del documento, y especificar que las técnicas de rutina que se utilizan para compactar archivos no resulten en pérdida de datos durante la telecomunicación. Hoy es posible el control de las normas de calidad digital, al igual que es posible el control de las normas de calidad en el microfilme. Sin embargo, los archivistas y bibliotecarios sólo pueden influir en el desarrollo de normas para la compactación de datos, la comunicación y la visualización y salida de los mismos. El mejoramiento de las capacidades técnicas de los hardwares y los softwares de conversión de imágenes está en las manos de la industria de la digitalización de imágenes.

La transformación de la integridad

El concepto de integridad tiene dos dimensiones dentro del contexto de la preservación tradicional -- la física y la intelectual -- y ambos están relacionados con la naturaleza de las evidencias contenidas en el documento. La integridad física se ocupa fundamentalmente del objeto como artefacto. Funciona de manera más directa en el estudio o

laboratorio de conservación, en el que el personal calificado utiliza adhesivos solubles en agua, técnicas de encuadernación ancestrales, y materiales de alta calidad para proteger las evidencias históricas del uso, de los tratamientos de conservación aplicados en el pasado, y de cambios intencionales o no intencionales en la estructura del objeto. La preservación de la integridad intelectual se basa en un interés de otro tipo por las evidencias. La autenticidad, o veracidad, del contenido de la información de un objeto que se ha mantenido a través de la documentación que certifica la procedencia del objeto -- la cadena de propietarios -- y el tratamiento, cuando es apropiado, es el corazón de la integridad intelectual. Más allá de la historia de un objeto está la preocupación por proteger y documentar las relaciones entre los objetos en una misma colección. En las prácticas de preservación tradicionales, los conceptos de calidad e integridad se refuerzan el uno al otro.

En el mundo digital, mantener la integridad física de una imagen digital tiene mucho menos que ver con el soporte que con la pérdida de la información cuando se crea un archivo por primera vez, después se compacta por medios matemáticos, se almacena en varios formatos y se envía por una red. En el reino de la integridad intelectual, los índices estructurales y las descripciones de datos que se publican tradicionalmente con un objeto en forma de tablas de contenidos o preparados como discretos medios de búsqueda o registros bibliográficos, deben estar muy estrechamente unidos y preservados junto con las imágenes digitales mismas. Preservar la integridad intelectual implica también procedimientos de autenticación, como por ejemplo verificar todas las informaciones, que garantizan que los archivos no se han alterado de manera intencional o accidental (Duranti, 1995). En última instancia, el mundo digital fundamentalmente transforma los principios tradicionales de preservación de garantizar la integridad física del objeto a especificar la creación del objeto cuya integridad intelectual es su primera característica.

Los bibliotecarios y archivistas pueden ejercer control sobre la integridad de los archivos de imágenes digitales al autenticar los procedimientos de acceso y documentar las sucesivas modificaciones realizadas a un determinado registro digital. También pueden crear y mantener índices estructurales y vínculos bibliotecarios dentro de normas de bases de datos bien desarrollados y bien entendidos. Los que desarrollan productos digitales también tienen un papel a desempeñar influyendo sobre el desarrollo de normas de intercambio de metadatos, incluidas las herramientas y las técnicas que posibilitarán compartir a través de plataformas, sistemas y fronteras internacionales, la información normalizada, documentada y estructurada. No obstante, es vano pensar que los bibliotecarios y archivistas no son más que

observadores pasivos del rápido desarrollo de las redes de protocolos, anchos de bandas o de las técnicas de seguridad de datos que son esenciales para la duración de los objetos digitales en el tiempo.

La transformación del acceso

En los cincuenta años que hace que la preservación surgió y se ha desarrollado como especialidad profesional en bibliotecas y archivos, las responsabilidades de preservación y acceso en un archivo o biblioteca siempre han estado en tensión. "Aunque la preservación es un objetivo o responsabilidad de primera línea, otro objetivo igualmente importante -- el del acceso y el uso -- crea un conflicto básico que debe ser arbitrado por los custodios y guardianes de los registros archivísticos", dice un libro de texto fundamental en este campo (Ritzenthaler, 1993). La estrecha relación existente entre preservación y acceso ha cambiado en formas que reflejan el ambiente tecnológico de las instituciones culturales.

Preservación O Acceso. En los primeros años de existencia de las modernas instituciones archivísticas -- previos a la II Guerra Mundial -- preservar quería decir simplemente coleccionar. El mero acto de rescatar una colección de manuscritos de un granero, un sótano o un estacionamiento o garaje, y colocarlos intactos en una edificación en un ambiente seco y con cerraduras en las puertas, satisfacía los objetivos fundamentales de preservación de la institución. En este sentido, la preservación y el acceso son actividades mutuamente excluyentes. El uso expone a las colecciones a riesgos de robo, daños o uso incorrecto del contenido o del objeto mismo. La forma más segura de garantizar que un libro perdure mucho tiempo es encerrarlo en lugar seguro o hacer una copia para usar la misma.

Preservación Y Acceso. Las estrategias modernas de gestión para la preservación establecen que la preservación y el acceso son ideas que se refuerzan mutuamente. Se ejecuta una acción de preservación en un objeto para que ese objeto se pueda utilizar. Siguiendo esta opinión, crear una copia de preservación en microfilme de un libro deteriorado sin posibilitar el acceso al filme, es una pérdida de dinero. En el mundo de la preservación **Y** el acceso, sin embargo, es posible, desde el punto de vista teórico satisfacer una necesidad de preservación sin resolver los problemas de acceso. En contraste, es posible garantizar el acceso a los materiales académicos por un largo período de tiempo sin, en realidad, adoptar ninguna medida concreta de preservación directamente en los materiales.

Preservación ES Acceso. Los archivistas y bibliotecarios interesados en la preservación de los registros electrónicos algunas veces ven los dos

conceptos como causa y efecto. El acto de preservar hace posible el acceso. Sin embargo, igualar la preservación con el acceso implica que la preservación se define como la disponibilidad, cuando en realidad esta teoría puede confundir todo el concepto. La preservación es tan acceso como el acceso es preservación. Simplemente si enfocamos la cuestión de la preservación en el acceso simplifica en exceso las cuestiones de la preservación, ya que sugiere que el acceso es el motor impulsor de la preservación sin tratar la naturaleza del objeto que se preserva.

Preservación DEL Acceso. En el mundo digital, la preservación es la acción, y el acceso es la cosa -- el acto de preservar el acceso. Un concepto más exacto plantea simplemente "preservar la accesibilidad". Cuando se transforma de esta manera, surgen toda una serie de nuevas complejidades. ¿Preservar el acceso a qué? La respuesta que sugiere este capítulo es: a un producto digital de alta calidad, alto valor, bien protegido y completamente integrado que se deriva de fuentes documentales originales pero a la vez es independiente de ellas. El contenido, la estructura y la integridad del producto digital asumen importancia central -- y la habilidad de una máquina para transportar y exhibir este producto se convierte en el resultado final asumido de la acción de preservación más que su objetivo principal.

El control sobre la accesibilidad, especialmente sobre la capacidad del sistema para exportar imágenes digitales (e índices asociados) a generaciones futuras de tecnologías, se puede ejercer, en parte a través de adquisiciones prudentes sólo de componentes de hardware y software no patentados. En el ambiente actual, hay una más amplia disponibilidad de componentes verdaderos de plug-and-play. El compromiso financiero de bibliotecarios y archivistas es uno de los únicos incentivos que tienen los proveedores para adoptar arquitecturas de sistemas abiertos o al menos proporcionar mejor documentación acerca del funcionamiento interno de sus sistemas. Además, los bibliotecarios y archivistas pueden influenciar a los proveedores y fabricantes para que proporcionen nuevos equipos que sean compatibles en retroceso con los sistemas existentes en la actualidad. Esta capacidad ayudaría a la migración de los sistemas de archivos de imágenes en la misma forma que los softwares de procesamiento de textos permiten el acceso a documentos creados en versiones anteriores. Por mucho que quisieran lo contrario, los que desarrollan productos digitalizados tienen poco o ningún control sobre la esperanza de vida de un determinado sistema de imágenes digitales y sobre la decisión de abandonar ese sistema.

Conclusión

Hace cincuenta años, uno de los más destacados y persistentes defensores de la calidad en la encuadernación de libros señaló la importancia de la preservación en la misión de las bibliotecas de investigación y los archivos modernos. La preservación, escribió Pelham Barr en su obra más citada, "como custodia responsable, es la única función de la biblioteca que debe estar funcionando constantemente, las veinticuatro horas del día. Es la única función que debe preocuparse de cada una de las piezas y materiales de la biblioteca desde el momento en que el que selecciona esa pieza se hace consciente de su existencia hasta el día en que la misma se descarta " (Barr, 1946). La alusión de Barr al ciclo de vida de las fuentes de información no tiene límite de tiempo. En la actualidad, ese concepto es el centro de la teoría y la práctica de gestión de la información, incluidas las especificaciones para la disposición de los archivos gubernamentales, la gestión de colecciones de libros, y el mantenimiento de los sistemas tecnológicos de información a gran escala. La custodia responsable circunscribe la preservación en el mundo digital también, mientras que la creación de productos digitales que valga la pena mantener en el transcurso del tiempo es la medida del éxito. La idea de la custodia responsable debe gobernar las acciones cuando construimos productos digitalizados imbuidos del valor de las empresas intelectuales.

Resumen de principios y puntos claves

- Defina fronteras muy claras para un proyecto de conversión a formato digital, especialmente el objetivo final.
- Piense con cuidado: en términos no técnicos, establezca cuáles son los resultados que desea obtener para las fuentes originales y los requisitos funcionales para las reproducciones digitales.
- Justifique por qué es necesaria la reproducción en formato digital en lugar del analógico.
  -- Describa el público al que está dirigido y las necesidades del mismo.
  -- Describa las cosas que se podrán hacer con las copias digitalizadas que no se podrán hacer con las copias en formato analógico.
- Proyecte una duración para la reproducción digital.
- Plan: Escriba un plan para el proyecto que incluya el presupuesto, el tiempo de realización y otros documentos de planificación.
- Presupuesto y plan de flujo de trabajo basados en los resultados del escaneo y la catalogación de una muestra representativa del material.
- Presupuesto (tiempo, o dinero) para adiestramiento.

- Ejecución: Coordinar flujos de trabajo simultáneos o superpuestos.
- Separe los materiales en lotes para facilitar la conversión y el control de la calidad.
- Escriba toda la documentación durante la ejecución del proyecto.
- Informe sobre las lecciones aprendidas, especialmente los fracasos y los callejones sin salida: ayúdese a si mismo y a sus colegas para aprender a partir de los errores que Ud. haya cometido.

Bibliografía

Barr, Pelham. "Book Conservation and University Library Administration," *College & Research Libraries 7* (julio 1946): 218-19.

Besser, Howard y Jennifer Trant. *Introduction to Imaging: Issues in Constructing an Image Database.* Santa Monica: Getty Art History Information Program, 1995. http://www.gii.getty.edu/intro_imaging/

Buckland, Michael K. "Information as Thing," *Journal of the American Society for Information Science 42* (junio 1991): 351-60.

Conway, Paul. *Preservation in the Digital World.* Washington, DC: Commission on Preservation and Access, marzo 1996. http://www.clir.org/cpa/reports/conway2/

Digital Imaging Technology for Preservation. Proceedings from an RLG Symposium Held March 17 and 18, 1994. Nancy E. Elkington, ed. Mountain View, CA: Research Libraries Group, 1994.

Duranti, Luciana. *The Preservation of the Integrity of Electronic Records.* School of Library, Archival, and Information Studies, University of British Columbia, 1994-97. http://www.slais.ubc.ca/users/duranti/intro.htm

-----. "Reliability and Authenticity: The Concepts and Their Implications." *Archivaria 39* (Primavera 1995): 5-10.

Ester, Michael. *Digital Image Collections: Issues and Practice.* Washington, DC: Commission on Preservation and Access, 1996.

Frey, Franziska. "Digital Imaging for Photographic Collections: Foundations for Technical Standards." *RLG DigiNews 1* (3), diciembre15, 1997. http://www.rlg.org/preserv/diginews/

Graham, Peter S. "Requirements for the Digital Research Library." *College & Research Libraries 56* (julio 1995): 331-39.

Hazen, Dan, Jeffrey Horrell, y Jan Merrill-Oldham. *Selecting Research Collections for Digitization.* Washington, DC: Council on Library and Information Resources, 1998.
http://www.clir.org/pubs/reports/hazen/pub74.html

Kenney, Anne R. y Stephen Chapman. *Digital Imaging for Libraries and Archives.* Ithaca, NY: Dept. of Preservation and Conservation, Cornell University Library, 1996.

Levy, David M. "Heroic Measures: Reflections on the Possibility and Purpose of Digital Preservation." *Proceedings of the Third ACM Conference on Digital Libraries,*1998, pág. 152-61.
http://www.acm.org/pubs/citations/proceedings/dl/276675/p152-levy/

Lynch, Clifford. "The Integrity of Digital Information: Mechanics and Definitional Issues." *Journal of the American Society for Information Science 45* (diciembre 1994): 737-44.

Macaulay, David. *Motel of the Mysteries.* Boston: Houghton Mifflin, 1979.

Mohlhenrich, Janice. *Preservation of Electronic Formats: Electronic Formats for Preservation.* Fort Atkinson, WI: Highsmith, 1993.

O'Toole, James M. "On the Idea of Permanence." *American Archivist 52* (Invierno1989): 10-25.

Reilly , James M. y Franziska A. Frey. *Recommendations for the Evaluation of Digital Images Produced from Photographic, Microphotographic, and Various Paper Formats.* Report to the Library of Congress National Digital Library Project. Rochester, NY: Image Permanence Institute, mayo1996.
http://memory.loc.gov/ammem/ipirpt.html

Ritzenthaler, Mary Lynn. *Preserving Archives and Manuscripts.* Chicago: Society of American Archivists, 1993, p.1.

Robinson, Peter. *The Digitization of Primary Textual Sources.* Office for Humanities Communication Publication, no. 4. Oxford: Oxford University Computing Services, 1993.

Rothenberg, Jeff. "Ensuring the Longevity of Digital Documents." *Scientific American 272* (enero, 1995): 42-47.

Smith, Abby. *The Future of the Past: Preservation in American Research Libraries.* Washington, DC: Council on Library and Information Resources, 1999.
http://www.clir.org/pubs/reports/pub82/pub82text.html

-----. *Why Digitize?* Washington, DC: Council on Library and Information Resources, 1999. http://www.clir.org/pubs/reports/pub80-smith/pub80.html

Van Bogart, John W. *Magnetic Tape Storage and Handling: A Guide for Libraries and Archives.* Washington, DC: Commission on Preservation and Access, 1995.

Waters, Donald J. "Digital Preservation?" *CLIR Issues* (noviembre/diciembre 1998): 1. http://www.clir.org/pubs/issues/issues.html

-----. "What Are Digital Libraries?" *CLIR Issues 4* (julio/agosto, 1998). http://www.clir.org/pubs/issues/issues04.html

Waters, Donald and John Garrett. *Preserving Digital Information: Report of the Task Force on Archiving of Digital Information.* Washington, DC: Research Libraries Group and Commission on Preservation and Access, mayo 1996. http://www.rlg.org/ArchTF/

Buscar Indice Sitio del NEDCC

Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

III

Consideraciones para la gestión del proyecto

Stephen Chapman
Harvard University Library

Los bibliotecarios y los archivistas son expertos en la gestión de proyectos. Para ellos es cuestión de rutina el procesamiento de grupos de materiales en flujos de trabajo de selección, procesos, catalogación y preservación. Sin embargo, los proyectos de digitalización presentan nuevos retos, y quizás el reto más difícil sea establecer límites muy claros, especialmente el punto donde hay que detenerse. Los administradores de algunos proyectos que vale la pena destacar han escrito acerca de sus experiencias en la creación de colecciones que exigen modificaciones constantes para mantenerse al día en los avances tecnológicos (Thomas, 1998).

Esto no quiere decir que los proyectos de conversión a formato digital no se puedan planificar bien con tiempo suficiente y ejecutarlos de manera exitosa hasta su conclusión. Se pueden anticipar muchas cuestiones y retos, y también gran parte del flujo de trabajo se puede estructurar en forma de actividades en lotes con resultados predecibles. El propósito de este capítulo es que los administradores comprendan con claridad cuáles son las decisiones que típicamente estarán bajo su control, de manera que puedan trazar estrategias eficaces para diseñar, financiar y administrar proyectos de digitalización.

El establecimiento de los objetivos

Los proyectos de conversión mejor administrados tienen objetivos muy claros. La primera fase de la gestión del proyecto, pensar la estructura del mismo, es el momento de hablar acerca de los resultados del mismo.

"Comenzar por el final" es una forma efectiva de garantizar un buen comienzo. La tendencia más frecuente es ir directo a las cuestiones de la tecnología -- *por ejemplo,* ¿qué escáner debo comprar? -- antes de articular los propósitos que debe cumplir el proyecto de reconversión a formato digital. Trazarse objetivos es un proceso que implica pensar acerca de las cosas desde varios ángulos antes de escribir el plan del proyecto. ¿Cuáles son los posibles resultados para la colección? ¿Cuáles son los beneficios potenciales para los usuarios, para los administradores de la colección, y para la institución? ¿Cuál es el precio razonable -- en tiempo y dinero -- para invertir en nuevos procedimientos, sistemas y servicios? ¿Es buena idea que la institución por si misma lo publique, o sería mejor publicarlo en asociación (con otras instituciones o incluso editoriales)? ¿Es este el momento oportuno para comenzar a digitalizar las colecciones?

Una buena gestión es en gran medida un acto de comunicación. Si las personas que trabajan en el proyecto comprenden cuáles son los resultados que se desea obtener, ofrecerán mejores servicios; serán conscientes de su contribución individual y de cómo esa contribución se relaciona con lo que están haciendo otros; sabrán por qué están digitalizando las colecciones *(la cuestión de la visión)*; y, quizás esto sea lo más importante, podrán reconocer mejor cuándo hay algo que está mal.

El enfoque de comenzar-por-el-principio se refiere a concentrarse más en los resultados antes de analizar los materiales que se van a utilizar como fuentes o de evaluar los procesos de conversión. Como se describe en las secciones siguientes, por lo general los resultados se dividen en tres categorías: colecciones, reproducciones digitalizadas, y beneficios institucionales. Antes de escribir el plan y el presupuesto de un proyecto, reúna a todos los depositarios que tienen interés en estas cuestiones y establezca prioridades que todos puedan aceptar.

Las colecciones

Cuando se habla de preservación y acceso como objetivos de un proyecto, esa afirmación tiene una cierta calidad transitiva. Los proyectos de conversión a formato digital se acometen en nombre de colecciones originales. (*Original* aquí se utiliza para referirse a cualquier material que se utilice como fuente para escanearlo, independientemente de su formato.)

Una de las razones más populares para invertir en la digitalización de colecciones es que con las copias se disminuye, o se elimina completamente, la manipulación física que amenaza materiales frágiles o únicos (Noerr, 1999). Esto parece algo sensato, pero cuidado con la responsabilidad que implica continuar avanzando sobre la base de esta lógica. Recuerden que las colecciones digitalizadas no se hacen solas, y

tengan en cuenta que es muy probable que durante el proceso de conversión se manipule mucho más la colección que en el resto de su vida en la institución. Por tanto, cuando se digitaliza para preservar, no sólo hay que pensar en los resultados, sino también en los lineamientos de manipulación que serán de obligatorio cumplimiento durante el proceso de conversión. Hay que recordar también que, por lo general, la manipulación más cuidadosa se traduce en un aumento de los costos.

Una vez que se han seleccionado los materiales para la reconversión, deben articularse los resultados físicos específicos que se desea obtener con los materiales originales. Cada vez que se retiran los originales de la circulación -- ya sea por un cambio en la política, por un traslado a otro lugar de almacenamiento, o, aunque en raras ocasiones, para descartarlo -- los requisitos para la digitalización serán muy altos. Como se especifica en la Sección 1 del Capítulo VII, "Cómo se trabaja con materiales impresos y con manuscritos", alta calidad no quiere decir necesariamente costos elevados, sino control de la calidad, autenticación de los ficheros y sus fuentes, y otras cuestiones que se vuelven más críticas en casos en que los materiales originales no se pueden extraer o consultar fácilmente.

En todos los proyectos, lo mismo si el objetivo de la digitalización es acceso o preservación, o ambos, tenga en cuenta las siguientes cuestiones:

- ¿Se valorarán, tratarán o conservarán los materiales antes de que se devuelvan a los depósitos?
- ¿Será necesario, o aconsejable, limpiar o estabilizar los materiales antes de escanearlos?
- ¿Uno de los objetivos es envolverlos de nuevo?
- ¿Se devolverán los objetos a sus ubicaciones originales o se enviarán a otros lugares de almacenamiento fuera de la institución para recuperar espacio necesario y/o mejorar las condiciones de almacenamiento de los materiales?

Otras cuestiones abordan las políticas de acceso y catalogación.

- ¿Después de escanearlos, los objetos que antes eran circulantes pasarán a categoría no circulante?
- ¿Se cambiarán las políticas de acceso a los archivos y manuscritos?
- ¿Se utilizarán las copias para exhibiciones?
- Si el objetivo es mejorar el acceso a las colecciones, ¿hasta qué punto entonces debe mejorarse la descripción bibliográfica (*ej.* registros de catalogación, medios de búsqueda)?
- ¿Es necesario crear nuevos registros?

- ¿Es necesario codificar los medios de búsqueda?

Sentar primero los objetivos para las colecciones originales facilitará reducir la amplia selección de posibilidades en cuanto a tecnologías y metodologías para escaneo. Con materiales que se deterioran -- como los periódicos, libros y revistas científicas de papel muy quebradizo, libros de notas y de recortes -- puede ser aconsejable un enfoque híbrido. Esto exige que se planifiquen dos, o hasta tres, flujos de trabajo, la creación de copias digitales para el acceso, crear microfilmes para la preservación y, si fuera necesario, envolver de nuevo o aplicar tratamiento a los originales.

Las reproducciones digitales

No hay un enfoque único que sirva para todos los proyectos de escaneo porque hay muchos tipos de materiales originales, diversos públicos con un amplio rango de intereses, y muchas posibilidades de formatos digitales que se amplían constantemente. El más diligente estudiante de tecnología -- incluso en lo que se refiere al campo de las bibliotecas digitales -- no podría mantenerse actualizado con todos los productos nuevos o con los que están en desarrollo. Incluso cuando las personas son conocedoras de los formatos digitales, es aconsejable prepararse para analizar diversas estrategias. No debe darse por sentado que habrá consenso en cuanto a qué es lo mejor.

En última instancia, es el administrador del proyecto, no el fabricante de la tecnología o el distribuidor el que debe decidir si se necesita un sistema determinado para realizar el proyecto. Son los bibliotecarios y los archivistas, más que los ingenieros, los que tienen los conocimientos para describir en términos prácticos cuál será la función de las reproducciones digitales. Parafraseando a Michael Ester, Presidente de Luna Imaging, Inc., formular el razonamiento para digitalizar una colección se basa en las habilidades de los curadores para ejercer su buen juicio. Sólo entonces se podrá valorar la tecnología según los objetivos del proyecto y no *viceversa* (Ester, 1997).

Existen dos escuelas de pensamiento acerca del desarrollo de las especificaciones para la reconversión a formato digital. Una aboga por una cuidadosa valoración de las fuentes, para después relacionar los atributos de las reproducciones digitales a los de los originales. A veces se denomina esta práctica como marcaje (Kenney, 1999). La otra, recomienda que se relacione los atributos de las reproducciones digitales a los de los sistemas de hardware y software con los que se visualizarán o procesarán. Veamos un ejemplo: suponga que se va a realizar un trabajo con originales impresos, por ejemplo un papiro. Escanear para crear una impresión de alta calidad no necesariamente tiene que satisfacer los

requisitos para ampliar los detalles en la pantalla en una proporción de 10:1.

Caracterizar los requisitos funcionales desde el punto de vista del usuario puede facilitar muchísimo la tarea de definir las especificaciones técnicas. ¿De qué recursos dispone el usuario(s) al que se pretende dar servicio? Dar respuesta a esta interrogante es inquietante, sobre todo cuando uno pretende que las colecciones digitales perduren. ¿Cuáles, debemos pensar, serán los sistemas de que dispondrán los usuarios dentro de diez o veinte años? Cuando piense en todas las formas en que se puede utilizar la tecnología para mejorar el acceso a las colecciones, tenga en cuenta:

- ¿Quiénes son los posibles usuarios de las colecciones digitales?
- ¿Cómo localizarán esos usuarios las colecciones, los objetos dentro de las colecciones, e importantes sub-series dentro de cada objeto?
- Si las imágenes se entregan a la pantalla, ¿cómo se visualizarán? ¿De una en una? ¿Varias a la vez para facilitar las comparaciones? ¿Será necesario ampliarlas?
- Si fuera necesario imprimir, ¿se entregarán las imágenes (o el texto completo) por páginas de una en una o en segmentos de tamaño adecuado?
- Desde su propia perspectiva como usuario, ¿qué constituye la legibilidad? ¿Es necesario reproducir los colores? De ser así, ¿con fidelidad hacia el original? ¿es necesario que las reproducciones imiten la apariencia contemporánea (*ej.*, desvanecimiento del color) las condiciones originales, o ambas?

El marcaje, en contraste, considera que los intereses de los creadores de las colecciones (los artistas y editores originales) y los custodios de las mismas son tan importantes como los usuarios de hoy. En este enfoque, los atributos de las fuentes que es necesario transmitir en las reproducciones digitales (ya sean pictóricas, en metadatos textuales, o ambos) son:

- Organización y presentación (valor colectivo de una página de un libro de recortes, por ejemplo; o presentación una al lado de la otra de páginas publicadas originalmente en forma de códices)
- Tamaño y dimensiones
- Detalles, tonalidad y color
- Antigüedad y estado de conservación.

Al considerar el proyecto desde la perspectiva del usuario y del propietario, el administrador del proyecto estará en una posición mejor para articular los objetivos del proyecto ante el personal que trabajará en el mismo y/o los proveedores que ofrecen sistemas y servicios. Se pueden establecer relaciones de trabajo exitosas cuando los representantes de las instituciones culturales pueden describir los requisitos de funcionamiento de las reproducciones digitales; cuando los representantes de la industria pueden entonces responder con ofertas de lo que puede hacer la tecnología -- pueden incluso sentirse motivados a crear nuevos sistemas.

Quizás el objetivo más importante en este punto de la planificación es responder la siguiente pregunta: ¿Pudiera expresar los requisitos funcionales que sólo se pueden satisfacer con reproducciones digitales? Si no puede, reconvierta sus colecciones con cualquier otro proceso analógico (National Endowment for the Humanities, 1999).

Un comentario final acerca de la definición de requisitos de las reproducciones digitales: Debe asignarse una duración, aunque sólo sea aproximada, a las ediciones digitales para ayudar a definir los requisitos técnicos de conversión así como el presupuesto general del proyecto. Con los formatos analógicos, podemos dar por sentado que la institución se hará cargo de los costos corrientes de almacenamiento, catalogación, y de proporcionar acceso a las reproducciones. Se considera que se puede afrontar el costo general por concepto de instalaciones para almacenamiento y tecnologías de apoyo como los sistemas de circulación, fotocopiadoras y lectores de microformas.

Con los formatos digitales, serán relativamente frecuentes las intervenciones, y se podrá definir el mantenimiento desde una simple copia (a nuevos soportes y/o nuevos formatos) hasta buscar un presupuesto completo para una conversión de formato digital a otro formato digital para mantener un nivel mínimo de servicios. Una cosa es preservar el contenido, otra es preservar un nivel de servicios. Todo esto quiere decir que la longevidad no es un atributo físico de las reproducciones digitales, sino una duración asignada que se respalda con el reconocimiento de que las decisiones de hoy con respecto a la calidad y funcionalidad digital la tendrán que respaldar los administradores de mañana con una parte de sus presupuestos operacionales.

Beneficios para la institución

En años recientes, muchas organizaciones han invertido en proyectos digitales con el ojo puesto en la obtención de beneficios para la institución, así como en el mejoramiento de sus colecciones. La Universidad de Oxford, por ejemplo, otorga una categoría a los

proyectos de digitalización de acuerdo a cuatro objetivos: Acceso, Infraestructura, Preservación, y Factibilidad (Lee, 1999).

Por varios años, las bibliotecas de investigación, en particular, se han interesado en los proyectos de factibilidad e infraestructura. Estos son parte importante del esfuerzo colectivo por probar y diseminar herramientas, procedimientos, y metodologías. Los administradores en todas las organizaciones están más interesados en monitorear los procesos de los primeros proyectos de digitalización con el objetivo de llevar a cabo análisis de costo-beneficio. La experiencia obtenida en la realización de proyectos dentro de la misma institución ayuda a que las organizaciones comprendan el costo adicional que implica no sólo la creación de las colecciones digitales, sino también el mantenimiento y la entrega de las mismas.

A continuación citamos una de esas experiencias de la vida real en la administración de proyectos digitales que ilustra cómo diferentes objetivos institucionales pueden conducir a diferentes filosofías con respecto a la creación de colecciones electrónicas (o *viceversa*):

Cuando evaluamos nuevas tecnologías para la reconversión de formato, podemos 'mantenernos dentro de lo más sencillo' si trabajamos con grandes cantidades de materiales con pocos problemas antes de trabajar con cantidades más pequeñas de materiales que presentan grandes dificultades o problemas (Waters, 1999).
Si un proyecto académico electrónico no puede fracasar y no produce más ignorancia, entonces el proyecto no vale un comino (Unsworth, 1999).

En el caso anterior se aplican los principios KISS, y la lógica es que resolver problemas pequeños ayuda a las instituciones a prepararse para enfrentar problemas mayores. En este último caso, los problemas mayores son más atractivos, pues el fracaso seguro representará en si mismo un paso significativo hacia el desarrollo del conocimiento experto.

La experiencia también puede producir beneficios tangibles. Entre estos se incluyen:

- Procedimientos y directivas de distribución para uso de la institución o los contratistas
- Integración de nuevos sistemas de gestión digital con los catálogos y bases de datos existentes
- Instituir servicios locales, como reservas electrónicas, para mejorar los tradicionales
- Recuperación de espacio

- Establecer relaciones contractuales con socios editores quienes prestarán asistencia en la creación y/o distribución de contenido digital (vea la Barra de Costos más adelante en este capítulo)
- Facilitar las relaciones públicas y la recaudación de fondos
- Generar ingresos (al menos si no se recupera completamente el costo) a través de la venta de productos (*ej. de museos*).

La planificación del proyecto: creación de un plan de trabajo y un presupuesto

Plantearse objetivos o metas, representa la primera etapa, la de pensar o diseñar el proyecto, y un buen administrador sabe cuándo debe pasar a la transición de la planificación, la segunda etapa.

Si un departamento o institución fuera a llevar a cabo un solo proyecto -- y a buscar el financiamiento necesario -- entonces sería posible pasar por alto la etapa de planificación e ir directamente al trabajo en si mismo. No obstante, el tiempo invertido en escribir los documentos del plan de trabajo se amortizará con buenos resultados durante la etapa de producción. Estos documentos son escalones intermedios también fundamentales que dan paso del primer proyecto al segundo proyecto y después al tercero. Si los proyectos se publican, pueden servir de guía a otras instituciones para planificar sus proyectos de conversión a formato digital (Library of Congress, 1999). Entre los documentos de planificación se incluyen:

- Solicitud de financiamiento (*es decir,* la solicitud del otorgamiento de fondos)
- Una solicitud de información (SDI)/Solicitud de propuesta (SDP) en caso de que el trabajo de conversión de formato se vaya a contratar fuera de la institución
- Descripciones del trabajo a realizar
- Manuales de procedimientos (u hojas de instrucciones) para selección, manipulación, escaneo, creación de metadatos y control de la calidad
- Diagramas de flujo y otros diagramas de producción o de trabajo
- Listas de los elementos de datos
- Plan de trabajo y presupuesto de trabajo.

Desde el punto de vista interno, estos primeros documentos de gestión quizás sean los productos más importantes que surgirán de un proyecto. Algunos de los documentos, como por ejemplo las SDP o los contratos, tendrán un impacto directo en la calidad del producto. En comparación, el plan de trabajo tendrá un impacto directo en los procesos de inicio, ejecución y terminado del proyecto.

Hay algunos elementos que son esenciales para el plan de trabajo, independientemente de la naturaleza de los originales o de las exigencias de los usuarios principales a los que está dirigido. A continuación les presentamos cinco preguntas cuyas respuestas específicas les ayudarán a garantizar que encuentren menos problemas cuando comience el trabajo.

(1) ¿Quién va a realizar el trabajo?

En la práctica, esta es la primera pregunta porque muchas de las tareas las realizarán personas que ya trabajan dentro de la organización. Cuando se trata de crear un equipo de trabajo, quizás sea mucho mejor hacer una encuesta y preguntar dentro de la misma organización "¿Quién puede realizar este tipo de trabajo?" o "¿Quién tiene los conocimientos adecuados y las habilidades para participar en un proyecto de conversión a formato digital?"

La segunda fase de la organización del personal que integrará el equipo de trabajo es determinar cuántos especialistas se necesitarán. Siempre deberá tenerse en cuenta que será necesario contratar a alguien para que se haga el trabajo. No importa lo simple que parezca el trabajo o lo pequeño que pueda ser, siempre será bueno tener como regla "el proyecto siempre es más que lo que puede hacer una persona". Los proyectos de mediana envergadura exigen que varios departamentos trabajen juntos. Los proyectos de gran escala exigen coordinación entre múltiples agencias, instituciones, burós de servicios y asociaciones publicitarias.

Los proyectos de gran escala no sólo necesitan múltiples puestos de trabajo, sino además varias personas con el conocimiento experto en cada categoría de trabajo. Los proyectos pequeños, sin embargo, no necesitarán una decena de empleados a tiempo completo, pero si el trabajo se va a ejecutar con niveles de responsabilidad razonables, alguien deberá asumir ese papel. (Naturalmente, varios de estos trabajos se pueden subcontratar.) Cada uno de los puestos siguientes, o tareas, es demasiado importante para que se excluya de un proyecto que tiene como objetivo convertir el formato de determinados materiales, mantenerlos por un período de tiempo razonable, y ponerlos al acceso de los usuarios por vía de las redes de computación (o incluso en forma de CD-ROM).

Personal que debe integrar el proyecto -- Papel de cada uno

- Administrador del proyecto
- Seleccionador
- Conservador, curador, u otro analista de la fuente o material original
- Técnico a cargo de los preparativos (puede también ser curados, que, entonces también puede ser el seleccionador)
- Catalogador para crear o mejorar los registros bibliográficos y para que se encargue de extraer los materiales para la conversión
- Fotógrafo o técnico encargado del escaneo
- Técnico de control de la calidad (puede ser también el técnico de escaneo)
- Analista de los metadatos (puede también ser catalogador)
- Técnico que introduzca los datos
- Programador u otro experto en bases de datos que integre los metadatos y las imágenes en recursos coherentes (conocidos también como el objeto digital)
- Administrador de los sistemas u otro administrador de registros electrónicos y sistemas
- Administrador de la red para poner en práctica la seguridad y otros requisitos de acceso (puede también ser el administrador de los sistemas)
- El que desarrolla o diseña la interfase del usuario

Para cada uno de los puestos de trabajo del proyecto, debe decidir si se necesita algún entrenamiento, en qué momento pudiera ser necesario (o es obligatoriamente necesario), y cuánto costará.

(2) ¿Qué sistemas será necesario utilizar o desarrollar durante el proyecto?

En este contexto, sistemas se refiere a software, hardware, y a las buenas y anticuadas instalaciones de ladrillo y cemento que se necesitan para almacenar los medios. Aunque son muy flexibles, los productos digitales son objetos que hay que colocar en alguna parte. Es importante especificar antes de comenzar el trabajo, dónde se van a almacenar los objetos digitales, cuánto tiempo van a residir en ese lugar en forma leíble y accesible, y quién será la persona responsable de los mismos.

Los requisitos del software y el hardware varían, pero el número de sistemas será directamente proporcional al número de procesos y tareas especificadas para estar bajo el control local. En otras palabras, las

capacidades de la infraestructura local definen los límites del trabajo que se puede ejecutar dentro de la propia institución.

Tenga en cuenta las consecuencias a largo y mediano plazo que pueden resultar de decisiones con respecto a hardware y software que Ud. se sienta inclinado a tomar en nombre de las necesidades de un proyecto de corta duración. ¿Estaría Ud. dispuesto a construir sistemas desechables? ¿Sería aceptable abandonar las aplicaciones comunes cuando se marcha un programador? ¿O se necesitarán soluciones comerciales?

(3) ¿Cuáles son las especificaciones técnicas para los archivos de imágenes y los metadatos?

Las imágenes digitales y los metadatos asociados a las mismas (en una serie de categorías) incluyen el material bruto de las bases de datos de imágenes. Si se desea una consistencia en la búsqueda y en la presentación, entonces es esencial que se hagan obligatorias las especificaciones técnicas para los elementos de datos, imágenes, formatos y protocolos de acceso. Estas especificaciones se hacen mucho más importantes cuando se desea lograr inter-operabilidad o interacción con otras colecciones.

El Capítulo VII describe alguna de las prácticas y las especificaciones que se emplean hasta la fecha. Si no se puede permitir especificaciones exactas antes de escanear o catalogar los materiales, el plan del proyecto deberá por lo menos plantear las opciones que se podrán tener en cuenta.

(4) ¿Cuánto costará el proyecto?

Las evidencias empíricas que hemos recogido en nuestras propias colecciones son más convincentes que las anécdotas de otros proyectos. Una de las mejores formas de predecir el costo de un proyecto es crear una muestra representativa de los materiales seleccionados para la reconversión. En muchos casos, será suficiente con cinco o seis objetos. Si se va a escanear con medios no propios, se puede finalizar el presupuesto del proyecto sólo cuando se ha pasado una muestra por un flujo de trabajo completo -- escaneo, procesamiento, creación de metadatos (incluido el texto completo), y control de la calidad -- y los participantes adecuados en el proyecto han inspeccionado y aprobado los resultados. Muchos proveedores están dispuestos a ofrecer este servicio como parte de su respuesta a SDI o SDP para competir por un contrato.

Todas las actividades que se llevan a cabo dentro de la institución deberán declararse como costos del proyecto o costo compartido. Realizar una especie de recorrido de avance de todos los flujos de trabajo propuestos puede revelar de inmediato hasta qué punto el administrador del mismo ha previsto bien el proceso, desde que comenzó a engendrarse

el mismo hasta la realidad. Pueden ocurrir sorpresas e imprevistos. Puede que se necesite mucho tiempo para retirar los materiales del almacenamiento y embalarlos para enviárselos al proveedor de los servicios; y por supuesto, deberá duplicarse ese tiempo para permitir el proceso inverso de retorno. Los masters digitales que Ud. inspeccionó durante una comprobación del 100% del control de la calidad pueden demorarse cinco minutos en la computadora que Ud. ha dispuesto para esa tarea. Los registros de catálogos que parecían adecuados en la revisión inicial necesitan actualización o información adicional. Trate de identificar de antemano dónde se pueden presentar cuellos de botella en la producción y asegúrese de que los niveles del presupuesto y de personal en el plan de trabajo tengan en cuenta esas contingencias.

Por último, tenga en cuenta también el impacto del cronograma en todos los costos del proyecto, sobre todo los costos del personal en los casos en que los salarios y los beneficios salariales deben programarse por periodos fijos.

(5) ¿Quién será el propietario y el administrador de los productos digitales que se producirán?

Esta pregunta se aplica al personal, al flujo de trabajo y al presupuesto. Como señalara un artículo del *New York Times* en abril de 1999, las cuestiones referentes al almacenamiento "no se pueden resolver sin analizar la cuestión de la propiedad". En el mismo artículo, Ann Okerson de la Universidad de Yale observó, "No sé cómo alguien puede preservar algo de lo que no es propietario" (Hafner, 1999). La sección de *Costos,* al final de este capítulo, muestra el impacto financiero de asumir las responsabilidades de almacenamiento y entrega de las colecciones digitales. Debido a los altos costos de la infraestructura necesaria para administrar y distribuir objetos digitales, quizás no sea sorprendente que hayan surgido una serie de asociaciones entre bibliotecas universitarias y editoriales. (*Early American Fiction* publicada por Chadwyck-Healey y la Biblioteca de la Universidad de Virginia es un ejemplo representativo.) Si una institución desea tener la propiedad y distribuir las reproducciones digitales que ha creado durante un período de tiempo, será importante que articule estos objetivos en un proyecto con un plan, que adquiera los sistemas y prepare el personal necesario para administrar ese proyecto y que garantice que la institución o quien(es) financia(n) el proyecto apoyen también plenamente esos componentes.

La ejecución del proyecto: La gestión del flujo de trabajo

La tercera y última etapa de la gestión del proyecto es la ejecución del mismo. Prácticamente todos los proyectos de conversión exigen que se preparen varios diagramas de flujos de trabajo y que se administren los mismos. Una excepción podría ser un proyecto clave realizado dentro de la propia institución con codificación de luz. Los proyectos se terminarán más pronto si se organizan las tareas en paralelo o superpuestas en lugar de organizar un flujo de trabajo lineal. Puede que los costos sean lo primero en el presupuesto del proyecto y otros documentos del plan, pero por el tiempo es por lo que debemos rendir cuentas en la gestión del trabajo mismo de conversión de las colecciones.

Es típico que las siguientes actividades se separen en flujos de trabajo diferentes. Puede ser que distintos individuos o departamentos terminen encargándose de cada actividad.

- Selección
- Autorizaciones de derecho de autor y otras investigaciones referentes a derechos y permisos; creación de metadatos de derechos y permisos
- Preparación que incluye la evaluación de conservación y/o tratamiento si fuera necesario
- Creación de registros de catalogación, medios de búsqueda y otros punteros para un objeto o colección digital (metadatos descriptivos)
- Producción de imagen digital: Escaneo de los originales o fuentes para crear los masters digitales y metadatos técnicos asociados; procesamiento de imágenes masters para crear derivados para la pantalla o para imprimir
- Comprobación y control de la calidad para los originales y los masters digitales; transferencia de las fuentes a la nueva ubicación o a la original; reubicación de materiales, actualización de registros de catálogos según sea necesario
- Creación de metadatos estructurales
- Creación de texto completo, incluido el marcaje
- Gestión de archivos: Introducción de los datos en el depósito
- Integración de los masters digitales y los metadatos en bases de datos en bases de datos de imágenes; creación de hipervínculos asociados a los registros de los catálogos u otros puntos de acceso
- Entrega: puede variar desde páginas web artesanales hasta sistemas altamente automatizados
- Publicidad, promoción, evaluación por los usuarios.

El trabajo debe fluir de manera más o menos cronológica según la lista anterior si los materiales no fueron segregados en lotes durante la producción. Seleccionar el tamaño adecuado d cada lote y seguir el progreso de los mismos con cuidado desde el principio hasta el final es la responsabilidad principal del administrador del proyecto. La recopilación e información de las estadísticas de producción, de los registros de los problemas, de la retroalimentación del personal y de los gastos son todos indicadores de una gestión eficaz.

Cuando se piensa en la conversión a formato digital desde la perspectiva de la práctica, es fácil apreciar la eficacia de trabajar en lotes de materiales similares siempre que sea posible. Esto es válido para la catalogación al igual que para el proceso de escaneo. Si los materiales no se pueden agrupar en categorías similares, con frecuencia será mejor estructurar el trabajo en una serie de pasos, en los que los técnicos se concentran en tareas específicas por períodos significativos ininterrumpidos. Con la configuración adecuada de la instalación donde se va a ejecutar un proyecto, se les puede dar la oportunidad a los trabajadores de romper el ciclo repetitivo de una tarea (como el escaneo, por ejemplo) realizando otra tarea (por ejemplo, control de la calidad y creación de metadatos). Esta práctica facilita una producción elevada y ayuda a garantizar una calidad constante.

Lineamientos y prácticas recomendables para la gestión

Con frecuencia uno escucha acerca de la necesidad de prácticas o lineamientos para conversión a formato digital. En el área de la gestión del proyecto, la primera medida para una mejor ejecución probablemente sea que el fin justifica los medios. Si las reproducciones digitales son bien recibidas y se han realizado en un momento oportuno y de manera eficiente en cuanto a los costos, naturalmente se considerará que el proyecto ha sido un éxito. Otra medida de éxito, que se recibiría de instituciones similares y otros colegas, se podría ver en la calidad de la documentación recopilada durante la ejecución del proyecto. Registrar en documentos la justificación, la metodología, los sistemas, los modelos de captación del personal, los costos, y lo más importante, las lecciones aprendidas del proyecto ayudan a que toda la comunidad (*es decir,* a la institución misma, a los que aportaron el financiamiento y a otros que realizan ese mismo tipo de proyecto) se beneficie de la experiencia obtenida en un solo proyecto. El administrador del proyecto ha realizado bien su trabajo si las personas que trabajaron en el proyecto han tenido una experiencia satisfactoria y si el futuro(s) administrador(es) de la colección digital pueden interpretar fácilmente las razones por las cuales se crearon las cosas en una forma específica y

qué es necesario hacer para mantener, o incluso mejorar, esta primera generación de objetos digitales.

Costos

Es difícil generalizar con respecto a los costos debido al amplio espectro de procesos y productos digitales que existen. Incluso cuando los originales o fuentes y las reproducciones digitales sean comparables, las inversiones pueden variar de manera considerable en actividades como planificación y gestión del proyecto, así como en la ~~infraestructura para almacenar y entregar los objetos digitales.~~

Costos de conversión

La conversión es una actividad limitada, que por lo general se contrata a especialistas fuera de la institución. Los costos de escaneo, del OCR, del marcaje del texto, y de la fotografía digital son relativamente predecibles. Las tendencias durante los últimos años sugieren que los costos de conversión de texto e imágenes permanecerán estables o ascenderán ligeramente -- no disminuirán -- aunque la calidad del producto puede mejorar en varias áreas. Las cifras básicas que se ofrecen a continuación deberán considerarse como puntos de partida muy reales para preparar el presupuesto. Los incrementos sobre esta base son aproximaciones del impacto de las variables combinadas que introduce la naturaleza de las fuentes u originales, así como por las especificaciones técnicas de los objetos digitales.

Producto	**Precio básico**	**Factores de costo significativos**	**Incremento del producto(sobre el básico)**
Páginas de imágenes	$ .25/página	- tamaño (dimensiones de página) - formato (papel < microfilm) - montadura (se retira < intacta)- profundidad de bit (b/n < escala de grises < color) - metadatos (descriptivos y estructurales)	2-6X (1-bit) 4-25X (8-24 bit)
Texto completo	$ .50/página*	*Igual que el anterior, más* - nivel de exactitud necesario - extensión del marcaje ("ligero" a completo SGML)	6X + (clave) 2X + (marcaje)

imágenes			
	$3.00/imagen	tamaño (dimensiones de los originales) - requisitos de manipulación- tono/color requisitos de reproducción - metadatos	2-20X (pictóricas)

* incluye el costo de las imágenes de página

Costos totales del proyecto

Al destacar el aspecto de que la conversión a formato digitales sólo uno de los pasos que conducen a la entrega de los archivos digitales, el Internet Library of Early Journals Final Report afirma que los costos "por imagen de página indizada a la que se puede acceder en Internet" es aproximadamente siete veces mayor que el costo unitario de escanear y de los OCR sin corregir (Vea la Nota más adelante). Como las bibliotecas y archivos se encuentran en medio de un proceso muy activo de integración de las tecnologías digitales en sistemas de adquisición y catalogación, e integrar incluso los departamentos de preservación, quizás sea válido tener en cuenta estas actividades como costos fuera de un proyecto de conversión de formato. No obstante, es importante admitir que la publicación de objetos convertidos de formato analógico a digital (incluida la distribución de los mismos) requiere inversiones significativas -- JSTOR y el Programa Nacional de Biblioteca Digital de la Biblioteca del Congreso nos ofrecen dos ejemplos, aunque a gran escala -- para desarrollar e integrar nuevos sistemas, servicios y conocimiento experto. Los editores comerciales están dispuestos a ofrecer estos servicios, pero los términos de dichos acuerdos deben ser revisados con gran cuidado para balancear los intereses del presupuesto del proyecto con los de propiedad y control sobre la colección.

Nota: "Internet Library of Early Journals (enero 1996 - agosto 1998), Un proyecto en el programa eLib, Informe Final", marzo, 1999. [http://www.bodley.ox.ac.uk/ilej/papers/fr1999/] (24 de octubre de 1999). Vea, párrafo 80: "El total de £458,000 [aproximadamente $757,395 USD] representa un gasto de £4.21 [aproximadamente $6.96 USD] por imagen de página indizada accesible a través de Internet. Este estimado de gastos no tiene en cuenta los costos de la contribución de IT y las infraestructuras bibliotecarias de las cuatro instituciones". Vea también la página E14.

Bibliografía

Borghuis, Marthyn, *et al.* TULIIP Final Report. Elsevier Science, 1996. Vea Apéndice X, "Checklist of aspects to be considered for the implementation of a 'digital library,' 337-44.

Chapman, Stephen y Anne R. Kenney. "Digital Conversion of Research Library Materials: A Case for Full Informational Capture." Vea Tabla 4. *D-Lib Magazine* (octubre 1996). [http://www.dlib.org/dlib/october96/cornell/10chapman.html] (noviembre 4,1999).

Ester, Michael. *Digital Image Collections: Issues and Practice,* pp. 10-12.Washington, DC: Commission on Preservation and Access, 1996.

Hafner, Katie. "Books to Bytes: The Electronic Archive, Research Libraries Grapple With the Difficult Task of Preserving the Digital Present." *The New York Times* en Internet. 8 de abril, 1999. [http://www.nytimes.com/library/tech/99/04/circuits/articles/08arch.html] (4 de noviembre de 1999).

Internet Library of Early Journals (enero 1996-agosto 1998), A project in the eLib programme, *Final Report.* marzo 1999. [http://www.bodley.ox.ac.uk/ilej/papers/fr1999/] (24 de octubre de 1999).

Kenney, Anne R. "Digital to Microfilm Conversion: A Demonstration Project, 1994-1996, Final Report to the National Endowment for the Humanities, PS-20781-94." Vea también otras publicaciones sobre la metodología utilizada por el Departamento de Preservación y Conservación de la Universidad de Cornell. [http://www.library.cornell.edu/preservation/pub.htm] (24 de octubre de 1999).

Lee, Stuart D. *Scoping the Future of the University of Oxford's Digital Library Collections, funded by the Andrew W. Mellon Foundation, Final Report.* Appendix D, 1. Oxford University, agosto 1999.

Por ejemplo, la Biblioteca del Congreso ha publicado muchos Folletos de Información Técnica, a los que se puede acceder en Internet, durante la creación de las colecciones American Memory. Vea Technical Operations Documentation y White Papers, n.d. [http://memory.loc.gov/ammem/ftpfiles.html] (24 de octubre de 1999).

La División de Acceso y Preservación del NEH, por ejemplo, acepta solicitudes para proyectos de conversión a formato digital, pero dichas solicitudes deberán regirse por la norma de que "... la digitalización [tiene que] mejorar de manera significativa el acceso a la colección y las formas en que [la misma] podrá ser utilizada para la educación, por los académicos o para la programación pública". Vea "Considerations for Reviewers", agosto 1999.

Noerr, Peter. *The Digital Library Tool Kit.* Sun Microsystems, Inc. (abril 1998):21. [http://www.sun.com/products-n-solutions/edu/libraries/digitaltoolkit.html] (24 de octubre de 1999).

Puglia, Steven. "The Costs of Digital Imaging Projects." *RLG DigiNews.* 3:5 (15 de octubre de 1999). [http://www.rlg.org/preserv/diginews/diginews3-5.html#feature] (4 de noviembre de 1999).

Thomas, Timothy. "Physical Review Online Archives (PROLA): An Image Archive for the Journal *Physical Review," D-lib magazine* (junio 1998). [http://www.dlib.org/dlib/june98/06thomas.html] (24 de octubre de 1999). El Sr. Thomas afirma que "un archivo electrónico no es estático bajo ningún concepto. [El mismo] requiere modificación constante para mantenerse al día con la velocidad actual de los cambios tecnológicos".

Unsworth, John. "The Importance of Failure," *Journal of Electronic Publishing* 3:2 (diciembre 1997). [http://www.press.umich.edu/jep/03-02/unsworth.html] (24 de octubre de 1999).

Waters, Donald J. *Electronic Technologies and Preservation.* Washington, DC: Commission on Preservation and Access, 1992. [http://www.clir.org/pubs/ reports/waters/waters2.html] (24 de octubre de 1999).

Buscar Indice Sitio del NEDCC

Northeast Document Conservation Center

100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Modificado por última vez, 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

IV

La selección de los materiales que se van a escanear

Diane Vogt-O'Connor

Elegir es un problema. – Antiguo proverbio holandés

Introducción

Seleccionar implica elegir entre una serie de opciones utilizando un juicio inteligente y criterios de selección. Buenas técnicas de selección garantizan que los recursos se inviertan de manera inteligente para digitalizar las colecciones más significativas y útiles con el menor costo posible sin que exista riesgo legal o social para la institución. Una mala selección conduce a la digitalización de materiales que no tienen utilidad o son de poco valor.

La selección es un proceso familiar para archivistas, bibliotecarios y curadores que tienen que:

- Valorar materiales para adquisición
- Determinar las prioridades de salvamento de materiales durante una situación de emergencia o para la conservación
- Seleccionar contenidos para exhibición y publicaciones.

Durante la realización de estas tareas, el personal adopta decisiones que afectan de manera significativa la vida y la accesibilidad de los contenidos de las colecciones. Cuando se revisan materiales para valorarlos, considerar las prioridades de conservación o exhibiciones, el personal tiene en cuenta factores tales como:

- Adecuación de la misión del depósito y de los objetivos de las colecciones

- Adecuación a un objetivo más amplio, por ejemplo las metas del consorcio
- Valor en comparación con otros materiales que posee la organización
- Demandas de los usuarios en comparación con otros materiales que posee la organización
- Restricciones debido a la condición legal de las colecciones
- Disponibilidad para su uso debido a su ubicación en depósitos remotos o almacenamientos refrigerados, debido a malas condiciones de preservación, o dimensiones de difícil manipulación (*ej.* , planos arquitectónicos de 40" x 60") o formato (*ej.* Negativos de vidrio u objetos en marco o montados en un bastidor).

La selección de materiales para escanear implica que se tomen en cuenta estos factores, así como otros que se analizan más adelante. El proceso para seleccionar lo que se va a digitalizar puede ser muy similar al proceso de selección para otros propósitos. Este capítulo propone tres etapas.

- ***Nominación.*** La nominación implica una amplia participación de los creadores de la colección, de los donantes, de los investigadores, de los administradores, de grupos bien documentados y otros. Tanto los depositarios como el personal nominan los materiales para que se incluyan o no se incluyan, e indican en su propuesta las razones por las que deben seleccionarse o excluirse esos materiales.
- ***Evaluación.*** Durante la evaluación, un Comité de Selección revisa las nominaciones basándose en los criterios planteados y toma decisiones acerca de la inclusión o la exclusión de los materiales.
- ***Priorización.*** Durante la etapa de priorización, el comité otorga un rango a los materiales seleccionados según su valor, su uso y los riesgos a los que están expuestos, de manera que los materiales se digitalizan en orden según su valor para el depósito.

Los que utilizan este proceso de tres etapas identifican y descartan los materiales que pueden resultar problemáticos, a la vez que seleccionan y otorgan prioridad a los materiales adecuados para el trabajo de digitalización. Con este enfoque se logra un flujo de trabajo estable independientemente de la cantidad de fondos disponibles.

El proceso de selección	**¿Quién lo hace?**	**¿Cuándo?**
1. Nominación de los materiales: A. Nominación de qué se va a digitalizar (Modelo A) B. Nominación de qué no se va a digitalizar (Modelo B)	Depositarios de la colección y personal	Al comienzo
2. Evaluación de los materiales nominados: Se revisa la selección	Comité de selección	Durante la revisión formal

y se descartan los materiales entre los materiales nominados (Modelo C)		
3.Priorización de los materiales restantes: Dar rango a los materiales basándose en el uso, y los riesgos	Comité de selección	Durante la revisión formal

¿Cuál es el objetivo de seleccionar para escanear?

Muchas organizaciones se pueden sentir tentadas a decir "vamos a digitalizarlo todo y ya" porque la selección puede parecer un proceso de labor intensa y costosa. ¿No sería mejor entonces emplear los recursos en digitalizar más contenido? Varias razones poderosas para realizar la selección nos indican lo contrario.

- ***Acceso Web como publicación.*** Cuando se montan en Internet, las colecciones digitalizadas llegan a los más grandes y as diversos públicos. El público que accede a Internet incluye académicos, estudiantes, público general, periodistas, donantes, creadores de colecciones, cineastas, editores y especialistas de las disciplinas así como otros usuarios menos deseables como posibles ladrones, vándalos y los que infringen las leyes de la propiedad intelectual. Una vez que se les ha alertado acerca de las colecciones que posee una organización, este público puede exigir reproducciones de mayor calidad, permisos de publicación y acceso a los originales. La seguridad de las colecciones y las políticas de gestión para duplicación, publicación y acceso deberán implementarse mucho antes de que la visibilidad de la colección mejorada genere un aumento de las solicitudes de acceso y de las visitas. Una eficaz planificación previa puede mitigar el impacto del aumento de las solicitudes de acceso, duplicados y permisos de publicación o exhibición de los objetos digitalizados.

- ***Costos elevados y limitados presupuestos.*** La mayoría de las organizaciones carecen de recursos para digitalizar todas sus colecciones, por muy aconsejable que esto pudiera ser para mejorar el acceso a las mismas. Los costos iniciales de selección y digitalización no son nada comparados con los costos de control de la calidad, producción de metadatos e indización/catalogación. Cuando no podemos afrontar el costo de digitalizar toda la colección, no sólo la selección es factible, sino que es, además, esencial

- ***La hipoteca digital.*** Los ficheros digitales vienen acompañados de una hipoteca. Cada organización debe preparar un presupuesto para transferir los ficheros viejos a nuevos formatos según cambian los softwares y el hardware y los medios electrónicos llegan al fin de su relativamente corta esperanza de vida (Puglia, 1999). Este presupuesto no sólo financia

la sustancial inversión inicial necesaria para la digitalización, sino también una infraestructura digital que incluye el personal que trabajará, los contratos, el equipamiento y el software. Como los ficheros digitales son mucho más costosos de manejar con el tiempo que los archivos en soporte de papel, las organizaciones deberán identificar y programar cuáles son los significativos recursos necesarios para la manipulación de los grandes volúmenes de archivos digitales que se han creado (Lowry & Troll, 1996; Marcum, 1998). Los costos de los programas no terminan cuando aparece la página Web.

- ***Aspectos legales.*** Con frecuencia, las organizaciones carecen de los derechos y permisos de propiedad intelectual para los materiales que poseen. Los derechos de autor; los derechos de publicidad y privacidad; y las cuestiones de obscenidad, difamación, e información sobre ubicación protegida por las leyes (cavernas protegidas por el gobierno federal, sitios sagrados, pozos, y sitios arqueológicos), si no se tratan de manera adecuada, pueden dar lugar a demandas en los tribunales y a costosos arreglos legales. Si la digitalización no se hace mediante una cuidadosa selección, no se podrá utilizar de manera eficaz debido a restricciones legales (Smith Levine, 1995). Determinar la situación legal de los materiales candidatos es un paso crucial en cualquier proceso de selección para digitalización. La solicitud de los permisos para digitalizar los materiales seleccionados comienza inmediatamente después de la selección.

  Proporcionar acceso amplio a materiales sensibles en Internet sin consultar con los especialistas, los grupos afiliados y los donantes puede resultar en daños a las fuentes originales (si se divulgan locaciones protegidas) o a las relaciones de la institución con la comunidad de donantes. Cuando se seleccionan materiales para ponerlos a acceso en Internet, constituye una política inteligente y además es una buena costumbre realizar una catalogación y contextualización precisas, y consultar con los grupos de depositarios. Publicar materiales en Internet no es equivalente a proporcionar el acceso al material a varios cientos de personas anualmente en una sala de lectura, sino que está más cercano en el impacto que puede producir a un documental presentado en la televisión o a un reportaje en los periódicos. Si hay serias dudas con respecto a publicar algo en la primera plana de un diario, los materiales no deberán introducirse en Internet.

- ***Intereses de los depositarios.*** Publicar en Internet materiales que puedan ser sensibles desde el punto de vista cultural o ético pudiera crear una tormenta de protestas, incluso si la colección no tiene restricciones legales. Sus decisiones deberán balancear la necesidad de respetar los intereses válidos de los depositarios con el deseo de evitar una censura arbitraria. Consulte con los depositarios cuando digitalice los siguientes tipos de materiales:

- *Información cultural* cuyo acceso está normalmente está limitado a miembros de grupos o culturas especiales, como por ejemplo, ceremonias sagradas, lugares de enterramientos, sagrados y/o de recolección de alimentos y subsistencia, u otros conocimientos generales que pueden provocar daños en los recursos si se difunden.

- *Materiales restringidos por los donantes,* por ejemplo solicitudes de evitar la publicación de datos por un período de 25 años después de la muerte del donante.

- ***Documentación.*** Durante la selección, el comité determina si los materiales candidatos están bien documentados en términos de subtítulos individuales o de datos de catalogación completos y exactos a nivel

de material individual. De no ser así, la organización se está comprometiendo a un significativo gasto adicional en la investigación correspondiente a cada material, comprobación de toda la información obtenida, y a escribir las descripciones que acompañarán al material y que proporcionarán un contexto claro y adecuado al mismo. De hecho, el personal que labora en el Archivo Nacional considera que hasta dos terceras partes del costo real de la digitalización durante los diez primeros años de un proyecto de digitalización de imágenes corresponde a la creación de metadatos y al control de la calidad (Puglia, 1998). Los textos de exhibición son aún más costosos, y se asemejan más al costo de un libro tradicional o a los costos de preparación de exhibiciones.

- ***Credibilidad institucional.*** Durante la selección, el comité comprueba la exactitud y autoridad de toda la información que se deberá incluir junto con los recursos electrónicos. La institución rectifica la información incorrecta para proteger su credibilidad y su reputación. Los criterios más comunes incluyen la autoridad de una organización, los antecedentes y la reputación del creador individual y las notas o citas verificables que se incluyen en los materiales.

Muestreo estadístico

Independientemente de cuán sistemático sea el proceso de selección y cuán bien definidos estén los criterios, no hay dos comités que seleccionen los mismos materiales o exactamente en la misma forma, como tampoco hay dos grupos que posean la misma experiencia, las mismas capacidades, la misma preparación y los mismos puntos de vista. Algunos historiadores y científicos prefieren digitalizar un por ciento aleatorio de las colecciones que pueda reflejar mejor y de manera más precisa todo el universo de materiales. Si una parte significativa de los posibles usuarios incluye académicos que trabajan con técnicas de muestreo, piense en incorporar el muestreo en el proceso de selección.

El muestreo, aunque menos eficaz para preservar el contenido académico y el flujo narrativo, puede ser:

- Más objetivo
- Más eficaz para presentar un corte de todo el cuerpo del material
- Menos dominado por las tendencias académicas.

Al introducir el componente de muestreo en el proceso de selección, base su muestreo en una tabla de números aleatorios. Utilice la tabla de números aleatorios para seleccionar los contenedores y artículos en particular o combine el muestreo con un enfoque de selección más tradicional. Por ejemplo, seleccione el 5% de una colección utilizando técnicas de muestreo para dar le a los investigadores una muestra del todo. Como casi seguro que 2/3 de los materiales muestreados serán excluidos a causa de su estado de conservación, restricciones u otros problemas, seleccione muchos más objetos que los que en realidad se necesitan.

El Comité de Selección

Por su propia naturaleza, la selección es una labor muy intensiva de aplicación de conocimientos y capacidades. Un individuo solo no puede dominar todo el conocimiento necesario de la materia, las leyes, la conservación, la educación, la tecnología, y así

sucesivamente. La mayoría de las organizaciones que cuentan con una buena administración forman Comités de Selección para evitar cometer costosos errores. Los miembros del comité de selección, de conjunto, pueden evaluar los materiales candidatos de manera inteligente y desde varios puntos de vista, disciplinas y perspectivas.

Los candidatos para el Comité de Selección incluyen:

- Especialistas de las disciplinas relacionadas con los temas que se abarcan en el proyecto
- Especialistas en educación, adecuados al nivel del público al que está dirigido el proyecto
- Especialistas en digitalización, como por ejemplo personal especializado de fotografía y laboratorio digital
- Bibliotecarios, archivistas y curadores
- Investigadores con experiencia en el trabajo con recursos en línea
- Conservadores y especialistas en preservación
- Abogados.

Incluso un Comité de Selección grande puede necesitar ayuda adicional. Por lo tanto, una organización que actúa con sabiduría puede discernir cómo enfrentar los conocimientos y las habilidades de las comunidades de depositarios durante el proceso de nominación. (Nota: También se pueden incluir en los Comités de Selección a algunos representantes especialmente eficaces de esas comunidades.)

¿Para qué involucrar a los depositarios en los Comités de Selección?

Si está financiado por el erario público, el depósito donde se almacenan las colecciones es responsable ante cada uno de estos grupos por la forma en que la institución administra, preserva y hace posible el acceso a las colecciones. Los depósitos privados que no están financiados por los fondos federales o del estado pueden, legalmente, no tener en consideración a algunas de estas comunidades. No obstante, ignorar a una comunidad de depositarios durante el proceso de selección es algo profundamente erróneo, ya que afecta la reputación de la institución y puede afectar las donaciones futuras y la cooperación de la comunidad.

Los criterios de evaluación se pueden

adaptar a varios esquemas institucionales. Los criterios que se expresan en este capítulo están diseñados fundamentalmente para museos.

Entre los depositarios que participan en la nominación se incluyen:

- Los creadores de las colecciones y los donantes que poseen un conocimiento profundo de los recursos que se van a digitalizar
- El personal de la institución que maneja la colección y conoce la historia del uso de los originales y la situación de su estado de conservación y catalogación
- Individuos y grupos representados en las colecciones y así como sus herederos y las comunidades a las que pertenecen, que pueden descifrar los significados ocultos y explicar las preocupaciones culturales que puede presentar el acceso a las colecciones
- Los usuarios, académicos, estudiantes, educadores, cineastas y editores académicos que conocen el mercado y el público y los usuarios secundarios de los archivos del proyecto y pueden ayudar a confeccionar una planificación
- El público en general que puede articular los intereses de la comunidad con las normas para compartir el conocimiento local en el ámbito internacional
- Los intermediarios, incluidos los editores de los programas digitalizados, los servicios en línea por encargo, la maquinaria investigativa y las compañías tecnológicas, las estaciones de televisión públicas y las compañías de *edu-entretenimiento* que han desarrollado estrategias innovadoras para la distribución de los archivos y han promocionado los contenidos para obtener financiamiento
- Los colegas profesionales, los especialistas de las disciplinas y los que financian el proyecto que conocen el contenido de la institución, así como y el papel y los objetivos de la misma.

El proceso de selección

Las personas que participan en el proceso de selección deben familiarizarse con los objetivos del proyecto y descubrir en qué forma el proyecto respalda la misión de la organización, los objetivos de la colección, el público al que está dirigido la misma y las estrategias cooperativas. En los capítulos anteriores se analizan esas importantes cuestiones.

Las tareas de pre-selección

Desde el principio, los miembros del Comité de Selección deberán reunirse con otras organizaciones, consorcios o grupos que estén encargados de la gestión de proyectos digitales. Obtengan los objetivos y planes de los proyectos de esas instituciones. Pidan las copias de los estudios realizados acerca del público al que están dirigidos los proyectos de ellas y las evaluaciones. Investiguen los criterios de selección utilizados por socios y cooperantes (*ej.* Los Criterios de Evaluación de las Colecciones de la Biblioteca Digital Nacional de la Biblioteca del Congreso). Examine otros proyectos de la organización, tales como exhibiciones, publicaciones y cursos para determinar si los planes y objetivos de la misma sugieren un componente digital. Basarse en el conocimiento experto de otros acelerará el desarrollo de lineamientos internos, políticas y procedimientos fuertes y sistemáticos.

Nominación

Durante la fase de la nominación, los depositarios y el personal utilizarán los modelos A y B para nominar los materiales que evaluará el Comité de Selección. El personal -- especialmente los bibliotecarios, los archivistas y los curadores -- deberán participar en la nominación de grupos de materiales que estén relacionados para la evaluación de los mismos. Cuando se nominan los materiales, el personal deberá tener en cuenta las siguientes cuestiones:

- ¿Qué parte de la colección está bien documentada a nivel de objeto individual en índices confiables y completos y en medios de búsqueda, y dónde se encuentran estos objetos bien documentados?

- ¿Qué parte de la colección se encuentra en condiciones de conservación buenas o estables y en qué lugar se encuentran esos materiales en condiciones estables?

- ¿Qué porción de la colección tiene un tamaño estándar, consistente, contraste normal, qué parte son materiales en blanco y negro y/o impresos, y en qué categoría se clasifican esos materiales? Nota: Evite trabajar con materiales de formato variado, poco usual o de gran tamaño y con materiales manuscritos para los proyectos iniciales.

- ¿Qué materiales se pueden facilitar a los investigadores debido a sus dimensiones, formato o requisitos de visionaje (por ejemplo, un documento mecanografiado en tamaño 8" por 10" o un original fotográfico impreso de alta resolución y alto contraste, pero no un negativo tomado con una cámara de circuito o un microfilme), y en qué lugar se encuentran en las colecciones?

- ¿De qué por ciento de los materiales la institución tiene el derecho de autor o la licencia para utilizarlo y dónde se encuentran los materiales de dominio público?

- ¿Qué por ciento de los materiales no tiene ningún tipo de restricción o no son sensibles en absoluto (como restricciones por privacidad, publicidad, difamación, obscenidades, y sensibilidades o restricciones impuestas por el donante), y dónde se encuentran esos materiales no restringidos o no sensibles?

- ¿Cuáles son los materiales de más alto valor monetario y mejor asegurados y en qué lugar se encuentran en las colecciones?

- ¿Cuáles son los materiales que se consideran en mayor riesgo y por qué y dónde se encuentran ubicados en las colecciones? De ellos, ¿cuáles están en un estado de conservación lo suficientemente estable como para ser escaneados sin que se produzca daño alguno o cuáles ya han sido fotografiados?

- ¿Qué materiales son los que se utilizan con mayor frecuencia, cómo se utilizan y dónde se encuentran ubicados?

- ¿Qué materiales son únicos para la institución y dónde se encuentran ubicados?

Como se señala anteriormente, si se tienen en cuenta una variedad de puntos de vista en las nominaciones y cuando se descartan los materiales, se garantiza un proceso de selección más equitativo y balanceado que refleje todo el rango de intereses de los usuarios y los académicos. Promueva la utilización de los Modelos A y B por parte del personal, los investigadores, los miembros del comité y otros depositarios para nominar a los materiales candidatos o para identificar los materiales que no se deben digitalizar. Los materiales deberán estar en la sala de investigación al alcance de todas las partes interesadas que estén familiarizadas con las colecciones. No obstante, deberá informarse a los nominadores que el Comité de Selección evalúa todas las nominaciones y las recomendaciones para descartar materiales y toma las decisiones finales.

Evaluación

Los criterios de evaluación se pueden adaptar a las condiciones de diversas instituciones. Los criterios que se utilizan en este capítulo están preparados fundamentalmente para museos.

El Comité de Selección comparará en primer lugar los modelos de nominaciones (Modelo A) y de descarte (Modelo B) para ver si alguno de los materiales aparece en ambos modelos. Marque cualesquiera materiales que hayan sido nominados y recomendados para descarte y fundamente por qué requieren especial atención. El comité trabaja a través de las preguntas que aparecen en el Modelo C para cada uno de los grupos de materiales nominados. Si la respuesta a cualquiera de las preguntas que aparecen en el Modelo C es "no, no digitalice", elimine al material candidato de la lista hasta que se pueda resolver el problema. Algunos materiales califican para digitalizar más adelante cuando se pueden cumplir determinadas condiciones -- por ejemplo, cuando se puede obtener el permiso de los que poseen el derecho de autor, o cuando caduca el período de derecho de autor, o cuando los modelos o los entrevistados que aparecen en los materiales los liberan para su

publicación. Mantenga un registro de dichos candidatos de manera que el Comité de Selección pueda considerarlos de nuevo en el futuro.

Si la cantidad de los materiales seleccionados no es la adecuada para la envergadura del proyecto de digitalización analice la posibilidad de añadir una muestra aleatoria de la colección como se indica anteriormente (*Muestreo estadístico*). Por ejemplo, si se han seleccionado 4,000 objetos y el financiamiento es para 5,000 objetos, utilice una tabla de números aleatorios para seleccionar los 1,000 que restan. Revise las muestras para comprobar que no existen problemas legales, de preservación, de sensibilidad y otros problemas al igual que con otros materiales. Cuando haga el muestreo, seleccione una cantidad extra (el triple de la cantidad que necesita) para dejar un margen de materiales que se van a descartar.

Una vez que el Comité de Selección ha completado su evaluación todos los materiales inadecuados deberán descartarse, y se debe dejar sólo los candidatos excelentes para la digitalización.

Muestra

Cómo evaluar las colecciones

Observe cada grupo de artículos nominados y responda las preguntas que aparecen en el Modelo C.
Nombre: Academia Quonsethut de Bellas Artes en New York
Misión y Objetivo de las colecciones: Presta servicios a Historiadores de Arte, Editores y Educadores de Estados Unidos en documentación sobre historia del arte, con énfasis en arquitectura paisajista y escultura en la parte alta del estado New York.

Después de concluido el período de nominación, la Academia recibe tres nominaciones para digitalización. El cuadro que aparece a continuación describe las colecciones y cómo se evalúan las mismas utilizando el Modelo C. La línea de Situación indica si los materiales fueron nominados (Modelo A) o si se sugiere que se descarten (Modelo B).

Descripción de las colecciones	**Situación**	**Decisiones de evaluación**
Colección 1: Fotografías históricas de la Sección de New York del American Atheneum of Art, 1870-1956, incluye imágenes tomadas por Julia Margaret Cameron, Ansel Adams, y Alfred Stieglitz que documentan la historia de la institución. No hay modelos de liberación de las mismas, y la institución no posee los derechos de autor.	Nominada (Modelo A) por un fotógrafo historiador visitante y por curadores de la institución	**Seleccionada:** Casi el 90% de esta colección cumple los requisitos de Misión y Enfoque de las Colecciones de la Academia. Los derechos de autor de la mayoría de las imágenes han caducado. Se trató de obtener permisos para los materiales que todavía s encuentran bajo protección de derecho de autor. Como los modelos de las fotos ya no viven, no se necesitan los permisos de los modelos. No so necesarios derechos de publicidad, E-FOIA, o sensibilidad, excepto en el caso de un desnudo que no se digitalizará. Las imágenes son auténticas, tienen valor asociado, evidencial y como artefacto y se pueden escanear con precisión. La colección está bien catalogada y se puede

		contextualizar perfectamente debido a la existencia de catálogos de exhibición. Se incrementará su valor al hacer que imágenes muy valiosas tengan accesibilidad internacional conjuntamente con índices de búsqueda. Se puede evitar digitalizar negativos de vidrio que se están deteriorando y platinotipos de la colección. Los originales se colocarán en lugar bien seguro y en refrigeración. Aunque nadie ha digitalizado estas imágenes hasta ahora, el Consorcio de la Academia está interesado en compartir los costos, ya que otros miembros del consorcio poseen porciones de esta colección y se pudieran combinar para crear un museo virtual en línea
Colección 2: Los papeles de Matias Martin, 1890-1979, contienen correspondencia con los más importantes poetas y artistas del mundo, entre las que se incluyen seis cartas (<2% de la colección) al Director de la Academia Quonsethut de Bellas Artes. La mayoría de las hojas están muy quebradizas, y muchas cartas están escritas con tinta ferrogálica, lo que provoca roturas y rasgados.	Los abogados y conservadores sugieren que se descarten (Modelo B)	**Eliminado (Se escogen y descartan):** La colección contiene grandes cantidades de materiales que contienen cuestiones de difamación relacionadas con individuos privados que aún viven. La academia no posee los derechos de autor, y consumiría mucho tiempo y dinero obtener los permisos. Los abogados y conservadores recomiendan que se excluyan. Deberán volverse a considerar en el futuro (alrededor del 2024, cuando caduquen los derechos de autor y los poetas ya hayan fallecido), puesto que existe gran interés por parte de los académicos. Mientras, se buscará financiamiento para estabilizar la colección.

Colección 3: Colección Sylvia Hands, 1930-1990, contiene los documentos personales y de la oficina de la primera curadora, más los registros de las primeras 50 exhibiciones presentadas en la Academia Quonsethut de Bellas Artes, incluida una significativa documentación académica y notas sobre los temas de las exhibiciones que están relacionados con la misión de la organización. El 50% de la colección tiene restricciones de derecho de autor. Alrededor del 35% de los documentos de la colección tiene moho activo, pero en la herencia se lega un dinero para tratamiento de conservación. Sólo los investigadores de la institución han utilizado la colección.	Un heredero sugiere que se excluya, ya que se piensa que no es bueno para la familia	**Seleccionada con reservas,** para digitalizarla una vez que se haya estabilizado la colección. Se ajusta con exactitud a la Misión y Enfoque para las Colecciones de la Academia. Se considera que la propuesta de exclusión de los depositarios es poco seria y que se puede negociar para decidir sobre algunas cuestiones. Según el abogado de la Academia, la mayor parte de la colección se fue recopilando según la contrataban para realizar trabajos, por lo que las cartas no están protegidas por derecho de autor. Se puede tratar de conseguir el permiso de los autores o destinatarios. No existen problemas de privacidad, publicidad, E-FOIA, o sensibilidad. Los materiales poseen valor evidencial, asociado e informativo, son auténticos y poseen exactitud visual. Están bien documentados y contextualizados y añadirá un valor significativo a los catálogos de exhibición y la documentación de la colección que ya están digitalizados. La mayor parte del público al que está dirigida está compuesta por académicos de arte. La tecnología beneficiará mucho estos artículos. Algo más de un tercio de la colección necesitará estabilización y limpieza. Todos los objetos necesitarán catalogación a nivel de objeto para que sean accesibles a la comunidad internacional del arte. En la actualidad no se puede acceder a los materiales, y digitalizar no representaría duplicar esfuerzos. Este proyecto no podría ser un proyecto cooperativo, aunque se podría vincular a otras obras relacionadas con la colección que también pertenecen a la Academia.

Priorización según el valor, el uso y el riesgo

Si los materiales a digitalizar son demasiados, priorice los objetos según su *valor, uso, y riesgo* (Vogt-O'Connor, 1995). Durante varios años muchas organizaciones han utilizado estos criterios de evaluación. El comité asigna valores a los materiales y computa los totales para ayudar en esta fase final de la selección. No obstante, será necesario determinar cada uno de estos factores en relación con la misión y declaración de colección, no de manera aislada.

Valor

Los materiales que se van a digitalizar deben tener uno de los valores siguientes o más en relación con la declaración de enfoque de colecciones aprobada de la organización.

Valor informativo se refiere al contenido en cuanto al tema en relación con la misión y la declaración de enfoque de colecciones de la organización.

- Las colecciones de alto valor ofrecen información significativa sobre personas, lugares, acontecimientos, objetos, períodos, actividades, proyectos y procesos importantes (tanto natural como cultural) que se refleja en la declaración de la colección, en el marco temático (si la organización lo posee), y en la misión.
- Las colecciones de valor moderado dicen algo de los tópicos y temas (como por ejemplo el "quién, qué, dónde, cuándo, por qué, y cómo") que se refleja en la declaración de misión y la política de colección.
- Las colecciones de poco valor ofrecen poca información acerca de factores claves que se reflejan en la declaración de misión y de colecciones.

Valor administrativo se refiere a la utilidad funcional del material para la organización creadora, siempre sobe una base de regularidad, como por ejemplo, la necesidad de disponer de planos arquitectónicos para realizar trabajos de restauración en un edificio o registros vitales para una operación.

- Las colecciones de alto valor están en uso constante por parte de la administración de la organización.
- Las colecciones de valor moderado se usan en algunas ocasiones.
- Las colecciones de poco valor se revisan sólo en raras ocasiones.

Valor como artefacto, según la definición de los archivistas, es el mismo valor intrínseco y se refiere a los materiales originales que tienen valor debido a su propia naturaleza.

- Entre los materiales de alto valor se incluyen objetos en buenas condiciones que son objetos raros o de interés para la cultura material. Por ejemplo, entre los materiales de alto valor se incluirían materiales visuales bien ejecutados, cartas holográficas con membretes poco comunes o diarios únicos, documentos en soportes o procesos históricos raros, como los platinotipos; materiales confeccionados en géneros y formatos poco comunes, como las fotos síquicas o daguerrotipos de media plancha.
- Entre los materiales de valor moderado se incluyen los procesos de uso extendido, como los impresos fotográficos en albúmina, libros encuadernados o cartas mecanografiadas y algunos formatos como las estereografías que se encuentren en buenas condiciones.
- Los materiales de poco valor son objetos en malas condiciones o copias o duplicados.

Valor asociativo se refiere a materiales originales que tienen relación con un individuo eminente, algún lugar, acontecimiento o grupo, como las cartas escritas, firmadas o que hayan sido propiedad de Tomás Alba Edison o fotos de soldados de la Guerra Civil de Estados Unidos o tomadas por los soldados participantes en la misma.

- Los materiales de alto valor incluyen objetos como los documentos personales de un individuo o grupo notable, o los qué estén asociados a algún proyecto como por ejemplo una excavación arqueológica.

- Las colecciones de valor moderado podrían contener alguna correspondencia o retratos de individuos destacados.

- Los materiales de poco valor incluyen las copias o duplicados.

Valor evidencial se refiere a la habilidad de los documentos para servir como prueba legal o histórica de una actividad, acontecimiento u ocupación.

- Los materiales de gran valor son los originales sin modificar.

- Las colecciones de valor moderado podrían incluir algunos registros de valor legal, como certificados de nacimiento o copias legales de registros de propiedades de tierras.

- Los materiales de poco valor son los registros o las copias modificados.

Valor monetario se refiere al valor de mercado actual de un objeto. Este valor puede cambiar diariamente.

Cómo medir el valor

Cuando se le otorga prioridad a diferentes grupos de materiales, por ejemplo varias colecciones de manuscritos o series dentro de una misma colección de archivos, o donaciones importantes:

- Marque 6 puntos si un grupo de objetos tiene alto valor en cualquiera de las categorías mencionadas anteriormente para una parte significativa de la colección, *es decir,* el 10% o más (alto valor).

- Marque 3 puntos si un grupo de materiales tiene menos del 10% o no posee materiales de alto valor, pero tiene valor moderado en cualquiera de las categorías anteriores (valor moderado).

- Marque 1 punto si la colección no tiene materiales de alto valor o de valor moderado (poco valor).

Riesgo

El riesgo puede asumir diversas formas: legal, social, y de preservación. Como los documentos que pueden presentar riesgo social o legal serán excluidos durante el proceso de selección, esta priorización se concentrará en los riesgos de preservación.

- *Alto riesgo.* Los materiales de más alto riesgo son fundamentalmente los que son químicamente inestables, lo que da como resultado que se auto destruyan y que dañen o contaminen los materiales que se encuentran junto o cerca de ellos, a la vez que representan un peligro para la salud del personal y los investigadores que los utilizan. Ejemplos clásicos de materiales de alto riesgo son los negativos de nitrato de celulosa y películas y otros materiales que presentan contaminación biológica o química, como hongos, insectos y

vectores que presentan riesgos de pérdida de la información así como peligro para la salud. Los ejemplos clásicos de materiales que presentan riesgo de seguridad y para la salud son aquellos contaminados con asbesto, hongos Aspergillum, y Hantavirus. Otros materiales de alto riesgo pueden ser materiales que se autodestruyen a causa de sus problemas inherentes (por ejemplo, la tinta ferrogálica, las encuadernaciones en piel afectadas por carcoma roja, papel muy quebradizo y ácido, y película de acetato de celulosa) y aquellos objetos que pudieran estar provocando daños en los materiales cercanos (como los materiales con cintas adhesivas).

- *Riesgo moderado.* Los materiales que están expuestos a riesgo moderado son aquellos que están experimentando daños mecánicos o físicos primarios debido a la forma en que están almacenados o a la manipulación y a las características del material (*ej.* la fuerza que se ejerce en el doblez). Los objetos que se están deteriorando y que están perdiendo el contenido informativo de manera natural o gradual debido a los procesos que experimentan los componentes y materiales con que están confeccionados son objetos de riesgo moderado. Algunos ejemplos incluyen los medios electrónicos y digitales de transporte de datos como los CD-ROM y los disquetes; la mayoría de las diapositivas en colores, los negativos, y las fotografías impresas y los materiales en soporte de celulosa-éster (acetato, diacetato, y triacetato); todas las imágenes que se descaman, las retocadas, friables o coloreadas a mano; los libros de impresión tipográfica, especialmente los confeccionados con tintas de grafito, las copias de papel carbón de la correspondencia y algunos papeles de dibujo. Otros factores que se pueden considerar de igual riesgo son los materiales de pequeño formato como las microformas, y se les debe otorgar la máxima prioridad cuando hay alguna pérdida de información. También se incluyen en la categoría de riesgo moderado, pero con menos prioridad, los artículos que tienen agujeros, rajaduras, piezas rotas o desgarradas, desgarraduras, rasgados, ponchaduras, pérdidas o aquellas que están torcidas, dobladas, estrujadas, arrugadas, plegadas, pandeadas, arañadas, dañadas por abrasión, manchadas, descoloridas, o dañadas estructuralmente de cualquier otra manera o cuya apariencia ha cambiado (*ej.* cambio del balance del color).

- *Bajo riesgo.* Los materiales de bajo riesgo tienden a ser aquellos materiales cuyos procesos son más duraderos, que no están dañados, y están en adecuadas condiciones de almacenamiento. Entre ellos podemos mencionar los que presentan pequeñas manchas en forma de moteado; documentos empolvados o sucios, y las copias de planos ligeramente descoloridas y los cianotipos que están bien almacenados.

Cómo clasificar el riesgo

Cuando se comparan grupos de materiales (por ejemplo, colecciones de manuscritos):

- Marque seis puntos si el 10% o más de la colección está en condiciones de alto riesgo según los criterios explicados anteriormente. Considere toda la colección de alto riesgo.
- Marque tres puntos si menos del 10% de la colección es de alto riesgo. Considere toda la colección en riesgo moderado.
- Marque tres puntos si no hay materiales en alto riesgo y el 10% o más de la colección están en riesgo moderado. Considere toda la colección en riesgo moderado.

- Marque un punto si menos del 1% de los materiales de la colección están en alto riesgo y menos del 10% son materiales en riesgo moderado. Considere toda la colección en bajo riesgo.

Las colecciones que también son de alto valor ameritan la digitalización cuando se puede minimizar o eliminar el riesgo. La institución puede decidir no digitalizar colecciones de alto riesgo que son de poco valor.

Uso

El tercer factor a tener en cuenta para determinar la prioridad de la colección para digitalización es el uso. Los materiales de uso elevado son aquellos que el personal y/o investigadores de fuera solicitan más frecuentemente para referencia. Si el proyecto de digitalización está dirigido a un público nuevo, no serán de mucha ayuda las estadísticas de uso anterior. Será necesario, por tanto, basar las decisiones en los pronósticos de uso esperado. Comuníquese con los depositarios de experiencia en el trabajo con públicos nuevos, así como con personas de ese posible público para poder pronosticar el uso futuro. Considere como un pequeño proyecto piloto comprobar la respuesta del público antes de comprometerse con una nueva iniciativa de esa envergadura.

Por lo general, los materiales que se usan mucho son materiales de alto valor. En algunas ocasiones, ciertos materiales que aparentemente no tienen mucho valor se hacen populares de pronto debido a un especial encanto de expresión -- por ejemplo, un giro del lenguaje en una carta, un ángulo singular en una fotografía, o un vínculo con un acontecimiento o actividad que antes no se celebraba. Según cambia el mundo académico, cambian también los valores que se le concedían a los materiales. Cuando se puede pronosticar un alto uso y minimizar el riesgo, digitalizar es una sabia solución para el acceso.

Cómo categorizar el uso

Cada institución debe establecer sus propios valores para este campo basándose en estadísticas como referencia y en los registros de visitantes. Para poder hacer esto, deben conocerse las estadísticas de uso de la institución. Después, halle la media de los valores de uso para una colección. Por ejemplo, si la media de los usos en el año por colección fue de 10, entonces un uso bajo para una colección sería entre 1-6, el uso moderado estaría entre 7-13, y un uso elevado sería de 14-20+.

Se unen todos los elementos: Cómo evaluar y dar un rango de prioridad a cada colección

---

El valor, el riesgo y el uso, cuando se analizan juntos y se les asignan valores (basándose en los valores numéricos de alto=6, moderado=3, y bajo=1), indican cuáles son las colecciones que requieren digitalización. La clave es que cada uno de estos factores debe determinarse

en relación con la misión y declaración de colección de la organización, y no de manera aislada.

Después de asignar un valor numérico al valor, el riesgo y el uso, el comité prioriza las colecciones según esos valores numéricos. En el caso de que haya coincidencia en los valores numéricos, compare las cifras de uso para determinar cuál es en realidad más alta. Compare también los tipos de deterioro para ver cuál representa una amenaza mayor para la vida de la colección.

En el ejemplo que aparece a continuación, se presentan dos colecciones con una puntuación de 15. La Colección 5 tiene un problema a mayor escala con el nitrato que el la Colección 1 tiene con el moho; además el uso de la Colección 5 fue mayor, lo que la hace que sea a todas luces la ganadora en términos de prioridades.

Ejemplo de cómo categorizar basándose en el valor, el uso y el riesgo

Nombre: Academia Quonsethut de Bellas Artes en New York

Misión y Propósito de la colección: Presta servicios a Historiadores de Arte, Editores, y Educadores en Estados Unidos en documentación de historia del arte, con énfasis en la arquitectura paisajista y la escultura en la parte alta del estado de New York.

Descripción de las colecciones	Valor	Riesgo	Uso	Puntos	Prioridad
Colección 1: Fotografías históricas del American Atheneum of Art, 1870-1956, incluye imágenes tomadas por Julia Margaret Cameron, Ansel Adams, y Alfred Stieglitz que documentan la historia de la institución; algunas presentan hongo (12%), y usos moderados (o 70) anualmente.	Alto (6)	Alto (6)	Moderado (3)	15	3 (tiene menos uso y riesgos que la Colección #5)
Colección 2: Papeles de Matin Femwit, 1900-1989, contiene correspondencia con importantes científicos de todo el mundo en los campos de la entomología, la física y la mamología, y dos cartas (<1% de la colección) al Director de la Academia Quonsethut de Bellas	Bajo (1)- No se adapta a la misión o al enfoque	Bajo (1)	Bajo (1)	3	5

Artes. En buenas condiciones con poco moteado, relativamente poco uso y no está catalogada aún. **Nota:** Se le podría otorgar mayor prioridad si estuviera vinculada a alguna necesidad del consorcio o tuviera alguna prioridad o por una fuente de financiamiento externa que incluya financiamiento para hacer el trabajo.					
Colección 3: Colección Sylvia Hands, 1930-1990, contiene los documentos personales y de la oficina de la primera curadora, más los registros de las primeras 50 exhibiciones presentadas en la Academia Quonsethut de Bellas Artes, incluida una significativa documentación académica y notas sobre los temas de las exhibiciones que están relacionados con la misión de la organización. El 50% de la colección tiene restricciones de derecho de autor. Alrededor del 35% de los documentos de la colección tiene moho activo, pero en la herencia se lega un dinero para tratamiento de conservación. Sólo los investigadores de la institución han utilizado la colección.	Alto (6)	Alto (6)	Bajo (1)	7	4

Colección 4: Colección de Historia Oral de la Academia Quonsethut de Bellas Artes, 1950-1999, incluye historias orales y en video, transcripciones y modelos de autorización de documentación sobre importantes artistas, especialmente escultores y arquitectos paisajistas del Nordeste de Estados Unidos. No se ha hecho migración alguna de las cintas ni se han refrescado las mismas y muchas todavía se encuentran en soportes de acetato o el adhesivo se está descamando. Tiene un uso elevado, por lo general más de 300 usuarios por año. Muchos solicitan copias en formato digital.	Alto (6)	Alto (6)	Alto (6)	18	1
Colección 5: Escultor Tom McMakeitup, 1935-1999, incluye los papeles personales del escultor y arquitecto paisajista neoyorquino Tom McMakeitup, afamado relativista. La colección contiene sus documentos personales, de la familia y de negocios, incluida la correspondencia, películas y videotapes, y documentación fotográfica de su trabajo para la Familia RockePont-Mellon en la propiedad de Kitchie en los Montes Adirondacks. La colección contiene alrededor de 6,000 negativos de nitrato; 4,000 pies de película de nitrato; y unas 5,000 diapositivas coloreadas en película de celulosa-éster. El uso es moderado entre los cineastas,	Alto (6)	Alto (6)	Moderado (3)	15	2 (Tiene más uso y riesgos que la Colección #1)

historiadores del arte y
arquitectos paisajistas,
con alrededor de 96 usos
al año.

Resumen de los aspectos claves

Las tres etapas de la selección son:

- ***Nominación de materiales*** para selección y descarte o exclusión (depositarios, público, personal y académicos);

- ***Evaluación de materiales*** y exclusión de materiales que no son adecuados para la digitalización (Comité de Selección);

- ***Priorización de los materiales restantes*** basándose en los criterios de valor, uso y riesgo para garantizar que se digitalicen primero los materiales más importantes (Comité de Selección).

Seguir los pasos de este proceso asegura que el proyecto de digitalización responda a las necesidades de los individuos que más lo necesitan y que tendrán mayor responsabilidad en el éxito del mismo. Pensar cuidadosamente en las principales cuestiones según se revisan los materiales, el comité evitará que se detenga el proceso en el medio del flujo. Por último una cuidadosa y bien pensada selección garantiza que la organización emplee el dinero del financiamiento en los objetos más importantes, más útiles y más en riesgo y que se adopten decisiones sabias.

Modelo A, Modelo de Nominación para Selección

Modelo de Nominación de Proyecto de Digitalización de la Institución X

1. Materiales que se nominan para digitalización (Por favor, indique el número de colección, series, número, número de caja, número de carpeta, número de control del objeto o equivalente y creador; título del objeto o citación bibliográfica con el máximo de detalle posible.)

2. Razones para la nominación (Describa por qué son importantes los materiales, quiénes podrían querer usarlos en formato digital, y los usos probables si se digitalizan.)

3. Fuentes Potenciales de Asistencia (Por favor, indique si tiene algún conocimiento o habilidad especial que pueda compartir con la institución X durante el proceso de selección. Por ejemplo, ¿puede Ud. proporcionar información, antecedentes históricos, o tiene Ud. conocimiento de posibles fuentes de financiamiento o de proyectos de digitalización que incluyan materiales similares a los que Ud. está nominando?)

4. Restricciones (Indique si Ud. conoce alguna razón por la que los materiales especificados no deben digitalizarse, como por ejemplo, cuestiones legales, éticas o culturales. Por favor, sea lo más específico que pueda al citar una fuente, como una ley o un grupo cultural y el nombre de alguna persona a la que se pueda contactar si fuera necesario.)

5. Su nombre: _____

6. Su dirección: _____

_____

7. Tel:_____ Fax:_____

8. E-Mail: _____

Nota:
El Comité de Selección tomará las decisiones finales con respecto a lo que se incluirá y lo que no se incluirá en el proyecto de digitalización. Si Ud. posee alguna información especial que Ud. desee compartir con el comité, por favor escríbala a continuación.

Modelo B, Modelo de Nominación para exclusión

Modelo de exclusión del Proyecto de Digitalización de la Institución X

1. Identifique los materiales que no se deben digitalizar (Por favor, indique el número de colección, series, número, número de caja, número de carpeta, número de control del objeto o equivalente y creador o título del objeto con el máximo de detalle posible.)

2.Razones de la exclusión (Describa por qué los materiales no deben digitalizarse o compartirse en formato electrónico. Identifique problemas o intereses que podrían surgir, incluidos los aspectos culturales, sociales o éticos. Identifique quién podría afectarse si los materiales se hacen accesibles en formato electrónico.)

3. Restricciones específicas (Indique si Ud. tiene conocimiento de alguna razón por la que no deberían digitalizarse los materiales, y cite leyes específicas, políticas, o documentación equivalente. Por favor, sea lo más específico posible citando alguna fuente, como por ejemplo, alguna ley o grupo cultural y persona a la que se puede contactar si fuera necesario.)

_____

_____

_____

_____

_____

_____

_____

_____

5. Su nombre: _____

6. Su dirección: _____

_____

7. Tel:_____ Fax:_____

8. E-Mail: _____

Nota: El Comité de Selección tomará todas las decisiones finales con respecto a qué se incluye y qué no se incluye e
proyecto de digitalización. Si Ud. posee alguna información específica que quisiera compartir con el comité, por fav
escríbala a continuación.

Modelo C, Lista Guía para la Evaluación

Responda Sí o No a cada pregunta.

Factores de evaluación:	Sí Digitalice	No, No Digitalice

***Declaración de Misión*:** ¿El proyecto se incluye dentro de la declaración demisión del depósito o de la institución? Si la respuesta es no, no digitalice.

Declaración de enfoque de las Colecciones: ¿Los materiales candidatos están dentro de la Declaración de Enfoque de Colecciones? (Política de desarrollo de Colecciones). Si la respuesta es **No**, no digitalice a menos que el depósito redefina la política de inclusión de materiales.

Solicitudes de Exclusión de los Depositarios: ¿Ha recibido el depósito solicitudes para excluir materiales de alguno de los depositarios o de alguna fuente confiable? De ser así, ¿han sido esas solicitudes impugnadas por solicitudes equivalentes para no digitalizar los materiales? De ser así, no digitalice los materiales. Nota: Si Ud. tiene solicitudes de exclusión que el Comité de Selección considera poco serias o sustanciales, ignórelas.

Restricciones del Donante: ¿El material donado no está restringido? De ser así digitalice. ¿Ha impuesto el donante o el creador de los materiales restricciones sustanciales o no negociables al uso de los mismos que pudieran evitar que se digitalizaran? De ser así, no digitalice los materiales. Nota: En ocasiones, se pueden negociar las restricciones impuestas por el donante.

Derechos de autor: ¿El material es del dominio público o está protegido por derechos de autor que su organización haya obtenido? De ser así, digitalice. En caso contrario, ¿tiene Ud. alguna razón para creer que Ud. no tendrá posibilidades de obtener una licencia para utilizar los materiales? De ser así, no digitalice hasta que haya obtenido los derechos de autor o licencias/permisos.

Derechos de privacidad: ¿El material contiene imágenes de individuos que están vivos y de las cuales Ud. posee los modelos de permisos (particularmente para historias orales y videos, registros personales, historias clínicas de siquiatría, o fotografías en los que se puede reconocer a los individuos)? Si la respuesta es Sí, digitalice. Si la respuesta es No, ¿Ud. tiene razones para creer que Ud. no puede localizar a esos individuos o que los mismos no le darán los permisos? De ser así, no digitalice hasta que haya obtenido los permisos.

Derechos de publicidad: ¿Existe en su estado una ley de publicidad (*ej.* California, Tennessee, New York)? De ser así, ¿incluye su material imágenes o grabaciones de individuos famosos, como por ejemplo estrellas de cine o músicos, científicos, artistas o autores (vivos o muertos)? De ser así, ¿tiene Ud. permisos o licencias que le hayan dado esos individuos o sus herederos para utilizar las imágenes? De ser así, digitalice. De lo contrario, no digitalice hasta que obtenga los permisos o licencias.

E-FOIA y Equivalentes Estatales o Locales: ¿Le exige la ley que digitalice los materiales candidatos de forma tal que cumplan la Ley de Libertad de Información, la Ley para la Reducción de la Burocracia, o iniciativas similares? De ser así, digitalice mientras no exista una razón poderosa para

no digitalizar los objetos que no pueda resolver en ese momento. Si existiera una razón poderosa, no digitalice los materiales.

Sensibilidad: ¿El material candidato contiene alguna información sensible (por ejemplo la ubicación de lugares sagrados, de enterramientos, donde viven especies en peligro de extinción, fósiles, recursos culturales en peligro [como petroglifos], o lugares de subsistencia o recolección de alimentos), o los materiales nominados presentan un punto de vista desbalanceado o carecen de perspectivas de contrapunteo? De ser así, ¿están programados los proyectos de manera adecuada y con el personal requerido para tratar de realizar consultas y realizar actividades para recopilar los permisos con los grupos afectados y para consultar con académicos de varios puntos de vista? Si no es así, no digitalice los materiales o digitalice solamente los materiales que el comité está plenamente equipado para evaluar y poner dentro de un contexto. Involucre a los depositarios en el Comité de Selección o en el personal del proyecto.

Valor evidencial: ¿Es el valor fundamental de los materiales de carácter evidencial, o se considera prueba legal o histórica de una acción o acontecimiento? ¿El material tiene también contenido sustancial informativo y/o asociativo que pueda ser de interés a un público clave? De ser así, digitalice. De no ser así, ¿el hecho de trasladar el objeto del ámbito analógico al ámbito digital erosionará su valor de tal manera que ya no servirá a los propósitos fundamentales del mismo? En ese caso, o en el caso en el que el valor se ve seriamente dañado o si no existe un público, no digitalice.

Autenticidad: ¿Es auténtico el artículo que se va a digitalizar y no se ha falsificado, copiado o alterado de manera sustancial? De ser así, digitalice. Si no es así, ¿digitalizar el objeto aportará una falsa autenticidad a un documento u objeto auténtico? ¿Es imposible rectificar las concepciones erróneas a través de documentación contextual, subtitulajes o metadatos realizados con todo cuidado? Si no es así, no digitalice. Nota: Si el proyecto implica alteraciones sustanciales o retocar una obra de arte visual para propósitos que no sean una parodia o una sátira en potencial violación de la Ley de Derechos de los Artistas Plásticos, no digitalice loe materiales.

Precisión visual: ¿La técnica de escaneo propuesta será capaz de capturar la apariencia del objeto con precisión? De ser así, digitalice. Si no, ¿puede avanzar el proyecto a una técnica de escaneo más sofisticada, como por ejemplo escaneo a color para capturar la información? Si no puede capturar la imagen con precisión, no digitalice el material.

Documentación: ¿Los materiales candidatos tienen buenos encabezamientos? En ese caso, digitalice. Si los subtítulos son incorrectos, malos o incompletos, o si las descripciones y rotulados tampoco son buenos, ¿el personal que trabajará en el proyecto y el presupuesto del mismo pueden proporcionar una buena documentación dentro del cronograma propuesto para el proyecto? Si no es así, no digitalice los materiales.

Contextualización: ¿Los materiales candidatos requieren una investigación sustancial y un contexto sofisticado y costoso para que sean de utilidad? Si este es el caso, ¿puede el proyecto proporcionar el contexto? De ser así, digitalice. En caso contrario, ¿la posibilidad de visualizar los materiales de forma seriada, pero no uno al lado del otro, disminuirá el valor de los

ficheros para el público de manera significativa? De ser así, ¿puede el proyecto proporcionar una vía para visualizar los materiales uno junto al otro? De lo contrario, ¿hay otros objetos dentro de la colección que se puedan seleccionar en lugar de ese tópico? Si se van a utilizar los ficheros, ¿deberá reconstruirse toda una excavación arqueológica o deberá colocarse un medio de búsqueda archivística en formato de Descripción Archivística Codificada [Encoded Archival Description (EAD)] o en algún equivalente? De ser así, ¿el personal y el presupuesto del proyecto son adecuados para producir este tratamiento contextualizado? Si Ud. no puede proporcionar el contexto necesario y el Comité de Selección considera que el contexto es esencial, no digitalice.

Valor agregado: ¿Los materiales candidatos son valiosos y a la vez se van a hacer accesibles por primera vez? De ser así, digitalice. ¿El proyecto añade valor a los materiales candidatos? Si es así, digitalice. ¿El proyecto simplemente repite algún trabajo que ya existe en una publicación en formato analógico o en papel (como algo copiado con carbón)? Si ese es el caso, ¿puede reelaborarse el proyecto de manera que añada valor a los materiales con el mejoramiento del acceso mediante la creación de:

- un público nuevo para materiales raros o únicos a los que en la actualidad sólo pocas personas tienen acceso?
- vínculos con materiales separados via códigos HTML, SGML, o XML?
- colecciones virtuales de materiales por el mismo creador, confeccionados con el mismo tipo de proceso, soporte, técnica o formato; u otro vínculo que de lo contrario esté físicamente separado o en un sitio Web o CD-ROM aislado?
- nuevos índices y medios de búsqueda que sean electrónicamente buscables?
- nueva buscabilidad mediante procesamiento posterior al escaneo vía OCR o reelaborándolos para que los ficheros textuales sean totalmente buscables?
- nuevas vías para analizar los originales mediante técnicas, como el escáner microscópico, escáners en 3-D u otras técnicas similares?
- ficheros usables para investigación cuando los originales están demasiado manchados, deteriorados o dañados por retoques anteriores u otros tratamientos para poder utilizarlos en investigación?

Si no, ¿el conocimiento experto del personal y el presupuesto del proyecto son adecuados para producir este nuevo tratamiento al material? De no ser así, no digitalice el material hasta que la digitalización le proporcione algún valor agregado.

Público: ¿El público nuevo que se proyecta para las imágenes digitales es el mismo que el de los originales? Si ese es el caso, ¿tendría en cuenta la institución la posibilidad de recontextualizar el producto digital para que llegue a un público más amplio? De ser así, digitalice. ¿Ayudará el proyecto

digital a que se llegue al mismo público de manera más efectiva? De no ser así, no digitalice los materiales.

Criterios suplementarios de selección: ¿El público ha establecido criterios suplementarios de evaluación que deben tenerse en cuenta dentro del proceso de evaluación, como por ejemplo el Criterio de Utilidad del Maestro desarrollado para la Biblioteca del Congreso? (EDC) ¿Se corresponde esta selección con estos criterios adicionales? De ser así, digitalice, de lo contrario no digitalice.

Tecnología: ¿El público proyectado exige técnicas de escaneo complejas o sofisticadas y equipos de visualización para utilizar las imágenes según se ha pensado? De ser así, ¿es posible que un porcentaje suficiente del público tenga este nivel de tecnología de visualización? De no ser así, planifique de nuevo el proyecto. Si ese fuera el caso, ¿los materiales textuales digitalizados necesitarán procesamiento OCR o armonización? ¿Las imágenes requieren retoques, copiado de alta resolución, captura del color, o codificación extensiva para mantener los vínculos y las jerarquías? De ser así, ¿se puede realizar el trabajo dentro del presupuesto y el cronograma del proyecto utilizando el personal del proyecto? Si no es así, no digitalice.

Estado de conservación: ¿Los objetos candidatos se encuentran en condiciones estables o se dispone de duplicados o copias para utilizar en la digitalización? De ser así, digitalice. De lo contrario, ¿los objetos candidatos están tan deteriorados o tan en riesgo que sería difícil digitalizarlos o copiarlos o se podrían dañar los originales? Por ejemplo, ¿hay alguna necesidad de desmontar la encuadernación de un libro de recortes único o un libro raro, desmontar objetos de los marcos o bastidores, o hacer presión sobre una imagen pandeada o muy quebradiza? De ser así, ¿es demasiado costoso estabilizar los originales o llevaría demasiado tiempo para hacerlo dentro del presupuesto y el cronograma del proyecto? Si es así, no digitalice.

Control: ¿Se accede a los materiales originales, se describen y se colocan en un lugar de almacenamiento seguro? De ser así, digitalice. De lo contrario, ¿la digitalización de los originales los pondría en riesgo al alertar a posibles ladrones de la existencia de originales valiosos y vulnerables? De ser así, no digitalice.

Duplicación de esfuerzos: ¿Ha comprobado que estos objetos no han sido digitalizados en alguna otra institución? Si no han sido bien duplicados en alguna otra institución, digitalice. Si ya han sido digitalizados, ¿la copia digital tiene la calidad adecuada? De ser así, obtenga una copia de la otra fuente y no digitalice los materiales.

Accesibilidad: ¿Los materiales candidatos no están accesibles, por ejemplo se encuentran en almacenamiento refrigerado? De ser así, digitalice. Si se tiene fácil acceso a los mismos en múltiples lugares — por ejemplo a través de copias en microfilme ampliamente distribuidas o en muchos catálogos de exhibición publicados — ¿existe alguna razón especial por la que sean necesarias las copias digitales? Si no la hay, no digitalice.

Proyecto: ¿Se le ha otorgado a los materiales candidatos cierta prioridad debido a alguna prioridad temática, cooperativa o de financiamiento? De ser así, ¿estas prioridades se ajustan a las declaraciones de misión y colección de

la institución? Si es así, digitalice. De lo contrario, no digitalice.

Acumulación: ¿El material candidato es parte de un grupo o de un cuerpo de materiales vinculados que atraen valor adicional al estar relacionados con otros materiales en la colección de la institución? ¿Ya están bien digitalizados, o seleccionados para digitalización o relacionados con otros materiales que ya han sido bien digitalizados por otras organizaciones? Si ese es el caso, digitalice. Si se trata de un solo objeto y el esfuerzo no es sólo para relaciones públicas o en respuesta a E-FOIA, o ha sido solicitado por uno de los depositarios, no digitalice.

Bibliografía

Atkinson, Ross. "Selection for Preservation: A Materialistic Approach." *Library Resources and Technical Services* 30 (octubre/diciembre 1986): 344-53.

Bagnell, Roger S. y Carolyn L. Harris. "Involving Scholars in Preservation Decisions: The Case of Classicists." *The Journal of Academic Librarianship* 13:3 (1987).

Besser, Howard y Jennifer Trant. *Introduction to Imaging.* Santa Monica, CA: Programa de Información deHistoria del Instituto de Arte Getty, 1995.

Child, Margaret S. "Further Thoughts on Selection for Preservation: A Materialistic Approach." *Library Resources and Technical Services* 30 (octubre/diciembre 1996): 354-62.

----. "Selection for Preservation" in *Advances in Preservation and Access.* Editado por Barbara Buckner Higginbotham y Mary E. Jackson. Westport, CT: Meckler Publishing, 1992.

Consejo para Bibliotecas y Recursos Informáticos. *Scholarship, Instruction, and Libraries at the Turn of the Century: Results from Five Task Forces Appointed by the American Council of Learned Societies and the Council on Library and Information Resources.* Washington, DC: Consejo para Bibliotecas y Recursos Informáticos, 1999.

DeStefano, Paula. "User-Based Selection for Preservation Microfilming." *College and Research Libraries*, septiembre 1995.

Ellis, Judith, ed. *Keeping Archives,* 2ª. edición. Melbourne, Australia: Thorpe and the Australian Society of Archivists, 1993.

Ester, Michael. *Digital Image Collections: Issues and Practice.* Washington, DC: Comisión de preservación y acceso, 1996.

----. "Digital Images in the Context of Visual Collections and Scholarship." *Visual Resources* 10 (1994): 11-24.

----. "Image Quality and Viewer Perception." *Leonardo:Journal of the International Society for the Arts, Sciences, and Technology,* Supplemental Issues (1990): 51-63. Reimpreso en *Visual Resources* 7:4 (1991): 327-52.

----. "Specifics of Imaging Practice." *Archives and Museum Informatics, Hands on Hypermedia and Interactivity in Museums.* Trabajos Seleccionados de la Tercera Conferencia Internacional (ICHIM/MCN95), San Diego, CA (Octubre 1995): 147-58.

George, Gerald. *Difficult Choices: How Can Scholars Help Save Endangered Research Resources? A Report to the Commission on Preservation and Access.* Washington, DC: Comisión de preservación y acceso, 1995.

Gertz, Janet. *Oversize Color Images Project, 1994-1995. Final Report on Phase I.* Washington, DC: Comisión de preservación y acceso, 1995.

Gonzales, Pedro. *Computerization of the Archivo General de Indias: Strategies and Results.* Washington, DC: Consejo para Bibliotecas y Recursos Informáticos, 1998.

Hazen, Dan, Jeffrey Horrell, and Jan Merrill-Oldham. *Selecting Research Collections for Digitization.* Washington, DC: Consejo para Bibliotecas y Recursos Informáticos, 1998.

Kenney, Anne R. y Stephen Chapman. *Digital Imaging for Libraries and Archives.* Ithaca, NY: Departamento de Preservación y Conservación, Cornell University Library, 1996.

----. *Digital Resolution Requirements for Replacing Text-Based Materials: Methods for Benchmarking Image Quality.* Washington, DC: Comisión de preservación y acceso, 1995.

Lesk, Michael. *Preservation of New Technology: A Report of the Technology Assessment Advisory Committee to the Commission on Preservation and Access.* Washington, DC: Comisión de preservación y acceso, 1996.

Library of Congress National Digital Library Collection Evaluation Criteria prepared by EDC Center for Children and Technology, 96 Morton Street, 7th Floor, New York, NY 10014 (sin fecha, 58 páginas, más bibliografía y modelos).

Lowry, Charles and Denise Troll. "The Virtual Library Project," en *Serials Librarian. ASIG Proceedings: Tradition, Technology, and Transformation Part I.* 28:1/2 (1996): 143-70.

McClung,Patricia A., ed. *RLG Digital Image Access Project.* Palo Alto, CA: Actas de un Simposio del RLG celebrado del 31 de marzo al 1° de abril de 1995.

Marcum, Deanna. *CLIR Issues 2* (marzo/abril 1998): 1-3.

National Research Council Committee of Preservation of Historical Records. *Preservation of Historical Records.* Washington, DC: National Academy Press, 1986.

Ostrow, Stephen E. *Digitizing Historical Pictorial Collections for the Internet.* Washington, DC: Consejo para Bibliotecas y Recursos Informáticos, 1998.

Puglia, Steven. "Cost Benefit Analysis for Reformatting Options," Discurso *Afterimages: Reformatting Visual Materials in a Digital World.* Ofrecido por el NEDCC y el Servicio Nacional de Parques el 17 de septiembre de 1998, en la Administración Nacional de Archivos y Registros Históricos, College Park, MD.

----. "Creating Permanent and Durable Information: Physical Media and Storage Requirements." *CRM* 22:2 (Febrero 1999): 25-27. En Internet en el sitio http://www.cr.nps.gov/crm.

Smith, Abby. *Future of the Past: Preservation in American Research Libraries.* Washington, DC: Consejo para Bibliotecas y Recursos Informáticos, 1999.

----. *Why Digitize?* Washington, DC: Consejo para Bibliotecas y Recursos Informáticos, 1999.

Smith, Levine, Melissa. "Electronic Publishing: A Legal and Practical Primer." *CRM* 18:9 (1995).

Vogt-O'Conner, Diane L. "Chapter 2: Legal Issues." *Museum Handbook, Part III: Access and Use.* Washington, DC: Servicio Nacional de Parques y GPO, 1998.

----. "Is the Record of the 20th Century at Risk?" *CRM* 22:2 (Febrero 1999): 21-25. En Internet en el sitio http://www.cr.nps.gov/crm.

----. "Preservation Reformatting: Selecting a Copy Technology." *Conserve O Gram* 19/11. Washington, DC: Servicio Nacional de Parques y GPO, 1995.

----. "Reformatting for Preservation and Access: Prioritizing Materials for Duplication." *Conserve O Gram* 19/10. Washington, DC: Servicio Nacional de Parques y GPO, 1995.

Vogt-O'Conner, Diane L., Virginia Kilby, and Joan Bacharach. "Chapter 1: Evaluating Museum Collections for Use." *Museum Handbook, Part III: Access and Use.* Washington, DC: Servicio Nacional de Parques y GPO, 1998.

----. "Chapter 3: Publications" *Museum Handbook, Part III: Access and Use.* Washington, DC: Servicio Nacional de Parques y GPO, 1998.

Waters, Donald J. "Archiving Digital Information" in *Tomorrow's Access--Today's Decisions: Ensuring Access to Today's Electronic Resources Tomorrow.* Actas de la 14ª. Conferencia Anual de Directores de Bibliotecas de Investigación. OCLC y elComité Asesor de la Bibliotecas de Investigación de OCLC, Dublin, OH: 1996.

----. *From Microfilm to Digital Imagery: On the Feasibility of a Project to Study the Means, Costs, and Benefits of Converting Large Quantities of Preserved Library Materials from Microfilm to Digital Images.* Washington, DC: Comisión de preservación y acceso, 1996.

Waters, Donald J. y J. Garrett. *Preserving Digital Information: Final Report and Recommendations of the Task Force on Archiving Digital Information.* Washington, DC: Comisión de preservación y acceso, 1996.

Weber, Harmut y Marianne Dorr. *Digitization as a Method of Preservation?* Washington, DC: Comisión de preservación y acceso, 1997.

Buscar Contenido Sitio del NEDCC

Northeast Document Conservation Center

100 Brickstone Square
Andover, MA 01810-1494
Telephone: (978) 470-1010
Fax: (978) 475-6021

última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

Manual de Proyectos de Digitalización: Una herramienta de gestión para preservación y acceso

V

Visión general de los aspectos legales en la digitalización

Melissa Smith Levine
Asesora legal, Proyecto Nacional de Biblioteca Digital, Biblioteca del Congreso

Introducción

Este capítulo ofrece un panorama general de los conceptos básicos del derecho de autor y menciona algunas de las cuestiones legales a tener en cuenta en el transcurso de un proyecto de digitalización. Sirve de breve introducción a una serie de aspectos que deberán tenerse en cuenta en cualquier proyecto de escaneo o en línea, incluidas preocupaciones que han surgido a la par que el Programa Nacional para la Biblioteca Digital ha digitalizado y montado algunas colecciones seleccionadas de la Biblioteca del Congreso para su presentación en Internet a través de la Memoria Americana. No puede ser muy abarcador. Las cuestiones son complejas. No existe una receta simple para identificar y solucionar las cuestiones relacionadas con una colección en particular. Por lo tanto, este análisis no sustituye el asesoramiento legal competente.

Al identificar estas cuestiones legales, este capítulo presta especial atención a las razones de posibles preocupaciones legales y a los conceptos que están en las leyes para ayudarles a reconocer esos posibles problemas y a actuar de manera responsable en un terreno que todavía está en desarrollo y con frecuencia es oscuro. Tanto las organizaciones culturales como las educacionales son a la vez creadores y usuarios de productos de creación protegidos. Con frecuencia estas organizaciones

dependen de artistas, escritores y otros, o están muy estrechamente relacionadas con ellos, quienes a su vez dependen de sus productos de creación para su sustento. Por tanto es vital que tomemos la administración de la propiedad intelectual intrínseca de las colecciones con la misma seriedad con que asumimos el cuidado de las colecciones físicas. Si se comprenden bien los conceptos legales, seremos más capaces de aplicarlos a los proyectos de digitalización. Es importante comprender las razones y políticas detrás de las leyes de derecho de autor para poder tomar las decisiones necesarias con respecto a cómo hacer que las colecciones estén accesibles en Internet en casos en que las leyes evolucionan con gran rapidez. Muchos de esos conceptos se aplicarán también a otros tipos de proyectos de las organizaciones culturales que en la actualidad y de manera constante acometen la tarea de hacer más accesibles sus colecciones y actividades, entre las que se incluyen las publicaciones impresas tradicionales, y proyectos de licencias, televisión, radio y video.

Nos concentramos aquí en las leyes y las obras de Estados Unidos protegidas por las leyes de Estados Unidos. No obstante se introducen algunos aspectos del sistema legal internacional. La presentación del material en Internet es inherente al ámbito internacional, de manera que es necesario al menos tener conocimiento del más amplio contexto de acceso y uso en el ámbito global de los materiales que uno va a colocar en línea. Como tantos de los participantes de los cursos de la Escuela de Escaneo (School for Scanning) han sido empleados del gobierno federal de Estados Unidos, también se tratarán, para su información, unos pocos problemas singulares de los proyectos federales.

Derecho de autor

El derecho de autor en Estados Unidos es el derecho exclusivo que poseen los autores sobre sus obras originales. Este existe desde el momento de la *fijación* en un medio de expresión tangible (incluido el software). Incluye el derecho que tiene el autor a controlar la reproducción, las copias, la exhibición, la ejecución, así como otros usos que se le den a la obra.

¿Qué leyes rigen el derecho de autor?

En Estados Unidos, existen varios cuerpos legales que rigen el derecho de autor. El primero es la Constitución de Estados Unidos, Artículo 1, Sección 8, que establece que: *El Congreso tendrá facultad para fomentar el progreso de la ciencia y de las artes útiles, garantizando por tiempo limitado, a los autores e inventores el derecho exclusivo a sus respectivos escritos y descubrimientos.* Esta breve declaración implica,

con simplicidad brillante, que la intención de la disposición sobre el derecho de autor es que la sociedad se desarrolle al promover la creatividad a la vez que reconoce la realidad de la naturaleza humana -- para alentar a los creadores a que compartan sus obras para el mejoramiento de la sociedad, es necesario que exista una forma de permitirles que se beneficien con su trabajo. James Madison describió este concepto del derecho de autor en los Papeles Federalistas: *El público coincide plenamente . . . con los reclamos de los individuos.*

El Congreso de los Estados Unidos interpreta la constitución y produce leyes que se plasman en el Código Civil de los Estados Unidos. Es en ese cuerpo legal donde encontramos la Ley de Derecho de Autor de Estados Unidos vigente. Esta es sólo la más reciente de varias anteriores que se han ido desarrollando, cada una de las cuales ha reflejado un nuevo intento de abordar las nuevas tecnologías y nuevos modelos de negocios. Por ejemplo, la fotografía es un soporte que no estaba protegido hasta los años 1870, y las películas cinematográficas no estuvieron protegidas tampoco hasta después de comenzado el siglo XX. Como muchos proyectos abordan colecciones históricas, se necesitará estar familiarizado, al menos de forma básica, con la Ley de Derecho de Autor de 1976, que es la que está vigente en la actualidad, con sus enmiendas, así como con la Ley de Derecho de Autor de 1909, la que anteriormente estaba vigente. Las reglas para calcular la duración del derecho de autor difieren, de manera que tendremos resultados diferentes en dependencia de la naturaleza del material. También el acto de *publicación* (que en este caso es un término de arte) fue de importancia vital según la Ley de 1909. Aunque en determinadas circunstancias, todavía es importante para la Ley de 1976, el acto de publicación ya no determina la duración del derecho de autor de la manera rígida en que ocurría en la ley anterior. La Publicación y el cálculo de duración del derecho de autor se abordan más adelante.

Otro ejemplo de una Ley promulgada por el Congreso es la Ley de Derecho de Autor del Milenio Digital (DMCA) de 1998. Esta es una nueva ley, que se promulgó como parte de la Ley de Derecho de Autor ya existente en 1998. Esta legislación está todavía en discusión, así como también su impacto de largo alcance. La DMCA puso en práctica ciertas obligaciones de los tratados internacionales de derecho de autor de Estados Unidos en conjunción con tratados promulgados bajo los auspicios de la Organización Mundial de la Propiedad Intelectual (OMPI). Otras áreas importantes que se abordan en la DMCA incluyen:

- Limitaciones de ciertas responsabilidades de los proveedores de servicios en línea

- Exención de infracción de las leyes de derecho de autor por hacer una copia de un programa de computadora si la copia se hace con el propósito de mantenimiento o reparación, y otras disposiciones misceláneas referentes a la educación a distancia

- Exención en la Ley de Derecho de Autor para las bibliotecas y archivos no lucrativos para que puedan hacer copias de preservación de ciertos materiales en formato digital

- "Webintroducción" de grabaciones sonoras en Internet.

Muchos aspectos de la DMCA pueden estar relacionados con proyectos emprendidos por las comunidades educacionales y culturales. La Oficina de Derecho de Autor de Estados Unidos ha preparado un resumen muy útil de la DMCA, que se puede obtener a través de Internet en el sitio http://www.loc.gov/copyright. Búsquelo bajo el acápite *Legislación.* La Oficina de Derecho de Autor de Estados Unidos también fue encargada de preparar un estudio acerca del impacto de la educación a distancia sobre el derecho de autor y acerca de la posible necesidad de cualquier nueva exención especial. Este detallado estudio también se puede obtener en el sitio Web http://www.loc.gov/copyright/docs/de_rprt.pdf.

El siguiente cuerpo legal a tener en cuenta es el derecho común o derecho basado en precedentes judiciales. Este es un derecho creado por los jueces según van interpretando la aplicación del Código Civil de Estados Unidos (y la Ley de Derecho de Autor incluida en el Código Civil) y otras leyes a un grupo de hechos particulares que se presentan ante un tribunal por los litigantes en un caso legal. Este capítulo sólo analiza unos pocos ejemplos, especialmente en el área de uso razonable, de manera que los resultados difieren en dependencia de los hechos en cada caso. Esta es una de las razones por las que es importante identificar qué es lo que se quiere hacer y por qué y cómo se piensa hacer. Estos son pasos importantes en el proceso de planificación de obtención de los documentos legales necesarios para los proyectos.

Por último, puede que también sea necesario tener en cuenta acuerdos y tratados internacionales, como la Unión Internacional para la Protección de Obras Literarias y Artísticas (conocida como "Convención de Berna"). La Convención de Berna proporciona mutualidad en la protección al derecho de autor entre una serie de naciones (las que son signatarias del tratado). Estados Unidos la firmó en 1989, aunque la Convención de Berna comenzó mucho antes, en la década de 1890. En general, las naciones tienden a no adherirse a la misma hasta que tienen algún interés comercial para hacerlo. Por ejemplo, hasta que la industria del cine y la música presionaron para que Estados Unidos firmara la Convención de Berna en interés de proteger las obras de Estados Unidos distribuidas o producidas en el extranjero, el Congreso no tuvo una razón de peso para

considerar seriamente el asunto. Presiones comerciales similares en la arena internacional condujeron a la puesta en práctica de una ampliación de los términos de derecho de autor bajo las leyes de Estados Unidos (que se abordan junto con la duración del derecho de autor más adelante).

Básicamente, la Convención de Berna ofrece protección mutua a las naciones miembros, a lo que también se denomina *tratamiento nacional.* En otras palabras, los estados miembros deben tratar a los extranjeros en términos al menos iguales a los que tratan a sus propios ciudadanos. Cuando se trata de distribución y acceso en Internet, no está claro en muchos aspectos como tratar con ciertas cuestiones en cuanto a qué ley y que jurisdicción se aplican. En Estados Unidos existe un cuerpo legal cada vez más amplio que crece en la medida en que los diferentes estados abordan las cuestiones jurisdiccionales análogas con respecto a cuáles leyes de cuál estado se aplican a una situación en particular. Profesores de derecho y expertos en leyes también abordan en diversos artículos la cuestión "¿cuál es la ley que se aplica?". Además se trata este asunto en el área de impuestos sobre las transacciones que se realizan en Internet. Es un área que debe observarse.

Un aspecto que debe mencionarse dentro del contexto de la Convención de Berna es el concepto de los derechos morales, que se describe en el Artículo 2b del tratado. El derecho moral (o *droit moral*) protege ciertos intereses ostensiblemente no comerciales de los creadores, fundamentalmente la *atribución* y la *integridad.* El derecho de atribución protege al creador al exigir que se reconozca o identifique al creador como creador de la obra, y así evitar que otras personas firmen con su nombre, de manera falsa, y al evitar que se atribuya la obra de otras personas a ese creador. Está relacionado con el derecho de una persona a asociarse con el producto de su propia creación. El derecho de integridad evita que otras personas distorsionen, mutilen, o tergiversar la obra del creador en una manera que afecte su honor o reputación de forma negativa. Por lo general, los derechos morales no se pueden transferir--son independientes de los derechos económicos y acompañan al creador incluso después que la obra ha sido vendida.

Este tipo de derecho no es enteramente consecuente con las nociones tradicionales de propiedad bajo las leyes de Estados Unidos. Por tanto, para poder cumplir con los términos de la Convención de Berna y que Estados Unidos pudiera firmar el tratado, se señalaron otras áreas de las leyes del país que proporcionan protección análoga a los creadores, como por ejemplo, las marcas registradas, los contratos y las leyes de competencia injusta. El Congreso enmendó la Ley de Derecho de Autor de Estados Unidos e incluyó la Ley de Derechos de las Artes Visuales de

1990 como Sección 106A (VARA). La VARA es válida sólo para una categoría muy limitada de obras de arte, entre las que se incluyen:

- Cuadros, dibujos, impresos (grabados), y esculturas en ejemplares únicos (una sola copia) o en ediciones limitadas de hasta 200 copias o menos, si están firmadas o numeradas por el artista

- Fotografías si están producidas sólo para exhibiciones y existen en una sola copia firmada por el artista o en edición limitada de hasta 200 copias firmadas y numeradas por el artista.

Observe que la VARA no se aplica a cualquier tipo de obra de creación. Se excluyen obras como (pero no sólo estas) carteles, mapas, globos, películas cinematográficas y otras obras audiovisuales, revistas, libros, publicaciones periódicas, periódicos y anuncios publicitarios.

Los derechos morales según la Convención de Berna y la VARA son derechos únicamente del artista -- y no del propietario de la obra si la obra ha sido transferida o asignada a otra persona. En la mayoría de los países europeos, los derechos morales no se pueden transferir o asignar. En Estados Unidos sucede lo mismo, sin embargo, en Estados Unidos un artista puede renunciar a sus derechos morales, cosa que debe hacer por escrito, declarando cuál es la obra y los usos para los que es válida esa renuncia. Observe que, aparentemente, los derechos morales según la VARA no se extienden a las obras electrónicas.

La Ley de Derecho de Autor

Como se señala anteriormente, el derecho de autor en Estados Unidos es el derecho exclusivo de los autores sobre sus obras originales. Existe desde el momento de la *fijación* en un medio de expresión tangible (incluidos los softwares). Incluye también el derecho del autor a controlar la reproducción, exhibición, ejecución, y otros usos de una obra. Observe que "autor" es la palabra utilizada en la ley para referirse a cualquier creador de una obra a la que se le pueda aplicar el derecho de autor, independientemente de que sea un escultor, un pintor, un fotógrafo, un escritor, o algún otro tipo de creador. La Ley de Derecho de Autor protege estos derechos en un amplio rango de productos de creación, incluidas las obras literarias; las musicales incluidas las de acompañamiento; las obras dramáticas; las obras coreográficas y las pantomimas; las pictóricas, gráficas y escultóricas, las películas cinematográficas y otras obras audiovisuales; las grabaciones sonoras; y las obras arquitectónicas.

La Ley protege también el derecho del autor a crear y controlar la creación de *obras derivadas.* Una obra derivada es una obra que se basa en una o más que ya existe, como una traducción, un arreglo musical, una

dramatización, una ficcionalización, una versión cinematográfica, una grabación de audio, la reproducción de una obra de arte, un compendio, una condensación, o cualquier otra forma en que se pueda rehacer, transformar, o adaptar una obra. Una obra que consiste en revisiones editoriales, anotaciones, elaboraciones, o cualquier otra modificación que, en su conjunto, represente una obra original de un autor es una obra derivada.

Es de especial importancia que se comprenda bien esta área par los proyectos de digitalización, porque muchas personas creen que la copia digital de una obra que ya existe es una obra derivada basada en un objeto existente. También es importante en relación con el concepto de *originalidad.* Una obra está sujeta a la protección del derecho de autor si es lo suficientemente original, y existe cierta duda con respecto a si una copia (digital o de otro tipo) de una obra que ya es del dominio público (es decir, sin protección alguna de derecho de autor) es una obra protegible independiente. Un tribunal en New York, en el caso *Corel v. Bridgeman Art Library,* ha dictaminado que en ese caso no puede haber una nueva protección para esa copia. Aquí podemos ver cómo la línea conceptual puede ser poco definida para algunas personas en la determinación de la diferencia entre las obras derivadas protegibles controladas por el propietario del derecho de autor y las copias de obras de dominio público realizadas en un nuevo soporte. Muchas organizaciones culturales y de negocios confunden esos conceptos y confían en conclusiones algunas veces inexactas cuando crean productos y reproducciones con el objetivo de generar ingresos.

Aspectos a tener en cuenta

En el área del derecho de autor, es importante tener en cuenta los siguientes aspectos:

- El derecho de autor protege la expresión de una idea, no la idea misma.
- Desde que Estados Unidos firmó la Convención de Berna en 1989 no es necesaria la notificación de derecho de autor (es decir, la "c" encerrada en un círculo (©), o la palabra "Copyright" o "Copr." con la fecha y el nombre del propietario) y la inscripción. No obstante, estas notificaciones son de importancia crucial para determinar si una obra más antigua está todavía protegida por el derecho de autor. Según legislaciones anteriores, la publicación sin notificación inserta una obra en el dominio público. Siempre es conveniente registrar las obras nuevas de manera que las personas que estén interesadas en obtener una licencia puedan encontrar al propietario del derecho de autor, y también porque se exige la inscripción en la Oficina de Derecho de Autor de Estados Unidos para establecer un pleito judicial por infracciones del derecho de autor y para obtener compensación por daños causados por violaciones.

- La ausencia de la notificación no quiere decir que no exista protección para el derecho, en particular en el caso de obras nuevas o para obras no publicadas (incluso las más antiguas)

- La posesión o propiedad de un artículo físico no quiere decir que el que la tiene en su poder o el propietario de la obra física es el propietario del derecho de autor.

Excepciones para el derecho de autor

A todas las obras no se le aplica la protección del derecho de autor, ni tampoco el derecho de autor es ilimitado en su duración. Se pretende que el derecho de autor sea un monopolio limitado que permita a los autores beneficiarse de su propio esfuerzo creador y, eventualmente, que el público tenga libertad para usar las obras de creación con la idea de que usa uso fluido inspire la nueva creación que en última instancia beneficie a toda la sociedad. Cuando una obra resultado de la creación no está protegida por el derecho de autor, se le denomina como "del dominio público".

Las obras creadas por empleados del gobierno de los Estados Unidos en el ejercicio de su trabajo son obras del dominio público según la Ley de Derecho de Autor. Esta es una cuestión de política pública, basada en la teoría de que los ciudadanos ya han pagado por ese trabajo con la contribución de sus impuestos. Este presupuesto no se aplica ni a las obras de los empleados estatales en Estados Unidos (los estados pueden hacer valer, y con frecuencia lo hacen, el derecho de autor de las obras de sus empleados) ni a las obras de los empleados de otras naciones.

La duración del derecho de autor en Estados Unidos puede variar en dependencia de si una obra ha estado sujeta a la Ley de 1909 o a la Ley de 1976 enmendada. Tenga en cuenta que existen disposiciones en la Ley 1976 enmendada que pueden aplicarse a obras anteriores. Por lo general, según la Ley de 1909, se protegían las obras por hasta dos períodos de 28 años, para un total de 56 años de protección posible. Según la Ley de Derecho de Autor de 1976 enmendada por la Ley de Extensión de Términos Sonny Bono efectiva a partir de octubre de 1998, las obras están protegidas durante toda la vida del creador más un período de 70 años. Para las obras conjuntas (las obras creadas por dos autores o más), la duración es de 70 años después de la muerte del último de los autores que haya sobrevivido. Para las obras anónimas, las escritas bajo seudónimo, y las obras realizadas por encargo, la duración es el término más corto entre 75 años a partir de la publicación o 100 años a partir de la creación.

Observe que una *obra por encargo* es una obra creada por un empleado dentro del ámbito de su empleo. Se supone que el derecho de autor en

obras realizadas por contratistas independientes debe permanecer con el creador -- es decir, el contratista independiente. Específicamente se supone que las obras no son obras por encargo a menos que exista un acuerdo por escrito que indique lo contrario firmado por el contratista independiente antes del comienzo de la ejecución de la obra. Si Ud. contrata a alguien para que escriba un informe o para que produzca para su organización, y si Ud. quiere que la organización retenga el derecho de autor sobre la obra, es de crucial importancia que firme un acuerdo por escrito con el contratista antes del comienzo de la realización de la obra, en el que establezca la situación de la obra en cuanto a la posesión de los derechos de autor. Muchos contratistas se niegan a firmar ese tipo de documento, de manera que tendrán que buscar otras vías para lograr sus objetivos, por ejemplo, obtener algún tipo de licencia que le permita lograr sus objetivos a la vez que deja al contratista alguna vía de escape para hacer algún otro uso del producto de la obra. No obstante, si el autor produce algo que sea de naturaleza confidencial, Ud. deberá hacer todos los esfuerzos posibles para obtener de manera bien clara todos los derechos o para abordar el asunto por escrito antes de que comience la obra.

El cuadro que aparece a continuación le ofrece una regla bastante útil para calcular la duración del derecho de autor en Estados Unidos. Tenga en cuenta que este cuadro no es tan aplicable a las grabaciones de audio y las obras audiovisuales.

CUÁNDO UNA OBRA PASA A SER DEL DOMINIO PÚBLICO

Incluye material de la Ley de Extensión de Términos, PL 105-298

FECHA DE LAOBRA	**PROTEGIDA DESDE**	**TÉRMINO**
Creada el 1-1-78 o después	Cuando se fija la obra en un soporte de expresión tangible	La vida + 70 años[1] (o si la obra es de autoría conjunta el término más corto entre 95 años a partir de la publicación, o 120 años a partir de la creación)[2]
Publicada antes de 1923	En dominio público	Ninguna
Publicada entre 1923 - 63	Cuando se publicó con notificación[3]	28 años + se pudiera renovar por 47 años, ahora se extiende por 20 años para un total de 67 años de renovación. Si no se renovó, ya es del dominio público
Publicada entre 1964 - 77	Cuándo se publicó con notificación	28 años para el primer término; ahora extensión automática de

		67 años para el segundo término
Creada antes del 1-1-78 pero no publicada	1-1-78, fecha efectiva de la Ley de 1976 que eliminó el Copyright del derecho común	La vida + 70 años, o el 31-12-2002, cualquiera que sea el más largo
Creada antes del 1-1-78 pero publicada entonces y el 31-12-2002	1-1-78, fecha efectiva de la Ley de 1976 que eliminó el Copyright del derecho común	La vida + 70 años, o el 31-12-2002, cualquiera que sea el más largo

[1] El término de las obras conjuntas se mide por la vida del autor que más tiempo viva.
[2] Obras por encargo, anónimas, y las escritas bajo seudónimo también se rigen por este término. 17 U.S.C. &sec; 302(c).
[3] Según la Ley de 1909, las obras publicadas sin notificación pasaron a ser del dominio público al ser publicadas. Las obras publicadas sin notificación entre el 1-1-78 e el 1-3-89, fecha efectiva de la Ley de Ejecución de la Convención de Berna, retuvieron el derecho de autor sólo si la inscripción se realizó dentro de un período de cinco años. 17 U.S.C. &sec; 405.

Notas cortesía del Profesor Tom Field, Franklin Pierce Law Center
Lolly Gassaway

Observe que en este cuadro se menciona la palabra "publicación". Como se explica anteriormente, publicación es un término de arte que se define en la Ley de Derecho de Autor. Por lo general se refiere a la venta o traspaso de una obra o una copia de una obra, pero no siempre es un cálculo sin ambages. Por ejemplo, la presentación de una obra no es una "publicación" para los propósitos del derecho de autor. Un tribunal consideró que el discurso de Martin Luther King Jr. conocido bajo el título "Yo tengo un sueño" no era una "publicación" a pesar de la presencia de miles de personas, la invitación de la prensa, y la distribución de copias del discurso a representantes de la prensa. Utilice el cuadro como regla, pero siempre tenga en cuenta que publicación puede ser un concepto muy complicado y que puede requerir en algunos casos que Ud. busque asistencia legal.

¿Qué es el Uso Justo?

Uso justo es un concepto singular de la legislación de Estados Unidos que proporciona defensa para las infracciones contra el derecho de autor en ciertas situaciones especiales. Si un uso es uso justo, entonces uno puede copiar una obra protegida sin permiso del propietario del derecho de autor y sin que se considere el uso como violación. Otros países tienen un concepto similar al que se denomina "trato justo", pero este concepto es por lo general más restrictivo y cargado de requisitos administrativos. El

concepto de uso justo se desarrolló en la legislación común o basada en precedentes de Estados Unidos (legislación hecha por los jueces) como forma equitativa de tratar ciertos tipos de copias que, en circunstancias específicas, se consideraban excusables por razones de política pública. El uso justo sólo existía en el derecho común hasta que se incluyó en la Ley de Derecho de Autor de 1976, que organiza estos conceptos de uso justo emanados de una larga lista de casos en un enunciado relativamente corto y que aún deja mucho a la interpretación personal en cada situación.

Uso justo es una excepción para los derechos exclusivos de los propietarios de derecho de autor, y sólo se dispone del mismo para usos limitados y con propósitos tales como la crítica, el comentario, información noticiosa, enseñanza (incluida la realización de copias múltiples para el uso en el aula de clases), becas o investigaciones. En dependencia del balance de los siguientes cuatro factores, la copia de obras con este objetivo no constituye una infracción bajo las leyes de Estados Unidos. La prueba de cuatro factores requiere una evaluación de cada uno de los grupos de factores bajo un balance de todas las consideraciones siguientes -- ninguna de las cuales es dispositiva en si misma.

¿Cuál es el propósito y el carácter del uso? ¿Es el uso intendido para propósitos comerciales o no comerciales? ¿El uso hace algo que sea de alguna manera *transformador*-- es decir, utiliza o copia la obra misma como una parodia que inherentemente requiere el uso de la obra en su totalidad?

¿Cuál es la naturaleza de la obra? ¿Es una obra de ficción o no es de ficción? Las obras de ficción tienden a recibir más protección porque se consideran "creativas" y, por tanto, están menos sujetas a los argumentos del uso justo que las obras factuales.

¿Cuál es la cantidad y la importancia de la parte de la obra usada en relación con el todo? ¿La copia utiliza un verso de un poema haiku o un poema épico de varios volúmenes? ¿La copia es una de las últimas páginas de una novela de misterio que, aunque sólo representa una pequeña porción de la obra, puede destruir el interés del comprador en comprar el libro pues ya conoce el final? Y esto nos conduce al último factor y con frecuencia al que más peso se le concede (aunque no siempre sucede así):

¿Cuál es el efecto del uso en el mercado potencial -- o en el valor de -- la obra protegida por derecho de autor? Observe que esto se concentra a la vez en el valor existente y en el mercado potencial de una obra. Los propietarios del derecho de autor están reactivando obras antiguas y las vuelven a poner en circulación de muchas maneras utilizando la

creatividad, con nuevas tecnologías. Por tanto, es más complicado argumentar que un libro no tiene mercado potencial porque está agotado. Este puede ser un factor difícil de superar si Ud. piensa en hacer una copia digital de una obra protegida para presentarla en Internet bajo el argumento de uso justo.

Algunos casos recientes demuestran la aplicación del uso justo, entre ellos casos en los que el hecho de que el uso era comercial por naturaleza el dictamen fue no dispositivo. En otras palabras, se dictaminó uso justo a pesar de la naturaleza del uso comercial. También se recuerda el caso del tribunal Supremo de 1994 *Campbell v. Acuff Rose Music, Inc.* En el que el grupo de rap 2LiveCrew fue demandado por utilizar la conocida canción de Roy Orbison *Pretty Woman* sin permiso o pago de royalties. El uso por 2LiveCrew's se dictaminó como uso justo a pesar de su naturaleza comercial. Los abogados del grupo argumentaron con todo éxito que la versión de 2LiveCrew no era una copia *per se* sino una parodia que exigía utilizar el "corazón de la obra" para lograr el objetivo que se planteaban. Este es un ejemplo de uso transformador.

En el caso *Leibovitz v. Paramount Pictures Corporation,* de 1998, un tribunal dictaminó uso justo en otra situación también con uso comercial, en la que, para poder hacer una parodia, era necesario utilizar el tema de la obra. La famosa fotografía de Demi Moore que se utilizó para la portada de la revista *Vanity Fair* y en la que aparecía la actriz desnuda en avanzado estado de gestación fue parodiada en un anuncio para la comedia *Naked Gun*. La estrella masculina de la película posó imitando a Demi Moore. La imagen original de la actriz se copió hasta el más mínimo detalle. El tribunal evaluó los detalles de todos los factores (cuatro). En este caso, la obra original fue copiada en su totalidad para un propósito comercial. No obstante, era válido el dictamen de uso justo para el objetivo de la parodia, el uso iba hasta el corazón del original -- era necesario hacer la copia para poder hacer la parodia. Se dictaminó que el uso era transformador y justo a pesar de la naturaleza del uso y del hecho de que no se solicitó permiso alguno ni se pagó nada tampoco. Observe que no se requiere pago de licencia se aplica uso justo.

Las situaciones en las que no se consideró que existía uso justo incluyen la reproducción de un cartel protegido por derecho de autor de una obra de arte por unos breves segundos en un anuncio de televisión. El cartel estaba visible en el fondo del set en el caso de *Ringgold v. Black Entertainment Television* de 1997. El artista que había creado el cartel y que poseía los derechos de autor y la obra original en la que se basaba el cartel presentó una demanda por violación de derecho de autor contra el productor y la emisora del programa de televisión. La obra se utilizaba como decoración en el set en una forma que en última instancia se

dictaminó significativa para los propósitos del derecho de autor, porque el uso tenía el mismo propósito decorativo para el cual se vendió el cartel. Esta situación decidió la balanza en contra del productor en cuanto al factor del uso justo. Además, el factor de uso justo que aborda el efecto del uso en el mercado potencial para la obra del artista también pesaba en contra del productor, porque el uso era comercial y utilizaba la obra sin transformarla. El productor utilizaba la obra completa para la que el artista ya tenía licencia de ganancias, y por lo tanto afectaba de manera negativa los ingresos del artista al no haber obtenido permiso y no haber pagado por la licencia para utilizar la obra. En este caso no había parodia. Obsérvese que el cartel apareció en pantalla sólo unos segundos.

En muchas situaciones es difícil determinar si el uso justo se aplica o no. Este es un concepto que se desarrolló en principio para aplicar a texto, lo que hace difícil aplicar todos los factores a fotos, obras de arte y otros materiales que no sean texto.

Otras Consideraciones

El derecho de autor está en el punto de mira como el aspecto más importante para los proyectos de digitalización, pero hay también otros conceptos legales que requieren atención. A continuación aparece una breve descripción de algunos de estos conceptos.

El derecho de publicidad

Por lo general el derecho de publicidad se asocia con las figuras públicas, con frecuencia, aunque no siempre celebridades y artistas. El derecho de publicidad aborda el aspecto de la ganancia comercial por el nombre de una persona, por el parecido, por la voz, o por otros aspectos explotables de la personalidad. Las leyes varían de un estado a otro. A diferencia del derecho de autor, no existe una ley federal por el derecho de publicidad, lo que hace difícil determinar qué ley se aplica en cada situación en particular. El hecho de que en algunos estados este derecho continúa después de la muerte de la persona complica aún más las cosas. En otros estados termina a la muerte del sujeto. Ahora es cosa de rutina los anuncios en los que se transfiere en formato digital películas de celebridades ya fallecidas -- desde John Wayne vendiendo cerveza, hasta Elvis Presley vendiendo pizza, hasta Fred Astaire bailando con aspiradoras.

Aunque muchos estados cuentan con leyes referentes al derecho de publicidad, si uno está trabajando con materiales que puedan presentar problemas de publicidad y estén asociados en particular con los estados de Tennessee, California, o New York, preste mayor cuidado, ya que

estos estados cuentan con leyes bien desarrolladas en esta área. Se puede obtener el permiso de personas que poseen derechos de publicidad directamente con ellas -- aunque casi siempre se obtienen los permisos con los representantes o los abogados. Muchas veces, en dependencia de la naturaleza del uso, se puede obtener una licencia gratuita o al costo nominal. Encontrar a la persona adecuada a quien se debe contactar es fundamental, y un trabajo casi detectivesco. Como los derechos de publicidad abordan derechos económicos, puede argumentarse que no se han violado los derechos de publicidad si el uso es estrictamente no comercial. No está muy definido el alcance de una posible responsabilidad penal por utilizar ese tipo de materiales en Internet, y por tanto contribuir al uso ilegal por otra persona. Siempre sería prudente insertar advertencias en el sitio web que explique a los usuarios que será necesario que ellos obtengan los derechos adecuados si copian el contenido de esa página para sus propios propósitos comerciales, aunque esas advertencias no lo protegen o aíslan a uno de posibles responsabilidades penales.

Derecho de Privacidad

Con frecuencia se denomina el derecho de privacidad como "el derecho de que lo dejen a uno en paz". A diferencia de los derechos de publicidad, este cuerpo legal por lo general se refiere a ciudadanos privados más que a celebridades, aunque existen significativas excepciones. También en contraste con los derechos de publicidad, los derechos de privacidad no son comerciales por naturaleza y protegen a las personas de las intrusiones en su soledad o sus asuntos privados, de la revelación al público de información privada, y de que sean presentados al público bajo una falsa luz. Si se está trabajando con materiales que colocan a personas privadas en situaciones embarazosas o que documentan de alguna manera la vida privada de esas personas, como por ejemplo, fotografías de personas desnudas (en las que las que se identifica a las personas) debe tenerse especial cuidado. Pudiera ser cuando menos inapropiado y en el peor de los casos ilegal publicar esos materiales en un sitio Web al que tiene acceso el público. Si los materiales son lo suficientemente intrusivos o embarazosos, la probabilidad de obtener el permiso de la persona es mínima. El derecho de privacidad por lo general caduca a la muerte del sujeto.

Difamación: Libelo y Calumnia

La difamación incluye la calumnia y la detracción. Implica la publicación a una tercera parte de materiales escritos falsamente (libelo) o comentarios hablados (calumnia) que ponen en evidencia a personas vivas y las expone al odio, el desprecio, o el ridículo. Esta área del derecho varía de un estado a otro. Si se está trabajando con materiales

sobre una persona ya fallecida, tenga en cuenta que el derecho por lo general caduca a la muerte de la persona. No se puede difamar a un muerto. Como cuestión práctica, pueden existir teorías legales alternativas en las cuales los hijos o los herederos pueden basar un pleito judicial, así que piense con cuidado antes de poner en Internet materiales que puedan ser difamatorios. Esta cuestión puede surgir en proyectos de digitalización de correspondencia o papeles administrativos en los que el escritor hace varias décadas hizo algún tipo de comentario negativo por supuesto sin pensar que esos comentarios saldrían a la luz pública.

Obscenidad y Pornografía

La obscenidad y la pornografía son cuestiones tan complejas y amplias de las leyes que será suficiente en este caso mencionarlas. Estén preparados para reconocerlas y abordarlas según sea necesario. La luz roja en los proyectos de digitalización incluye los casos de desnudez, sobre todo en el caso de los niños, y cualquier representación de menores que aparezcan en una postura o conducta explícitamente sexual. En la actualidad a nivel de estado y también a nivel federal se están realizando esfuerzos para controlar la distribución de materiales pornográficos a través de Internet, en especial la pornografía infantil, y también se están realizando grandes esfuerzos para proteger a los niños usuarios de Internet para que no tengan acceso a materiales más adecuados para los adultos. Los primeros esfuerzos en esta dirección se consideraron inconstitucionales. Por ejemplo, la Ley para la Decencia en las Comunicaciones por vía Electrónica y toda su progenie fueron consideradas inconstitucionales en parte basándose en la teoría de que controlar la pornografía según establece esa ley teóricamente protegería a los niños con toda probabilidad a expensas del derecho de los adultos según la Primera enmienda a la Constitución, y en parte basándose en la suposición práctica de que los menores usuarios pueden ser protegidos de tener acceso a la pornografía a través del uso de softwares de filtro. Es importante cooperar con los esfuerzos legislativos en este sentido. Con toda probabilidad, el balance adecuado se logrará eventualmente en el punto medio entre el manejo de contenido ilegal como por ejemplo la pornografía infantil en línea y la Primera Enmienda a la Constitución.

Sensibilidad en cuanto a los contenidos

En la práctica, esta es un área que no siempre será de interés legal, pero deberá abordarse en todo proyecto de digitalización. Si se está trabajando con materiales antropológicos o con materiales que involucran personas que han sido fotografiadas en contra de su voluntad o en situaciones de explotación, como los prisioneros de guerra, pueden ser sensibles en cuanto al contexto en el que se recopiló el material. Determine si es adecuado consultar con los sujetos o los descendientes si fuera posible.

Por ejemplo, hay que otorgar especial atención a los materiales que involucran a los Indios norteamericanos, sus objetos sagrados y sus ceremonias. Examine el contexto en el que el se recopiló el material original y tenga en cuenta la manera en que presentará dicho contenido.

Ley de Libertad de Información

Esta es un área de interés para los proyectos producidos por agencias federales del gobierno de los Estados Unidos en particular. La Ley de Libertad de Información exige que el gobierno proporcione acceso del público a ciertos registros, fundamentalmente las de la administración del estado y las de las cuestiones de política como descripciones organizativas y procedimientos. Los sitios y páginas Web son herramientas de costo eficiente para proporcionar acceso a materiales que se solicitan constantemente y que se puede exigir por vía legal que se pongan a disposición del público. Advertencia: Algunos materiales específicamente están exentos de acceso por parte del público, por ejemplo, aquellos materiales cuya publicación puede constituir una invasión no autorizada de la privacidad (*ej.* registros de personal o registros médicos), ciertas cuestiones relativas a la defensa nacional y la política exterior, secretos comerciales, información privilegiada o confidencial comercial o financiera y ciertos registros relativos a la puesta en vigor de las leyes.

Vínculos

En años reciente se han producido varios casos que se refieren a los vínculos (o hipervínculos) entre los sitios Web. Uno de los casos que más publicidad ha recibido es el caso de Ticketmaster y Microsoft. Microsoft se vinculó a un área del sitio Web de Ticketmaster "pasando" la página principal. Cualquiera que accediera al sitio de Ticketmaster por la vía del vínculo de materia de Microsoft pasaría por alto la publicidad sobre la página principal de Ticketmaster. Ticketmaster argumentó que, en última instancia esta acción privaría a la compañía de sus ingresos por publicidad, mientras que Microsoft declaró, entre otras cosas, que el vínculo en realidad creaba más exposición para los productos de Ticketmaster e incrementaría las ventas de esa compañía. En este caso se llegó a un arreglo.

Más recientemente, la Iglesia de Jesús Cristo -Todos los Santos demandó al Ministerio encargado de Farolas del estado de Utah en un tribunal federal distrital en ese estado por crear hipervínculos en una página Web que contenía trabajos de la Iglesia protegidos por derecho de autor sin haber pedido permiso alguno. En diciembre de 1999, el tribunal emitió una orden preliminar que prohibía al Ministerio la utilización de dichos vínculos. Este caso plantea el aspecto de *infracción contributoria.* Según

el caso, si Ud. crea un vínculo con sitios que contienen materiales que infringen o violan un derecho de autor, puede ser acusado por infracción contra derecho de autor, a pesar de que Ud. no colocó ("copió") el material Ud. mismo. El caso además tenía ramificaciones escalofriantes por la fluidez de Internet así como las implicaciones referentes a la Primera Enmienda a la Constitución. Este caso debe estudiarse cuando Ud. esté planificando la preparación de su sitio Web -- una búsqueda de los nombres de las partes en cualquier sistema de búsqueda le proporcionará referencias.

Se ha hecho práctica común obtener permiso para establecer vínculos con otros sitios -- o al menos notificar a dicho sitio. Esto es válido en especial con los sitios comerciales. Es un enfoque conservador, y por supuesto no es un requisito en este momento. Al igual que muchas otras cuestiones, es importante tener en cuenta el contexto del vínculo. Por analogía, no es necesario solicitar permiso alguno para citar una obra en una biografía. Si estamos creando una lista de recursos de referencia en un sitio educacional no comercial, esta es una cuestión de interés limitado. Debe navegarse con cuidado si se está estableciendo un vínculo de alguna forma que implique algún tipo de endose o apoyo, o de lo contrario puede echar a rodar alguna bola como la de las marcas registradas o la competencia desleal. La existencia y el alcance de la responsabilidad en estas áreas no están bien delimitados todavía.

El procedimiento a seguir

Un tema de este análisis es la naturaleza en desarrollo del ambiente legal según surgen nuevos hechos y situaciones *vis à vis* con Internet. El primer paso en el avance en esta cambiante escena es educarse uno mismo. Obtener el sentido general o la imagen del asunto; este artículo es el punto de partida. Una vez que se ha logrado cierta familiaridad con los aspectos legales y también con los financieros y administrativos, es aconsejable establecer dentro de la organización, algún lineamiento o política para poder desarrollar prácticas consecuentes y vías para abordar los proyectos.

Cada proyecto deberá incluir para su uso interno y para comunicación, una descripción del material seleccionado para digitalización. ¿Qué es, de dónde es, cómo y por qué ha sido seleccionado para digitalización? ¿El material es de interés extraordinario? ¿El entusiasmo del donante es lo que impulsa el proyecto? ¿El material se va a conservar en el proceso, o este es un proyecto en el se destruirán los materiales en el curso del esfuerzo, como en los proyectos de los llamados "libros quebradizos"? Determinar la procedencia de los materiales es un primer paso de

importancia crítica en la evaluación legal de la situación de los derechos del material en cuestión. ¿Cuáles fueron los términos de la transferencia -- el material fue entregado a su organización o fue comprado por la misma? ¿Hubo algún término o condición por escrito, por ejemplo una escritura de entrega de la donación o un testamento? En ese caso, hay que obtener una copia y leerla con cuidado. Puede ser necesaria la asistencia de un abogado para interpretar los términos o condiciones.

Una vez hecho todo lo anterior se está en condiciones de determinar si se necesita adquirir algún permiso para reproducir a través de Internet los contenidos seleccionados. Tenga en cuenta también que la situación legal del material original es un factor de importancia en cualquier proceso de selección. Si contamos con materiales interesantes muy afectados por la necesidad de pasar por todo el oneroso proceso de obtención de permisos, se pudiera determinar que, al hacer un balance, es más expedito analizar la posibilidad de reproducir otros materiales. De lo contrario, incluya el costo de la obtención de permisos, y también los gastos administrativos que eso implica (tiempo que debe dedicar el personal a esa tarea, fuentes de referencia, llamadas telefónicas, fax, envíos por correos...).

Cuando se está planificando un proyecto, es necesario tener en mente muy claro el propósito de la institución: presupuesto, tiempo y recursos en correspondencia. Los proyectos de digitalización por lo general requieren de una labor muy intensa -- ¡recuerden que esto no es simplemente el equivalente de fotocopiar! El tiempo del personal y de las otras personas que participan en el proyecto es costoso, y eso es algo que casi siempre se subestima. Decida, como cuestión práctica si está dispuesto y puede pagar por los permisos si fueran necesarios. Empiece a trabajar a partir de cartas de solicitud de permisos para reducir al mínimo los esfuerzos administrativos que Ud. necesita. Como política a seguir, puede que Ud. prefiera no incluir materiales que estén embargados por los derechos. Piense en qué forma Ud. espera que las personas utilicen las ofertas que aparecen en Internet y planifique alrededor de esas posibles expectativas. Documente todos sus esfuerzos. Planifique introducir notificaciones en línea que incluyan la información acerca de la obtención de los permisos y cómo contactar con las personas que poseen esos derechos ("metadatos legales"). Estas deberán estar incorporadas en un diseño gráfico de manera consistente, para que los usuarios puedan saber dónde encontrar la información sobre los derechos en forma predecible. Tenga en cuenta que los individuos que conceden los permisos de uso, en comparación con los de negocios, pudieran preferir que Ud. no incluya la dirección de su casa en la página Web por razones de privacidad. Por tanto es posible que Ud. tenga que servir como punto de contacto para referir las solicitudes de permisos de terceras partes que

desean utilizar los materiales que, con los correspondientes permisos, se han introducido en la página de Ud. -- otra consideración administrativa.

Es posible que Ud. desee estudiar las declaraciones de notificación que se utilizan en la página principal del sitio Web Memoria Americana como punto de partida. Observe que hay una declaración para la Memoria Americana en conjunto y una declaración específica de cada colección que se presenta en la página respectiva de cada colección. Por tanto, las notificaciones deberán ajustarse a las características de su proyecto, su organización y otras características que son únicas de sus esfuerzos. Las declaraciones de notificación ofrecen información importante a los usuarios de su sitio, aunque todavía no está bien definido hasta qué punto las notificaciones lo protegen a uno de ciertas responsabilidades. En contraste con las notificaciones de la Memoria Americana, muchos sitios Web incorporan notificaciones muy legalistas. El tono y el contenido varían en dependencia del proyecto, el nivel de seguridad de la organización y la asesoría legal que se haya podido obtener.

Declaraciones de Muestra de Notificación de la Memoria Americana

Declaración General
http://lcweb2.loc.gov/ammem/copyrit2.html

Gottscho-Schliesner (derechos comerciales reservados)
http://lcweb2.loc.gov/ammem/gschtml/gottres.html

Colección de conservación (colección en soporte mixto)
http://lcweb2.loc.gov/ammem/amrvhtml/consres.html

Palabras y Hechos (lo que se puede obtener con permisos)
http://memory.loc.gov/ammem/mcchtml/corres.html

Referencias para información acerca de las leyes y el mundo en línea

---

La "ley del ciberespacio" evoluciona a diario con el surgimiento de nuevas tecnologías, con el desarrollo de nuevas analogías legales y con la respuesta de nuevas leyes y opiniones por parte de los tribunales y el Congreso. Cuando hayan leído este documento, ya habrá nuevos materiales en línea. La práctica demuestra que muchos sitios Web se actualizan o se cambian con frecuencia. Algunas direcciones pueden desaparecer. Observe que el material que se cita se refiere al texto del sitio descrito.

Parte del interés de utilizar recursos en línea es que nunca se sabe lo que se va a encontrar. Es muy probable que cada sitio vincule con otra información interesante. Algunas de las direcciones que presentamos aquí, como la del sitio de la Oficina de Derecho de Autor, ofrecen recursos sólidos a los que uno se puede referir una y otra vez. Algunos incluso ofrecen actualizaciones con respecto a la legislación, como por ejemplo el sitio **THOMAS** de la Biblioteca del Congreso. Otras direcciones están más políticamente orientadas y contienen datos y comentarios sobre los acontecimientos legales relacionados con el ciberespacio. Esta lista no constituye un apoyo ni una recomendación para ningún sitio en especial, ni de ningún punto de vista de ninguno de ellos tampoco. Constantemente se genera una increíble cantidad de material sobre el tema, tanto impresa como en formato electrónico -- bajo ningún concepto puede considerarse que esta lista es exhaustiva. Es sólo un punto de partida para obtener información acerca de cuestiones legales relacionadas con el uso y con los participantes de Internet.

Los invito a que miren a su alrededor, busquen y se eduquen -- tanto acerca del estado actual de la legislación como de la variedad de opiniones acerca del futuro legal de Internet. Estos debates sólo comienzan con preguntas acerca del derecho de autor. Traten de mantenerse actualizados con los debates que se producen para poder comprender en general los conceptos que hay detrás de las leyes vigentes. Entonces apliquen el sentido común a su trabajo diario. Lo ideal es que tengan a su disposición a un abogado que los pueda asesorar. Coordinen sus actividades con la Oficina General de Asesoramiento Legal que corresponda, si fuera necesario. No hay nada que sustituya el asesoramiento legal calificado.

Comiencen por la página principal del sitio de la **Biblioteca del Congreso**:
http://www.loc.gov/

Este sitio los llevará a **National Digital Library Project/American Memory Collection.** Ahí, busquen ejemplos de restricciones y declaraciones de notificación. Busquen en otros sitios -- comerciales, no lucrativos, y gubernamentales -- para ver de qué manera las diferentes organizaciones manejan esta cuestión.

También de la página principal de la Biblioteca del Congreso, véase el sitio de la **Oficina de Derecho de Autor**. Este sitio proporciona una plétora de información sobre derecho de autor. Es indispensable obtener la información básica sobre derecho de autor. Busque también en Internet las circulares de la Oficina de Derecho de Autor, materiales indispensables que describen de forma simple los aspectos prácticos y las aplicaciones de las leyes de Derecho de Autor -- también hay

instrucciones acerca de cómo obtener de forma gratuita una copia impresa de las circulares. Este sitio es indispensable para obtener el texto de la legislación sobre derecho de autor, las convenciones internacionales y las actualizaciones sobre la legislación. Fíjese en la información sobre CORDS (The Copyright Office Electronic Registration, Recordation and Deposit System [Registro, Recordatorio y Sistema de Depósito por vía Electrónica de la Oficina de Derecho de Autor]), un interesante proyecto para el registro y el depósito de obras en línea para propósitos de derecho de autor.

La página Web de la Biblioteca del Congreso también contiene el sitio **THOMAS.** Denominado así en honor a Thomas Jefferson, cuya biblioteca personal fue la que inició y durante mucho tiempo fue la colección más importante de la Biblioteca del Congreso, THOMAS ofrece información que se actualiza con regularidad sobre la legislación aprobada por el Congreso de Estados Unidos. Estén atentos a la legislación sobre derecho de autor y otras cuestiones relacionadas con la privacidad, obscenidad / pornografía (para sólo nombrar alguna) y la relación de las mismas con Internet.

El **Cyberspace Law Institute** [Instituto de Leyes del Ciberespacio](CLI) pretende "estudiar y ayudar a desarrollar las nuevas formas de legislar que exige el crecimiento de las redes de comunicación global y las comunidades en línea". Hay abundante información dirigida a personas que no sean abogados; busque el vínculo con **E-mail Course on Cyberspace Law for Non-Lawyers (Curso por correo electrónico sobre Leyes del ciberespacio para personas que no son abogados).** Este curso virtual ya habrá terminado cuando este material se publique, pero las lecciones seguramente estarán archivadas en el sitio Web del CLI. Esta bien preparada presentación se puede encontrar en:
http://www.cli.org

El **UVa Copyright Resources** de la Universidad de Virginia ofrece información general sobre derecho de autor; declaraciones sobre política referente a las cuestiones de derecho de autor, por ejemplo el derecho de propiedad sobre obras producidas por los profesores y los estudiantes de la Universidad de Virginia; y lineamientos para cartas de solicitud de permisos. Búsquelo en:
http://wwwllib.virginia.edu/dmmc/copyright/local.html

When Works Pass Into the Public Domain (Cuando las obras pasan a ser del dominio público) es un cuadro de trabajo muy útil confeccionado por Laura N. Gassaway, Directora de la Biblioteca de la Facultad de Derecho & Profesora de Derecho, Universidad de Carolina del Norte, Chapel Hill. El cuadro, que reproducimos en este capítulo, nos ayuda a

determinar la duración del derecho de autor de muchas obras en los casos en que los hechos están muy claros; es menos útil en situaciones más complejas.
http://www.smartbiz.com/sbs/arts/ipi5.htm

El cuadro lo obtuvimos mediante un vínculo con el sitio **U.S. Intellectual Property Information** (Información sobre la Propiedad Intelectual en Estados Unidos). Búsquelo en:
http://www.fplc.edu/tfield/order.htm

El **Legal Information Institute, Cornell Law School** (Instituto de Información Legal, Facultad de Derecho de la Universidad de Cornell) presenta copias de las decisiones más recientes del tribunal Supremo de Estados Unidos, el Código Civil de Estados Unidos (leyes promulgadas por el Congreso de Estados Unidos), decisiones históricas, y más. Es un poco leguleyesco o legalístico pero es un recurso muy completo y útil. Consulte el área **material organized by legal topic (material organizado por tema legal).** Después marque **intellectual property (propiedad intelectual).** Después **copyright (derecho de autor)** o **right of Publicity/Right of Privacy (Derecho de publicidad/derecho de privacidad).** Comience en:
http://www.law.cornell.edu/lii.table.html

La **Guide to Copyright Bibliography (Guía para la Bibliografía sobre el Derecho de Autor)** es una lista de fuentes muy útiles, en su mayoría artículos sobre leyes. Algunos están accesibles para los que no son abogados. Acceda a la lista y vea que le puede ser útil:
http://www.music.indiana.edu/tech_s/mla/legcom/bib.htm

La **International Federation of Library Associations (Federación Internacional de Asociaciones de Bibliotecas) (IFLA)** cuya página Web está en:
http://www.nlc-bnc.ca/ifla/ Home.htm

Contiene muchos términos útiles, incluida una bibliografía de artículos, revistas científicas, organizaciones, compañías y colecciones en línea de relevancia. Incluye también ejemplos de políticas de derecho de autor y tratados internacionales a los que se puede acceder a través del sitio de IFLA o directamente en:
http://www.nlc-bnc.ca/ifla/ll/cpyright.htm

El sitio de la **American Library Association (Asociación de Bibliotecas de Estados Unidos) (ALA)** ofrece información acerca de los últimos acontecimientos legislativos que afectan a las bibliotecas; el derecho de autor también continúa siendo uno de los aspectos de mayor interés. En su página principal, busque en las secciones, **Copyright and Intellectual**

Property (Derecho de Autor y Propiedad Intelectual) e **Intellectual Freedom (Libertad Intelectual)**. También tienen una lista de servicio -- obtenga instrucciones para inscribirse en la página Web en: **http://www.ala.org/alawashington.html**

La **Coalition for Networked Information (Coalición para la Información a través de Redes)** (CNI) tiene una lista de servicios en **Copyright and Intellectual Property Forum (Forum sobre Derecho de Autor y Propiedad Intelectual)** cuyo objetivo principal es "proporcionar a aquellos que preguntan, responden, y analizan las cuestiones sobre derecho de autor, cualquier tipo de foro para el debate...sin limitaciones en cuanto a las áreas. Por ejemplo el derecho de autor sobre los materiales electrónicos". El archivo se puede encontrar en: **ftp://ftp.cni.org/CNI/forums/cni-copyright.** O mejor aún, suscríbase enviando una nota por correo electrónico a Coalition ListProcessor (LISTPROC@CNI.ORG) con el siguiente mensaje:
subscribe CNI-COPYRIGHT <nombre de la persona>

La **Sexta Conferencia sobre Computadoras, Libertad y Privacidad celebrada en el Massachusetts Institute of Technology, 27-30 de marzo de 1996** abordó una amplia variedad de aspectos en el contexto de Internet entre los que se incluyen el derecho de autor, la privacidad y la libertad de expresión. Incluye vínculos con el contenido que abordaron distinguidos panelistas en los debates. Esto le ayudará a tener una idea más clara de la variedad de temas:
http://swissnet.ai.mit.edu/ switz/cfp96/index.html

El **Journal of Online Law** (JOL), "es una publicación electrónica de ensayos de académicos sobre leyes y comunicaciones en línea -- la Ley y el Ciberespacio". Estos magníficos ensayos son útiles tanto para abogados como para los que no lo son. Busque la revista en:
http://www.wm.edu/law/publications/jol

A License to Copy (Una licencia para copiar) describe una propuesta para establecer un esquema de licencias para las instituciones académicas basadas en la cantidad de estudiantes matriculados para utilizar cierto número de materiales protegidos por derecho de autor. El marco es similar a las prácticas de obtención de licencias que se aceptan en la industria de la música y que controlan ASCAP y BMI. Los proponentes afirman que reducirá el pesado trabajo de obtención y pago de permisos en el sistema de educación superior. Los oponentes afirman que destruye el uso justo:
http://www.copyright.com/chronicl.html

La **Electronic Frontier Foundation (Fundación de Fronteras Electrónicas)** es "una organización no lucrativa de libertades civiles que

trabaja en interés del público para proteger la privacidad, la libertad de expresión y el acceso a los recursos públicos y a la información en los nuevos medios de comunicación". Busque información legislativa y más en:
http://www.eff.org/

The **Digital Future Coalition (Coalición del Futuro Digital)** incluye el derecho de autor e información sobre otras cuestiones legales y sus relaciones con Internet. La DFC fue creada en el otoño de 1995 "para trabajar en pro de un debate profundo, amplio y balanceado en el Congreso de los Estados Unidos sobre la ley de derecho de autor y la política del mismo". Véase la lista de las organizaciones miembros, que vincula con otros sitios útiles:
http://www.dfc.org/

Un artículo interesante por David Post, Profesor Invitado de Derecho, Georgetown University Law Center (Centro de Derecho de la Universidad de Georgetown), es **The State of Nature and the First Internet War: Scientology, Its Critics, Anarchy and Law in Cyberspace (El estado de la naturaleza en la Primera Guerra de Internet: la Cienciología, La crítica, la Anarquía y la Ley en el Ciberespacio)** en:
http://www.reasonmag.com/reason/9604/Fe.POST.text.html

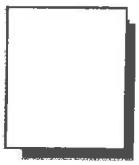
Otro artículo por Ann Okerson de Yale University, **Who Owns Digital Works (¿Quién es el propietario de las Obras digitales?),** está accesible y plantea muchos de los retos que presenta aplicar el derecho de autor a las actividades en línea. La versión impresa se encuentra en *Scientific American,* julio 1995, Pág. 80. En Internet:
http://www.sciam.com/WEB/0796issue/0796okerson.html

Bibliografía

Nicholas Negroponte. *Being Digital.* New York: Knopf, 1995. Una obra interesante para analizar los conceptos generales asociados a la información digital. Útil para los que no estén muy inclinados hacia la técnica.

A Museum Guide to Copyright and Trademark (American Association of Museums) por Michael S. Shapiro y Brett I. Miller, Morgan, Lewis & Bockius, LLP. 1999. Este libro fue publicado con el apoyo de financiamiento de The Pew Charitable Trusts, con apoyo adicional del The J. Paul Getty Trust.

Lance Rose. NetLaw: *Your Rights in the Online World,* 1st ed. New York: Osbourne/McGraw-Hill,1995. Una introducción a la aplicación de las leyes de Estados Unidos al mundo en línea. Palabrería legal mínima, útil para abogados y no abogados.



Buscar Contenidos Sitio del NEDCC

Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Modificado por última vez: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

VI

Las Primeras Cosas Técnicas

Steven Puglia
National Archives and Records Administration

Introducción

Este capítulo expone a los lectores a la terminología y los conceptos técnicos del proceso de digitalización. Específicamente, proporciona la información técnica básica relacionada con la digitalización de colecciones de bibliotecas, depósitos de archivos y otros materiales de instituciones culturales. Como panorama general, el capítulo no se adentra mucho en los detalles técnicos para realizar la digitalización en si misma. En realidad, está dirigido a aquellos que son responsables de la gestión de las actividades o del trabajo en otros aspectos del proyecto de digitalización. Para algunos lectores, este capítulo puede ser demasiado básico, para otros, puede resultar algo complejo, pero al menos trata de lograr un balance razonable que puede ser muy útil.

La imagen digital

La digitalización convierte una imagen en una serie de elementos pictóricos o pixels, pequeños cuadritos que son o blancos o negros (código binario), o de una tonalidad específica de gris (escala de los grises) o de color. Cada pixel se representa por un solo dígito binario o por una serie de dígitos binarios, que pueden ser 1s o 0s. Los pixels están dispuestos en una matriz bidimensional que se denomina mapa de bits.

A esto, a su vez, se le denomina trama o imagen de red. Si uno amplía una imagen digital de trama, podremos ver que la imagen está compuesta por una serie de filas y columnas de pixels cuadrados. Una imagen de trama es relativamente análoga a las fotografías tradicionales, que están compuestas por granos o partículas de plata o de colorante que son las que forman la imagen. Siempre que sea posible, este capítulo hará una comparación entre aspectos de la tecnología digital y la fotografía tradicional como punto de referencia.

Los ficheros de imágenes de vectores son otro tipo de imagen computarizada diferentes de la imagen de trama. Muchos programas de computación (dibujos, ilustraciones, modelos representados en 3D/diseño por computadora / fabricación con ayuda de computadoras, y diseño arquitectónico) utilizan vectores – flechas de dirección, puntos y líneas – que definen formas, en comparación con los elementos pictóricos que se utilizan para representar una trama. Este capítulo no volverá a analizar este aspecto.

El proceso de digitalización

La digitalización es el proceso de convertir una señal analógica en señal digital, que se conoce como conversión A/D (analógico a digital). En el caso de las imágenes de trama, se divide una señal de voltaje analógico (de cualquiera entre diversos tipos de sensores de imágenes), proporcional a la cantidad de luz reflejada o transmitida por un objeto que se digitaliza, entre valores numéricos discretos. El número del valor es la profundidad de bit para cada pixel.

Sensores comunes de Digitalización de Imágenes (detectores de imágenes)

Una parte importante del proceso de digitalización es el tipo de sensor utilizado. Estos detectores de imágenes se pueden comparar con la película utilizada en fotografía. Los sensores comunes de digitalización de imágenes son:

- CCD: Dispositivos acoplados cargados (Charged coupled devices), o CCDs, que se utilizan en escáners planos y en cámaras digitales
- PMT: Tubos foto multiplicadores (Photo-multiplier tubes), o PMTs, utilizados en escáners de cilindro
- CMOS: Semiconductores de óxido de metal complementarios (Complementary metal oxide semiconductors), o chips CMOS, utilizados en escáners planos baratos y en cámaras digitales también baratas.

El sensor más común que se utiliza en los escáners es el de dispositivo acoplado cargado o CCD, utilizado en todos los tipos de cámaras digitales y escáners. El tubo foto multiplicador, o PMT, se utiliza sólo en los escáners de cilindro para artes gráficas o para la pre-impresión, es decir, para el mercado de publicaciones y la imprenta. Más recientemente, se han introducido los semiconductores de óxido de metal complementarios o sensores CMOS como una alternativa de bajo costo con respecto a los CCDs. Los chips CMOS se fabrican en la misma forma que los chips estándar de computadoras, y su fabricación es, por tanto, menos costosa. Eventualmente, los chips CMOS podrían reemplazar a los CCD como sensor predominante en el mercado, pero en la actualidad, debido a ciertas deficiencias técnicas, no producen la misma calidad de imagen que los CCDs y no pueden igualar la resolución de los CCDs. En este momento, los chips CMOS se utilizan sólo en cámaras digitales y escáners.

Los CCDs se producen en gran variedad de diseños y formas. Una sola hilera de sensores CCD (o foto diodos) dispuestos en una línea recta se denomina disposición en línea. Las disposiciones en línea se utilizan por lo general en los escáners planos. Las disposiciones de área abarcan un conjunto bidimensional de fotosensores dispuestos en filas y columnas. Las disposiciones de área se usan por lo general en las cámaras digitales.

Tanto en los escáners como en las cámaras digitales, un lente o juego de lentes se utilizan para enfocar una imagen en el sensor -- un CCD, un PMT, o un chip CMOS. Algunas veces las personas se refieren a la digitalización de imágenes como una actividad que se realiza sin lentes. Esto no es cierto. Sin un lente con el que enfocar la luz, las imágenes digitales serían borrosas, al igual que en la fotografía. Todos los escáners tienen lentes, lo que, en la mayoría de los casos, el diseño del escáner esconde el lente (o los lentes) dentro del cuerpo del propio escáner.

Los escáners también tienen fuentes de luz incorporadas para iluminar los objetos que se escanean. La luz, o bien se refleja (al igual que con los documentos o impresiones fotográficas) o se transmiten (como con el microfilme, los negativos fotográficos o las transparencias a color) por el objeto que se escanea, y la imagen se enfoca por el lente(s) sobre el sensor que copia la imagen. En el caso de los CCDs, la luz cae sobre los pequeños sensores de luz o foto diodos que están en el CCD. Estos sitios o diodos generan una corriente eléctrica, o voltaje. La cantidad de voltaje generado es proporcional a la cantidad de luz que incide sobre el sensor individual. Mientras más brillante sea la luz, mayor el voltaje que genera el sitio en el CCD.

Conversión analógica a digital

La señal analógica digital que genera el sensor se procesa por la vía de la conversión analógica a digital. La señal eléctrica generada en cada sensor de luz o foto diodo se divide en valores numéricos discretos que son proporcionales a la cantidad de luz reflejada por el objeto que se escanea o que éste transmite. El número total de valores numéricos discretos posibles está determinado por el muestreo de la profundidad de bit, mientras que el valor numérico específico para un pixel individual se basa en la cantidad de luz específica que se refleja o se transmite desde ese punto sobre la imagen original.

Además, una vez que una imagen digital ha sido creada y almacenada en cualquier soporte, existe una conversión digital a analógica correspondiente que permite a la computadora presentar la imagen en forma leíble para el ojo humano ya sea en un display o en una impresora. Presentar una imagen en el monitor de una computadora o imprimir la imagen en una impresora son ejemplos de representación analógica de una imagen digital.

Dimensiones básicas de la imagen

Existen tres dimensiones importantes de toda imagen digital estática:

- Resolución. El número de puntos, o pixels (elemento pictórico), utilizado para representar la imagen. Ésta siempre se da como una medida de densidad lineal o de área (*ej.*, 300 puntos / pulgada).
- Profundidad Pixel Bit. Esta medida define la cantidad de sombras que se pueden representar en realidad por la cantidad de información salvada por cada pixel. Puede variad desde 1 bit/pixel para las imágenes binarias (tipo fax) hasta 24 bits por pixel en imágenes a color de alta calidad.
- Color. Existen muchas maneras de representar, compactar y distribuir las imágenes a color. Baste decir que mientras más pequeño sea el tamaño del fichero de la imagen, menos precisa será la imagen original entregada.

1. Resolución

La resolución, o frecuencia espacial, es el número de veces que se muestrea una imagen durante el proceso de escaneo. La resolución – o número de pixels en una imagen – se puede describir de varias formas:

DPI— puntos por pulgada
PPI-- pixels por pulgada
LPI— líneas por pulgada, utilizada para los medios tonos

La resolución de escaneo y la resolución de los ficheros de la imagen digital se refieren mejor como pixels por pulgada o PPI. El término puntos por pulgada o DPI se considera un término de impresión y es más adecuado cuando nos referimos a la resolución a la que la impresora de la computadora produce una impresión. No obstante, DPI es un término más genérico y se usa más comúnmente que el término PPI. LPI es un término que se refiere a los valores de medios tonos de la pantalla. Una pantalla de medio tono convierte una imagen en una serie de puntos que se pueden reproducir en la impresora de una computadora o en una imprenta; las imágenes digitales de tono continuo se convierten a medios tonos cuando se imprimen en la mayoría de las impresoras de computadoras, incluidas las de chorro de tinta y las impresoras láser. Algunas impresoras imprimen imágenes de tono continuo y no convierten la imagen.

1.a. La disposición en Pixels

La disposición en pixels es el número de pixels que hay en ambas dimensiones de una imagen en términos de hileras y columnas a través de las dimensiones de la imagen. Por ejemplo, una fotografía de 8" x 10" se escanea a 300 ppi. Esto produce un fichero que tiene una disposición de pixels de 2400 x 3000 pixels.

Por lo general, líneas por pulgada (LPI) es un término que se utiliza para los medios tonos (para la reproducción en una imprenta) y no se utiliza para imágenes de tonos continuos. Sin embargo, "líneas" o hileras de pixels es un término que se utiliza en la industria fotográfica como abreviatura común para el número de pixels a lo largo de la dimensión longitudinal de las imágenes de las fotografías. Como las fotografías pueden tener muchos formatos y dimensiones diferentes (que varían desde pequeños negativos hasta grandes impresiones), es difícil referirse a los pixels por pulgada (PPI) de la resolución cuando producimos imágenes digitales del mismo tamaño, porque los PPI variarán en dependencia del tamaño del original fotográfico.

Una imagen impresa de 8" x 10" escaneada a 300 ppi produce un fichero de 2400 x 3000 pixels.
Un negativo de 4" x 5" escaneado a 600 ppi produce un fichero de 2400 x 3000 pixels.
Un negativo de 35mm escaneado a 2100 ppi produce un fichero de 2000 x 3000 pixels.

Nos referimos a cada uno de estos ficheros como un fichero de 3000 líneas, y todas las dimensiones mencionadas son anteriores a la aplicación de cualquier tipo de compactación.

1.b. La resolución – La verdadera vs. La interpolada

... obtenga un escáner con la mayor resolución óptica que Ud. pueda adquirir

La resolución óptica (verdadera) es la resolución inherente del escáner basada en el tamaño del sensor y de la ampliación que pueda dar el sistema óptico. La resolución interpolada es resolución sintética o calculada. Interpolación es un proceso matemático que se utiliza para incrementar o disminuir la resolución de una imagen. Esto se puede hacer durante el escaneo o después del escaneo. Una mayor resolución óptica (verdadera) en un escáner proporcionará una mejor calidad de imagen que la resolución interpolada. Se recomienda que obtenga un escáner con la mayor resolución óptica que Ud. pueda adquirir, no sólo en términos del precio del escáner, sino también teniendo en cuenta el costo del almacenamiento de cada fichero de imagen.

Pruebe escanear los documentos con el escáner que va a utilizar en realidad

La interpolación puede ser tan simple como cambiar la resolución óptica a un valor menor para propósitos de presentación, o tan compleja como detectar y volver a presentar las áreas de medios tonos de un documento. Algunos algoritmos de interpolación funcionan mejor que los otros. Por lo general, los softwares de procesamiento de imágenes que son más costosos funcionan mejor que otros. No obstante, hay excepciones porque cada vez más algoritmos de interpolación se incorporan en el hardware del escáner. La mayoría de los algoritmos de interpolación representan un intermedio entre la calidad de la imagen, la velocidad y el tamaño del fichero de la imagen. Se recomienda con mucha fuerza que se haga la prueba de los documentos con los escáners que se van a utilizar en realidad y con los algoritmos de mejoramiento de la imagen que se van a utilizar. También es buena idea adquirir un escáner más costoso o añadir un paquete de software de procesamiento – uno que tenga mejores algoritmos de procesamiento de imágenes. A pesar de que la mayoría de los usuarios utilizan sólo un pequeño número de todas las características disponibles, en un momento o en otro será necesario procesar imágenes de alta calidad con algoritmos de procesamiento de imágenes.

1.c. Resolución de digitalización

La resolución de digitalización se puede dividir en dos categorías genéricas. La resolución de reproducción es la resolución necesaria para proporcionar la calidad de imagen deseada para un tipo específico de dispositivo de salida. La resolución de preservación es el nivel de resolución que reproduce toda la información en la imagen original o documento. Para utilizar fotografías como ejemplo, estos niveles podrían ser:

- Reproducción— resolución de pantalla, que es un mínimo de 600 x 400 pixels, o resolución de impresión, que es usualmente 300 dpi a 600 dpi.
- Preservación— hacer coincidir el original (ejemplos para negativos a color y transparencia a color). Este es un límite de resolución teórico basado en la granularidad de la película original y la resolución del lente utilizado para tomar la fotografía. La verdadera resolución digital deseada variará en dependencia de la película fotográfica (y del revelador utilizado), del lente de la cámara original, el tamaño característico significativo que se desea reproducir y la calidad del escáner utilizado para digitalizar.
- 3,000 a 4,000 líneas para 35mm
- 10,000 a 16,000 líneas para 4" x 5"
- 20,000 a 32,000 líneas para 8" x 10"

A continuación aparecen estimados de tamaños de ficheros para calidad de preservación en fotografías escaneadas --negativos o transparencias – basados en los límites menores de resolución citados arriba.

1.d. Requisitos generales o mínimos de digitalización para facilitar la reproducción y el acceso

Cornell recomienda 600 ppi para escaneado de 1-bit o 400 ppi para escaneado de 8-bit de tipos impresos para obtener un escaneado con calidad de preservación. Otras recomendaciones generales para la reproducción son:

Registros de texto
200 a 600 ppi para 1-bit
200 a 400 ppi para 8-bit escala de grises
200 a 300 ppi para 24-bit color
Fotografías
3000 a 5000 líneas para 8-bit escala de grises
3000 a 5000 líneas para 24-bit color
Mapas/Planos/Gran tamaño
200 a 300 ppi para 8-bit escala de grises
200 a 300 ppi para 24-bit color

Según las computadoras se hacen más rápidas y las memorias se hacen más baratas, es probable que aumenten las recomendaciones para la resolución de escaneado. En la actualidad, los proyectos seleccionan una mayor resolución de escaneado que los primeros proyectos de digitalización.

2. Profundidad Pixel Bit

Las computadoras trabajan según un sistema binario; cada bit de datos es o un 1 o un 0. Cada pixel en una imagen de trama se representa por una cadena de dígitos binarios. Un pixel de 1-bit se representa por un dígito binario, o bien un 1 o un 0. Un pixel de 2-bit se representa por dos dígitos binarios, o bien – 0+0, 0 + 1, 1 + 0, o 1 + 1. La profundidad de bit determina el número de combinaciones posibles de 1s y 0s para ese número de dígitos binarios y por tanto el número de tonalidades de grises o tonalidades de color que se pueden representar por cada pixel, como se ilustra en la siguiente fórmula.

Cantidad de tonalidades = 2^X X = la profundidad de bit

1 bit = 2 tonos (un solo dígito binario- un solo 1 o un solo 0 – (negro o blanco)
2 bits = 4 tonos (dos dígitos binarios forman cuatro combinaciones posibles -- negro, gris oscuro, gris claro, y blanco)
3 bits = 8 tonos (tres dígitos binarios forman 8 combinaciones posibles)
4 bits = 16 tonos (cuatro dígitos binarios forman 16 combinaciones posibles)
5 bits = 32 tonos (cinco dígitos binarios forman 32 combinaciones posibles)
6 bits = 64 tonos (seis dígitos binarios forman 64 combinaciones posibles)
7 bits = 128 tonos (siete dígitos binarios forman 128 combinaciones posibles)
8 bits = 256 tonos (ocho dígitos binarios forman 256 combinaciones posibles
10 bits = 1,024 tonos (diez dígitos binarios forman 1,024 combinaciones posibles)
12 bits = 4,096 tonos (doce dígitos binarios forman 4,096 combinaciones posibles)
14 bits = 16,384 tonos (catorce dígitos binarios forman 16,384 combinaciones posibles)
16 bits = 65,536 tonos (dieciséis dígitos binarios forman 65,536 combinaciones posibles)

Profundidad de Bit de ilustraciones

A continuación aparecen las normas actuales de profundidad de bit para ficheros de imágenes.

1-bit	blanco y negro
8-bit escala de grises	256 tonos de grises
8-bit color	256 colores
24-bit RGB*	aproximadamente 17 millones de colores, tres canales de 8-bit

**Véase la siguiente sección*

La profundidad de bits influye en la representación de las imágenes. Es evidente que a 1-bit sólo hay valores en blanco y negro y no hay tonalidades de grises. La textura y otros valores de tonalidades sutiles no se reproducen. A 2-bits, se reproducen cuatro tonalidades – blanco, negro y dos tonalidades intermedias de gris. A 4-bits, se reproducen 16 tonalidades, y se presenta la textura de fondo de un documento. A 6-bit escala de grises, 64 tonalidades de grises, la imagen digital se aproxima a la respuesta perceptual típica del ser humano. Los estudios psicométricos han determinado que la mayoría de las personas puede diferenciar aproximadamente 64 tonalidades de grises. Hace años, cuando los científicos de la computación estaban estableciendo convenciones para la digitalización de imágenes, la memoria de las computadoras era costosa y la velocidad de la CPU era lenta. Era fácil tomar la decisión de limitar los ficheros de imágenes de escalas de grises a 8-bits para ahorrar espacio de almacenamiento, ya que las 256 tonalidades reproducidas exceden la percepción humana. Sin embargo, la representación de 256 tonalidades en escala de grises de 8-bit es limitada en comparación con los materiales fotográficos y pueden presentar problemas cuando es necesario ajustar el contraste y la brillantez de las imágenes digitales. El uso de ficheros de imágenes de escala de grises a 8-bit, y los correspondientes ficheros de imágenes a color a 24-bit RGB (tres canales de información a color de 8-bit), fue un compromiso razonable y, en muchos casos todavía lo es.

3. Color

Sistemas de Color

Se utilizan varios sistemas diferentes para representar imágenes a color. Los más comunes son el RGB (sistema de color aditivo), el CMYK (sistema de color sustractivo), y el CIE-L*A*B* color space, una modelación matemática del color.

RGB El sistema de color aditivo combina variaciones de rojo, verde y azul (RedGreenBlue) para formar el blanco. Este método se utiliza en el diseño de televisores, monitores de computadoras y grabadoras de películas. Piense en una imagen RGB a color como tres imágenes superpuestas una sobre la otra. La superposición se realiza de forma matemática. En principio, una imagen RGB consta de tres imágenes de

8-bit en escala de grises o canales; un canal representa la información en rojo, otro canal representa la información en verde y el tercer canal representa la información en azul. La computadora combina por medios matemáticos los tres canales para cada pixel para determinar el color final. Un fichero RGB de imagen digital a color de 24-bit consta de tres canales, cada uno con 8-bits de datos (3 canales x 8-bits = 24-bits).

CMYK El sistema de color sustractivo combina variaciones de cyan, magenta, y amarillo para formar negro. Este método se utiliza en el proceso de impresión en las artes gráficas y con las impresoras de computadoras. Con frecuencia, la impresora utiliza un cuarto color de tina, el negro, para incrementar el rango de densidades que se puede reproducir. Las impresoras a cuatro colores utilizan cyan, magenta, amarillo y negro (CyanMagentaYellowblacK). Casi todos los materiales fotográficos a color se basan en color sustractivo, utilizando diversas cantidades de pigmentos cyan, magenta, y amarillo. La mayoría de las impresoras utilizan cuatro colores, aunque ahora existen en el mercado impresoras de seis colores; se podría añadir un magenta claro y un cyan claro para mejorar la calidad de la imagen cuando se imprimen imágenes fotográficas. Un fichero CMYK de imagen digital a color de 32-bit consta de cuatro canales, cada uno con 8-bits de datos (4 canales x 8-bits = 32-bits).

*CIE-L*A*B** El CIE-L*A*B* color space es un modelo matemático de color que divide el color en luminosidad (L) que pudiera considerarse como información de la escala de los grises, información en rojo (+A) a verde (-A), y azul (+B) a amarillo (-B). El L*A*B* color space se considera como un espacio de color independiente de dispositivos. No está vinculado a ningún tipo específico de dispositivo de salida como un monitor de computadora (RGB) o una impresora (CMYK). Un fichero de imagen digital L*A*B* de 24-bit a color consta de tres canales, cada uno con 8-bits (3 canales x 8-bits = 24-bits).

Recomendaciones

Para la mayoría de los proyectos, salve las imágenes de color como ficheros RGB.

La mayoría de los escáners utilizan escaneado en RGB aunque algunos convierten las imágenes a CMYK o imágenes L*A*B*. Para la mayoría de los proyectos de digitalización, se recomienda salvar las imágenes a color como ficheros RGB, y no como ficheros CMYK. El L*A*B* color space podría utilizarse, pero muy pocas aplicaciones de software son

capaces de interpretar y utilizar los ficheros L*A*B* en esta época. Como siempre es posible convertir los ficheros RGB para la salida, sólo deben utilizarse los ficheros CMYK para imprimir. El objetivo primordial en la preservación es salvar la mayor cantidad de información que sea económicamente posible, utilizando métodos que puedan ser reversibles si fuera necesario.

Cada vez más, los escáners y los softwares son capaces de manipular ficheros de imágenes de altos bit. Esto quiere decir que, en lugar de tener 8-bits por cada canal de color, el fichero puede tener 10-bits, 12-bits o 16-bits por cada canal de color. Una imagen RGB a color que cuenta con 16-bits por canal es un fichero de imagen a color de 48-bit (3 canales x 16-bits = 48-bits).

Comparación

Como se puede apreciar en el cuadro, según se aumenta la profundidad de bit, se incrementa de forma drástica la cantidad de tonalidades y de colores que se pueden reproducir. Los materiales fotográficos son capaces de presentar de manera efectiva varios miles de tonalidades. La profundidad de bit equivalente para la digitalización de imágenes es de al menos 12-bits por canal.

Gama de colores

Una gama de colores es el rango de colores que puede reproducir un sistema, por ejemplo un monitor o una impresora de computadora. Las gamas de colores se ilustran de forma gráfica para comparar diferentes espacios de colores, sistemas de colores o dispositivos. Los espacios de colores amplios de RGB y L*A*B* pueden brindar un rango mayor de colores y por lo general requieren del uso de mayores profundidades de bit para lograr una gama de colores amplia. El sistema de color CMYK tiene una gama de colores limitada y puede reproducir, por consiguiente, un rango de colores limitado; esta es otra razón para no utilizar ficheros CMYK para ficheros de imágenes masters.

Paletas de colores

Las paletas de colores son conjuntos discretos de colores definidos que usan las computadoras para representar imágenes de 8-bit o 256-colores. Los sistemas operativos Windows y Macintosh utilizan diferentes conjuntos de colores para las imágenes a color de 8-bit. La entrega de la imagen cambia en dependencia del tipo de computadora en que se vea la imagen. Un enfoque para los formatos a color de 8-bit – como los ficheros GIF que son para distribuir vía World Wide Web – es utilizar una paleta con seguridad Web. Una paleta con seguridad utiliza 212 a 216 colores que son comunes a las paletas de Windows y de Macintosh, y la imagen debe verse igual en cualquiera de los dos tipos de computadora. Otra opción es una paleta adaptable, en la que los 256 colores utilizados para la paleta se basan en los colores específicos en una imagen específica. En la mayoría de los casos, el uso de una paleta adaptable hará que una imagen a color de 8-bit se vea muy parecida a la imagen original de 24-bit, en comparación con el uso de una paleta Windows, Macintosh, o con seguridad Web.

Desbalance del color

Los desbalances de color ocurren cuando no se brindan valores neutrales con los mismos niveles de rojo, verde y azul (por supuesto, para un fichero de imagen RGB). Por ejemplo, una zona blanca en una imagen digital cambiará a un color en el que los tonos se unen en uno solo o en dos de los canales de colores.

Precisión del color

El manejo de la precisión del color para las imágenes digitales es complejo, e implica el ajuste y la calibración de los monitores, el ajuste de los controles del escáner, la corrección o mejoramiento de las imágenes utilizando softwares de procesamiento de imágenes, el ajuste y la calibración de los dispositivos de salida y el uso de softwares de manejo de colores. Este software transforma las imágenes entre los diferentes espacios de color para corregir las diferencias de las gamas de colores de los escáners, los monitores y los dispositivos de salida. Apple's Colorsync y Windows ICM son ejemplos de softwares de manejo de color que han sido incorporados en los sistemas operativos de computadoras.

La medición de los valores digitales

No todos los escáners o controles de procesamiento de imágenes funcionan tan bien como se espera que funcionen, por tanto, con frecuencia es necesario medir los valores digitales – ya sea los niveles

RGB para imágenes a color o el % de negro para las imágenes en escala de grises. La mayoría de las aplicaciones de softwares de procesamiento de imágenes poseen un control que permita al usuario medir los valores digitales para un solo pixel o para un grupo de pixels en una imagen, como el Eyedropper en Adobe Photoshop. Es importante ajustar las opciones para el Eyedropper en los valores o controles adecuados antes de medir los valores. Todas las imágenes digitales tienen ruido (pixels aleatorios de una tonalidad o color incorrectos) que hacen que la medición de pixels individuales sea problemática. Se recomienda ajustar el Eyedropper en los valores que promedian un conjunto de 5 x 5 pixels (un cuadrado de 25 pixels). Esto promediará la variación debido al ruido.

Procesamiento de la imagen digital

Sobremuestreo

Como se señaló anteriormente, las imágenes digitales tienen profundidades de bit de 1-bit por pixel para las imágenes en blanco y negro (común para la digitalización de documentos), 8-bits por pixel en escala de grises para las imágenes de tono continuo, y 24-bits por pixel para las imágenes a color. Por lo general, los escáners hacen un muestreo a profundidades de bits mayores que estas, y después la profundidad de bit se reduce para la imagen final. Esto se conoce como sobremuestreo. Los escáners están diseñados para sobremuestrear con el objetivo de mejorar la calidad de la imagen mediante la reducción del ruido (pixel aleatorio en la tonalidad incorrecta), y de extender la escala tonal efectiva del escáner (midiendo de inicio más tonalidades que las que se usan en la imagen final). Esto permite representar un rango de densidad mayor sin pérdida de los detalles – un problema que se presenta cuando se escanea diapositivas a color o transparencias a color y otros originales muy densos. Los escáners de documentos harán un muestreo a 8-bits para producir una imagen de 1-bit, y un escáner en escala de grises hará un muestreo a 10-bits o 12-bits para producir una imagen de 8-bit.

Filtros de procesamiento de imágenes

Los filtros de procesamiento de imágenes – fórmulas matemáticas que cambian la apariencia de las imágenes digitales – se pueden aplicar para mejorar la apariencia de las imágenes y para ayudar a redimensionar las imágenes. Es común que se utilicen los filtros de precisión para mejorar la apariencia de los ficheros de imágenes digitales. La necesidad de precisar es directamente proporcional a la resolución de la imagen digital: las resoluciones más bajas de las imágenes digitales tienden a

necesitar más precisado, y las resoluciones más altas o las imágenes digitales más grandes tienden a necesitar menos precisado. Muchas personas abogan por que no se precisen los ficheros de imágenes, pues consideran que el mejoramiento no se puede deshacer en el futuro. El filtro de precisado más fotográfico es una máscara no precisa. Este término proviene de la práctica en la industria de las artes gráficas de utilizar una máscara de tono invertido que está ligeramente fuera de foco para aumentar ligeramente la precisión de las imágenes. Es posible precisar excesivamente una imagen: El precisado excesivo con un filtro de máscara no precisa creará un ligero halo alrededor de los bordes dentro de las imágenes.

Otro filtro que se utiliza comúnmente cuando se redimensionan las imágenes es el filtro borroso. Hacer que una imagen sea ligeramente borrosa crea una tonalidad adicional a lo largo de los bordes bien definidos en una imagen, lo que puede permitir al software de interpolación realizar un mejor trabajo cuando la imagen se redimensiona. La mayoría de las imágenes deben volverse a mejorar después del redimensionamiento, se aplique o no un filtro de borrado.

Al igual que con los algoritmos de interpolación, algunos algoritmos filtros de procesamiento de imágenes serán mejores en cuanto a la calidad de la imagen obtenida que otros algoritmos, mientras que otros trabajan más rápidamente. Es necesario repetir que, por lo general, los filtros en los softwares de procesamiento de imágenes que son más costosos tenderán a realizar un trabajo de más calidad con la imagen si se comparan con otros filtros en los softwares menos expandidos.

Histograma

Una herramienta común de procesamiento de imágenes es el histograma, que se encuentra en la mayoría de los paquetes de softwares de procesamiento de imágenes. El histograma es una representación gráfica de la distribución de las tonalidades de grises en una imagen. La altura de cada línea vertical es proporcional al número de pixels que existen de esa tonalidad – mientras más larga sea la línea, más pixels de esa tonalidad habrá. Además, el histograma puede ofrecernos indicaciones de ciertos tipos de defectos de las imágenes, como por ejemplo, pérdida de tonalidades en las sombras (los valores oscuros de las tonalidades) o en las zonas claras (los valores de luz o tonos) de una imagen. El histograma ilustra y nos ayuda a comprender el concepto de umbral.

Umbral

El umbral es una técnica utilizada en el procesamiento de imágenes para convertir los tonos grises a negro o blanco. Todas las tonalidades por

debajo de un valor seleccionado se presentan como blanco y todas las tonalidades por encima de ese valor se presentan como negro. En dependencia del valor seleccionado para el umbral, la representación de la misma imagen se puede alterar de forma drástica. La mayoría de los escáners de 1-bit en realidad hacen un muestreo a 8 bits, pero después se usa un valor de umbral para convertir la imagen de 8-bit a una imagen de 1-bit.

¿Cuándo se presentan problemas en la utilización de digitalización a 1-bit y con el umbral? En los casos de procesos como thermofax, verifax, o copias al carbón, en los que el papel envejece según se oscurece y los tipos se desvanecen, es muy difícil reproducir la imagen con escaneo a 1-bit independientemente del nivel de umbral. A valores de umbral menores, los caracteres aparecen incompletos. Según aumenta el valor del umbral, los caracteres se rellenarán rápidamente (*ej.*, la letra "o" se convierte en un punto muy grande) y sólo el contexto dentro de la palabra o la oración nos proporciona una idea de cuál es el caracter. Si se aumenta aún más el nivel del umbral provocará que los pixels que representan las sombras en el fondo se vuelvan negros, efecto que se conoce con el nombre de moteado o punteado. Existen softwares diseñados para trabajar con escaneado a diseñados para desmotear una imagen. El software trata de eliminar los pixels negros extraños a la imagen. Desafortunadamente, esto no siempre funciona de la manera que uno quiere. Los parámetros para desmotear se pueden ajustar, basándonos en el tamaño de la mancha que uno quiere eliminar, pero, según aumenta el tamaño de la mancha a eliminar, el programa también comenzará a eliminar los puntos, los puntos de las íes y otros signos de puntuación necesarios.

Imprecisión

Cuando se usan imágenes a bajas profundidades de bit, es posible simular un mayor número de tonalidades con menor cantidad de tonalidades. Este proceso se conoce como imprecisión. La clave es redistribuir los pixels según una fórmula matemática para producir tonalidades de gris sintéticas basadas en la disposición de esos pixels y en la forma en que el ojo las percibe. Existen diferentes fórmulas para la imprecisión, y algunas funcionan mejor que otras. Si una imagen en escala de grises de 8-bit se convierte a imagen de 3-bit sin el proceso de imprecisión, se presentarán amplias áreas de tonalidades similares como si fuera una sola tonalidad. En términos fotográficos, este efecto se conoce como posterización. En imágenes digitales, algunas veces nos referimos a ese efecto como agrupación o apiñamiento, sobre todo cuando aparece en amplios gradientes de tonalidades, por ejemplo el cielo en las fotografías. Cuando una imagen a color de 24-bit se convierte a una imagen a color

de 8-bit, el fichero de 8-bit se puede imprecisar. Se puede usar el imprecisado y una paleta de escala de grises adaptable para proporcionar una presentación muy precisa de una imagen con profundidades de bit tan bajas como 4-bits o 16 tonalidades.

Controles tonales

Cada aplicación de software de procesamiento de imágenes tiene diferentes controles para ajustar los tonos y el balance del color de las imágenes digitales. En Adobe PhotoShop, uno de los softwares de procesamiento de imágenes más populares, los controles preferidos para el ajuste tonal y la corrección del color son *Levels and Curves (Niveles y Curvas).* También se dispone de otros controles en PhotoShop, como *Contrast and Brightness (Contraste y Brillantez)* y *Color Balance (Balance del color),* pero estas son correcciones globales que influyen en todos los tonos de la imagen. *Levels and Curves* ofrece mayor control con menos riesgo de perder información a la vez que ajusta las imágenes.

Comparaciones de escalas tonales

La profundidad de bit de una imagen digital ejerce una gran influencia en la precisión con que se presenta un documento original, fotografía o libro en cuanto a los tonos del original. La reproducción del tono para una imagen digital de 1-bit es en cierto modo similar a la respuesta tonal del microfilm – alto contraste y más adecuado para presentar tipos impresos limpios. Una imagen en escala de grises de 8-bit es más similar a la película fotográfica en blanco y negro de tono continuo que se utiliza en la fotografía fija – contraste más bajo y capaz de presentar mayor rango de tonalidades.

Una forma de comparar las respuestas tonales es observar las curvas características para la película fotográfica y un gráfico similar para la respuesta de la digitalización. Una curva característica es una representación gráfica de la respuesta de una película a la exposición y al revelado. El eje horizontal es la exposición; según nos movemos hacia la derecha en el eje de las x, aumenta la exposición. El eje vertical es la densidad. Una curva característica para el microfilm muestra que según aumenta la exposición, existe un gran incremento proporcional de la densidad. La tasa de incremento en la densidad comparada con la exposición, que es la parte curva de la línea, es el contraste. El microfilme es una película fotográfica de alto contraste con un rango de tonos limitado que se puede distinguir. Esto es en cierto modo similar a las imágenes digitales de 1-bit (bitonal). En las imágenes bitonales, todos los valores bitonales del original que son más claros que una tonalidad seleccionada se presentarán como blancos, y todos los valores tonales más oscuros que el tono seleccionado se presentarán como negros. El

punto en el que las tonalidades cambian de blanco a negro es el valor del umbral. Todo lo que sea más claro que el valor del umbral se presentará como blanco, todo lo que sea más oscuro que el umbral se presentará como negro. La respuesta de digitalización de 1-bit es similar al microfilm. No obstante, a pesar de ser de alto contraste, el microfilme tiene un rango de tonalidades a diferencia de las imágenes de 1-bit, que sólo tienen valores en blanco o negro (todas las tonalidades grises se eliminan).

Si uno observa la curva característica de una película fotográfica de fotografía fija en blanco y negro de tono continuo, la curva parece diferente porque el contraste es menor y, para la mayoría de estas películas, el largo de la curva característica es mayor. Ambas propiedades significan que la película fotográfica de foto fija puede presentar más cantidad de tonalidades, equivalente a escanear a mayor profundidad de bit, y por consiguiente se pueden distinguir más tonalidades. Una imagen de escala de grises de 8-bit tiene una respuesta relativamente similar a las películas fotográficas de tono continuo. No obstante, una imagen de 8-bit tiene un máximo de sólo 256 tonalidades o niveles. La mayoría de las películas fotográficas pueden distinguir de manera efectiva miles de tonalidades.

Recortado

El recortado se produce cuando el detalle de la imagen se presenta como blanco o negro y por lo tanto el detalle se pierde. Una vez que los tonos han sido recortados, no es posible volver a obtener las tonalidades. Es importante ajustar los controles del escáner para reducir al mínimo el recortado durante el escaneo. Después, es importante evitar el recortado cuando se utilizan los controles de ajuste de tono y de color del software de procesamiento de la imagen.

Estructura del fichero de la imagen digital

Los datos digitales que representan una imagen completa están contenidos dentro de un fichero de computadora. La cadena de dígitos binarios está colocada en una estructura organizada que permite a la computadora y al software interpretar los datos y recrear la imagen. Un fichero de imagen digital tiene varias partes principales.

Estructura de fichero de imagen simple (tomado de *"Structures and Metrics for Image Storage and Interchange,"* JEI, Journal of Electronic Imaging *co-published por SPIE [the International Society of Optical Engineering] and IS&T [the Society for Imaging Science and Technology], abril 1993).*

Encabezamiento

* Un identificador del fichero
* Especificación de la imagen

Datos de la imagen
* Tabla de búsqueda
* Retícula o cuadrícula de imagen

Pie
* Conclusión del fichero

El encabezamiento del fichero identifica un fichero digital para la computadora e incluye una especificación de la imagen que indica el formato del fichero. La sección de datos de la imagen del fichero, que en algunos casos incluye una tabla de búsqueda, aparece a continuación del encabezamiento. La tabla de búsqueda es un conjunto definido de colores o tonalidades de gris que le dicen a la computadora cómo representar la imagen en un monitor de computadora. La cuadrícula o retícula de la imagen es la cadena de 1s y 0s que representa a cada uno de los pixels individuales que a su vez representan la imagen de mapa de bit. La parte final del fichero, el pie del fichero, le dice a la computadora que se ha abierto o descargado el fichero completo.

Compactación de datos y de ficheros

La compactación de datos y ficheros es el proceso de reducir, por diversos medios, la cantidad de datos que se pueden almacenar o transmitir. Existen dos grandes categorías de compactación: la compactación **sin pérdida** permite la reconstrucción de un fichero que es idéntico al original y la compactación **con pérdida** descarta cierta cantidad de información original durante el proceso de compactación. Entre algunos de los algoritmos de compactación incluyen:

LZW (Lempel-Ziv-Welch) – sin pérdida
JPEG (Joint Photographic Experts Group) – con pérdida
MPEG (Moving Pictures Experts Group) -- con pérdida
Wavelet – con pérdida
Fractal -- con pérdida

Comparación de reconversión de formato

Los documentos originales con tipos impresos definidos, o textos (con un alto contraste inherente entre el tipo / texto y el fondo, y caracteres altamente definidos) se pueden reproducir utilizando escaneado de 1-bit en el ambiente digital. Esto es comparable al microfilme tradicional. Si los documentos poseen un bajo contraste inherente tipo / texto, si los caracteres tienen bordes difusos (como por ejemplo, las copias al carbón u otros tipos de procesos de copiado), o si hay fotografías, entonces usted deberá digitalizar a 8-bits (256 tonalidades de gris) como mínimo, o utilizar película fotográfica de bajo o medio contraste de tono continuo. Por último, para gráficos a color, textos a color o fotografías a color, se necesita capturar la información de color además de las tonalidades. Como mínimo, la digitalización deberá realizarse mediante escaneo a 24-bit RGB (16 millones de colores; también puede utilizar color fotográfico.

A este capítulo contribuyó Don Willis, cuyo papel precursor en la preservación se cita en el capítulo VII: Un estudio de caso— El trabajo con Microfilme.

Bibliografía

Besser, Howard y Jennifer Trant. *Introduction to Imaging: Issues in Constructing an Image Database*. Getty Art History Information Program, Santa Monica, CA, 1995. [Online] http://www.getty.edu/gri/standard/introimages/

Frey, Franziska y James Reilly. *Digital Imaging for Photographic Collections: Foundations for Technical Standards*. Image Permanence Institute, Rochester Institute of Technology, Rochester, NY, 1999. [Online] http://www.rit.edu/~661www1/sub_pages/frameset2.html

A Guide to Digital Photography: Theory and Basics. Agfa Educational Publishing, Randolph, MA. [Online] http://www.agfa Home.com/publications/

An Introduction to Digital Scanning. Agfa Educational Publishing, Randolph, MA. [Online] http://www.agfa Home.com/publications/

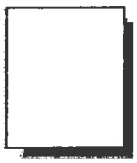
Kenney, Anne y Oya Rieger. *Moving Theory Into Practice: Digital Imaging for Libraries and Archives*. Research Libraries Group, Mountain View, CA, 2000. [Online] http://www.rlg.org/preserv/mtip-order.html

Kodak Digital Learning Center. [Online] http://www.kodak.com/US/en/digital/dlc/book3/

RLG DigiNews on-line newsletter. [Online] http://www.rlg.org/preserv/diginews/

The Secrets of Color Management. Agfa Educational Publishing, Randolph, MA. [Online] http://www.agfa Home.com/publications/

Buscar Contenido Sitio del NEDCC



Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

VII

Desarrollo de las mejores prácticas: Lineamientos a partir de estudios de casos

Introducción

Este capítulo contiene seis estudios de casos que propician al lector pasar de la visión teórica a la forma en que se debe conducir la digitalización en la práctica, desde la planificación del proyecto, hasta la ejecución y la evaluación del mismo. Algunas de las secciones se concentran fundamentalmente en las experiencias de una institución, mientras que otras secciones son compilaciones de lo aprendido en diversas situaciones.

Los estudios de casos incluyen descripciones de lo que ha funcionado y lo que no ha funcionado en los proyectos. Siempre que ha sido posible, los autores han incluido consejos para aquellos que se inician en nuevos proyectos de escaneo.

1. Cómo se trabaja con textos impresos y con manuscritos

Stephen Chapman
Harvard University Library

Si observamos el grupo cada vez más voluminoso de libros, revistas científicas y archivos publicados por instituciones culturales y editoras comerciales a los que se puede acceder a través de Internet, descubriremos que los textos en formato electrónico no son todos iguales. Algunas colecciones son buscables, otras no lo son, algunas cuentan con reproducciones a color de alta calidad, otras limitan su contenido a imágenes en blanco y negro (1-BIT); algunas proporcionan navegación del tipo "vaya a la página" o "vaya a la sección", y muchas simplemente ofrecen función "página anterior", "página siguiente". Más que presentar un solo estudio de caso de un tipo de texto electrónico, esta sección presenta una compilación de estudios de casos en los que se analizan los desafíos de cada uno de los varios tipos de conversión de textos y los lineamientos que han surgido en respuesta a esos desafíos.

Como los costos varían ampliamente entre una y otra versión, la primera tarea del gerente que está al tanto del presupuesto es seleccionar el modelo de publicación electrónica menos costoso y más apropiado según la(s) colección(es) que se han seleccionado para el proyecto de digitalización. En general, los textos electrónicos se dividen en tres categorías:

> ***Imágenes de páginas*** Estas *fotocopias digitales* se crean cuando se escanean páginas impresas o microfilme. Las imágenes de páginas no son buscables. Pueden ser en blanco y negro, en escala de grises, o en color. Se puede asumir que las páginas de imágenes en blanco y negro (1-BIT) son los productos menos costosos en la digitalización de textos, pero asegúrese de tener en cuenta los costos asociados en que hay que incurrir para preparar los metadatos estructurales necesarios y hacer que las imágenes sean adecuadas para buscar y para la navegación en línea.
>
> ***Texto pleno*** Para que un texto impreso sea plenamente buscable desde el punto de vista electrónico (texto pleno), es necesario traducir las letras de las páginas originales al código ASCII que la máquina puede procesar. Existen dos formas de hacer esto: se copia, mecanografiando, del original (lo que se conoce como *teclear*) o se utiliza un programa de reconocimiento de caracteres ópticos (OCR) para convertir las imágenes de páginas a código ASCII. El primer proceso es manual, el segundo automatizado. Como teclear puede ser diez veces más costoso que la combinación de escanear más OCR, con frecuencia se hacen las imágenes de páginas de forma tal que se facilite

la creación de texto pleno. Cuando se hacen estos dos productos (texto pleno e imágenes de página), se obtiene la ventaja adicional de poder presentar una versión *facsimilar* de la página original -- con las mismas fuentes, el formato, las anotaciones al margen, e ilustraciones intactas -- en respuesta a una búsqueda. En otras palabras, se utiliza el código ASCII para crear un índice de búsqueda, y sólo las imágenes de páginas se entregan a la pantalla o la impresora.

Pero uno podría preguntar: Si escanear con OCR es mucho más barato que teclear, ¿para qué ni siquiera pensar en la posibilidad de teclear? En primer lugar, el OCR es viable sólo para imágenes de páginas de texto impreso en máquinas de impresión. Es necesario teclear los originales manuscritos para que sean buscables. En segundo lugar, la precisión del OCR disminuye según aumenta la complejidad de los originales (número de fuentes, número de columnas, ilustraciones que acompañan al texto), y según disminuye la calidad de las imágenes de páginas. Por tanto, si se requiere casi el 100% de la precisión, podría ser menos costoso teclear que acometer los tres pasos del proceso de escaneo, aplicar el OCR, y rectificar los errores del OCR. Varios estudios confiables afirman que un técnico bien preparado puede corregir entre 6-10 páginas por hora. En dependencia del salario, esta tarea sola puede exceder fácilmente el costo de teclear.

Texto Codificado, o Texto Pleno con Marcaje Este tercer modelo de publicación para conversión de texto es el más costoso, pero también es el más funcional y flexible en el ambiente de las publicaciones en línea. Al igual que el texto pleno simple, la producción del texto codificado exige teclado u OCR de imágenes de página para crear ASCII. El paso final es codificar atributos de un determinado documento colocando etiquetas de Lenguaje Generalizado de Marcaje Estándar (Standard Generalized Markup Language) (SGML) alrededor del texto seleccionado. Existen cientos de elementos de SGML que se pueden utilizar para codificar. Los Lineamientos de la Iniciativa para la Codificación de Textos (TEI) se refieren a un grupo que ha sido ampliamente utilizado para publicaciones en las ciencias humanísticas. Por lo general, los textos se codifican en uno de los dos niveles siguientes o en ambos a la vez: (1) estructural: que se refiere a las

divisiones como los capítulos dentro de los libros, los artículos dentro de las revistas científicas, los poemas dentro de las antologías; o (2) descriptiva: que se refiere a los elementos como por ejemplo fechas, nombres de personas o lugares y ocupaciones. Cuando se apareja una aplicación / interfase de búsqueda que ha sido bien configurada con una base de datos SGML, la codificación posibilita la búsqueda por campos (*Ej.* , encontrar "esclavitud" en epígrafes), y también se puede utilizar para controlar la presentación del documento -- incluidas múltiples representaciones si se desea.

Nota: No es necesario crear imágenes de páginas para producir texto codificado si (1) teclear fuera un enfoque permisible para la producción, y (2) el objetivo es presentar páginas modernas y no facsimilares en la pantalla o la impresora.

Después de decidir cuál es el producto (o productos) electrónico que satisface los requisitos del proyecto, la segunda tarea del gerente debe ser especificar el resultado para los materiales originales después de la conversión. Como los originales impresos son producto de la conversión de textos, es importante determinar si deben presentarse exactamente en la forma que era el original o si se aceptan ciertas alteraciones. Es mucho más fácil crear un presupuesto y un plan de trabajo para el proyecto si se toman decisiones de disposición -- relativas a las políticas de acceso, la envoltura y ubicación de los materiales, e incluso la limitación de acceso -- desde el principio.

Las decisiones con respecto a los resultados adecuados para las fuentes informan, si no determinan, los lineamientos de manipulación para el escaneo. En algún momento del proyecto será necesario evaluar, limpiar, restaurar o reparar o reubicar los materiales, sobre todo si se van a mantener como copias circulantes. Por otra parte, los materiales que se van a trasladar para su almacenamiento en otro lugar, o incluso los que se van a desechar permiten un elevado número de opciones en técnicas de escaneo.

Los aspectos referentes a la manipulación y disposición de los materiales son especialmente importantes en el caso de los materiales encuadernados. Las páginas sueltas, aún cuando están en condiciones muy frágiles, se pueden escanear en escáners planos, o se pueden introducir de manera automática en escáners de alimentación automática (con esteras rectas para papel). En otras palabras, es mucho menos costoso escanear páginas sueltas que escanear libros encuadernados. Hasta 1999, las estadísticas de producción recogidas en varios proyectos

indican que, aunque los técnicos pueden escanear hasta cinco páginas por minuto, el promedio típico está entre dos y tres páginas. Por otra parte, los escáners de auto alimentación pueden escanear las dos caras de una hoja de una sola pasada. Si utilizamos estas mismas estadísticas de salida (*Ej.. ,* 600 dpi1-BIT TIFF), estos escáners producen 20 imágenes por minuto. Por tanto, la evaluación de los originales es un aspecto crítico, ya que siempre que sea necesaria alimentación manual (o volteo de páginas), los precios de escaneo estarán directamente relacionados con los costos de la mano de obra. En este modelo, los adelantos que se producen en la tecnología de escaneo, sólo pueden resultar en una mejor calidad, pero no necesariamente en mayor velocidad. Cuando se puede utilizar auto alimentación, los adelantos en la tecnología de escaneo no sólo pueden resultar en más calidad, sino también en más velocidad, y por tanto menor costo por unidad.

Las decisiones referentes a la manipulación adecuada son complejas, y el método utilizado debe ser probado y confirmado con una muestra de los materiales antes de acometer la producción a plenitud. No ha surgido aún un sistema único que sea el mejor para escanear. Todos los sistemas, el de auto alimentación, el escáner plano, el escáner de vista superior, e incluso los sistemas de cámara digital son viables. Cuando se selecciona un escáner o se escriben especificaciones para un sistema que presta servicios de este tipo, lo primero que debe especificarse son los requisitos de manipulación, después se presentan los requisitos de calidad de imagen y de velocidad. El software de escaneo desempeña un importante papel en estas áreas. Por ejemplo, los mismos parámetros de entrada -- *Ej.* , 600 dpi 1-BIT TIFF -- en escáners diferentes producirán diferentes resultados de salida. Los parámetros para lotes con frecuencia diferencian a los sistemas de alto precio de los de bajo precio y son de importancia crítica para aplicaciones de alto volumen.

Reglas

Aunque hay muchas variables asociadas con la selección de los mejores métodos para crear imágenes de páginas y/o textos plenos, afortunadamente existen algunas reglas comunes a muchos proyectos de conversión de textos.

- Para minimizar los costos de creación y de mantenimiento de las imágenes de páginas, siempre que sea posible deberá escanearse a 1-BIT con compactación sin pérdida; el uso de escáners de auto alimentación es la forma menos costosa de producir imágenes de calidad suficiente para OCR, impresión, y/o microfilme de computación (COM). La calidad que se obtiene con todos los escáners de 1-BIT -- los de auto alimentación, los planos, y los de imagen superior -- es el resultado de una combinación de ingeniería (hardware, óptica), software, y habilidad

del operador, por lo que debe asegurarse de confirmar que los requisitos de resolución citados en un proyecto funcionan de igual manera para los materiales y el escáner que Ud. ha seleccionado para su proyecto.

- Cuando se prefiere el uso de la escala de los grises o el color para textos impresos en imprenta, utilice un escáner o una cámara digital con resolución espacial suficiente para capturar las líneas, los bordes y otros detalles de los materiales originales. Compare los costos y la calidad de los sistemas de ordenamiento de líneas y ordenamiento de área para determinar cuál produce una calidad aceptable con el menor costo. Si se necesitara OCR, se necesitarán sistemas de procesamiento de imágenes sofisticados (inmediatamente después de escanear) para generar archivos de 1-BIT a partir de los archivos escaneados en escala de grises o a color.

- Cuando se ordena que se apliquen evaluaciones o diagnósticos de conservación y/o tratamientos a los materiales originales, los conservadores deberán participar en la selección de los equipos de escaneo que se utilizarán y en la redacción de los lineamientos de manipulación del proyecto.

- Los requisitos de calidad de imagen y de control de la calidad están directamente relacionados con la disposición de los originales. Los requisitos de calidad serán más elevados para los proyectos en los que se necesita acceso reducido a los originales o hasta reposición de los mismos. En ese caso, resulta irónico que los costos serían menores, ya que se puede considerar que la auto alimentación es una técnica más aceptable para esos objetos que para materiales únicos en buen estado de conservación.

- Es probable que los costos de preparación de los documentos (sin incluir los tratamientos de conservación), la creación de metadatos, y el control de la calidad excedan el costo de escanear, en especial para la digitalización a 1-BIT.

- Teniendo en cuenta el diseño de los escáners de vista superior, así como la profundidad limitada del campo visual de la mayoría de las cámaras digitales, será menos costoso escanear los volúmenes encuadernados si se pueden abrir bien (a 180 grados). Siempre será difícil capturar los textos impresos en el margen interior o muy cerca de esos márgenes -- y según aumenten los requisitos de manipulación, aumentarán también los costos.

- Siempre es más costoso escanear las páginas de grandes dimensiones (sobre todo cuando la dimensión más larga es mayor que 17"). La reproducción con alta calidad de textos se hace más difícil con escaneado directo; por ejemplo, los periódicos siempre se han microfilmado primero para poder producir páginas de imágenes de calidad adecuada.

- Muchas técnicas de mejoramiento de imágenes, como el desmoteado y la corrección de la oblicuidad se pueden automatizar inmediatamente después de escanear. El procesamiento de la imagen es importante, no sólo para la apariencia de las imágenes de páginas, sino también para la optimización del OCR.

- Los metadatos estructurales necesarios para organizar las imágenes de páginas se pueden crear antes, durante o después del escaneado. Teniendo en cuenta la idiosincrasia de la paginación y organización de muchas colecciones históricas, debe esperarse que estas tareas se realicen de forma manual, o en el mejor de los casos semiautomática.

- Los requisitos de la precisión del texto pleno y de la profundidad de la codificación son resultado de un cuidadoso análisis de los materiales originales y de las consultas con la(s) comunidad(es) interesada(s) en utilizar las colecciones digitales.

La tabla que aparece a continuación, resume las decisiones de más impacto en la calidad y el costo de los proyectos de conversión de textos. Se han propuesto muchos lineamientos a partir de la experiencia de los estudios de casos, y los mismos se han generalizado para la tabla. Sin embargo, como ya se ha analizado en capítulos anteriores, una buena gestión comienza con el planteamiento de objetivos, y no siguiendo lineamientos a ciegas. Relacione las decisiones que Ud. tome con los objetivos de la publicación y con los resultados deseados para los originales, y de manera natural vendrán los lineamientos de escaneo y los costos.

DECISIONES CLAVES PARA LA CALIDAD Y EL COSTO EN LOS TEXTOS DIGITALIZADOS

Producto	**Ejemplos de decisiones claves**	**Lineamientos**
Originales	*Manipulación*	
	* Se permite contacto con el vidrio	Viable con todos los tipos de escáners
	* Los volúmenes encuadernados deberán estar bien apoyados durante el escaneo (abiertos a menos de 180°)	Se necesita el tipo de escaneo de sólo cara superior, con un apoyo adecuado para el libro (cradle)

	Disposición * Mantener estándar de acceso: retorno en formato original a la ubicación original	Identifique los recursos de que dispone para tratamientos. Si cuenta con fondos y con personal, por ejemplo, para valorar, desmontar y re encuadernar los materiales, compare entonces los costos de escaneo de las páginas en contraposición con los de escanear libros antes de seleccionar el mejor enfoque.
	* Reducir el acceso cambiando la política de circulación o mediante reubicación	Para ahorrar costos, utilice auto alimentación si fuera posible, pero asigne un presupuesto para la preparación necesaria de los materiales y para los costos de re envoltura.
	* Reducir de manera drástica o incluso eliminar el acceso creando imágenes digitales con calidad de sustitución y/o deshaciéndose del acceso a los materiales después de escaneados	Los requisitos de control de la calidad y metadatos deben estar definidos de manera explícita (tenga en cuenta el uso de micromiras); desmontar la encuadernación puede ser más apropiado en estas circunstancias.
	Preparación * Facilite la más alta calidad de escaneo al menor costo	Separe los materiales en lotes siempre que sea posible (*Ej.* , por tamaño; o por contenido -- texto, ilustración, mezclado, color)
Imágenes de páginas	*Especificaciones para imágenes master (de archivo)* * Debe lograr la reproducción del tono apropiada según los originales y los requisitos de salida	Cuando el blanco y negro (1-BIT) no logra capturar la información esencial, use escáners que muestrean 12-bits por píxel y con al menos 8-bits de salida por píxel para escala de grises y 24-bits por píxel para color.
	* Para textos impresos en imprenta, debe lograr la reproducción de los detalles necesarios para satisfacer los requisitos de salida (pantalla, impresión, OCR para textos de imprenta)	400-600 dpi es lo que se usa; las capacidades de umbral y de procesamiento de imágenes también son críticas para la calidad de la imagen, especialmente para las imágenes de 1-BIT; el mejoramiento posterior al escaneo puede incrementar la precisión de OCR

*Para manuscritos escritos a mano y tipos de bordes blandos, como las copias fotostáticas, debe alcanzarse el detalle de reproducción necesario para satisfacer los requisitos de salida (pantalla, impresión, zoom)	300 dpi mínimo para 1-BIT, 200-400 dpi mínimo para escala de grises y color
*Use formato abierto	TIFF
*Use compactación segura	Compactación Grupo 4 (sin pérdida) para 1-BIT, para escala de grises e imágenes a color
* Implemente un programa de control de la calidad	Confirme que todos los archivos por objetos se han recibido, que la secuencia es la correcta, que los metadatos están completos y correctos (100%); compruebe la calidad de la imagen en la pantalla, impresa o de ambas formas (muestra)
Especificaciones para la entrega de la imagen (derivados)	
* Impresión, microfilme con salida por computadora (COM)	Imágenes master (alta resolución TIFF), PDF, o PostScript
* Imágenes en pantalla	La legibilidad por lo general se alcanza a 80-120 dpi; minimice el tamaño del archivo utilizando menos que los 8-bits para GIF siempre que sea posible (*Ej.* , 4-BIT); si fuera necesario usar entre 8 y 24-bits, considere JPEG
Especificaciones para la navegación	
* Página anterior, página siguiente	Incluya campo de *secuencia* en la base de datos de la imagen, o incluya la secuencia en los nombres de los archivos
* Ir a página	Incluya el campo del *número de página* en la base de datos de la imagen, o incluya el número de la página en el nombre del archivo (esta última es por lo general una solución más costosa)

	* Ir a sección	Incluya campo de *característica* o *código de característica* en la base de datos de la imagen, o marcaje de texto pleno (véase a continuación)
Texto pleno	*Especificaciones para precisión (solo caracteres)* * 100%	Primero obtenga precios para teclear, después realice pruebas de muestreo OCR de imágenes de páginas
	* menos de 100%	Realice pruebas de muestra de OCR de las imágenes de página y revise la aceptabilidad de la salida; evite a toda costa la necesidad de corregir textos generados en OCR
Texto marcado	*Especificaciones de precisión (caracteres y formateado)* * La fidelidad al original escaneado que se recomienda	Teclear / codificar puede ser la variante menos costosa; pruebe el escaneado/ OCR solo si el diseño original y las fuentes son relativamente simples
	Especificaciones para codificar * Acomode los atributos de los materiales en la mano mientras utiliza prácticas endosadas por la gran comunidad	Consulte TEI LITE y cree DTD para acomodar las divisiones estructurales y las características descriptivas de los textos en la mano; son posibles las interpretaciones locales de los lineamientos generales

Bibliografía

Bicknese, Douglas A. *Measuring the Accuracy of the OCR in the Making of America.* Invierno1998. http://moa.umdl.umich.edu/moaocr.html

Guthrie, Kevin M. "JSTOR: From Project to Independent Organization." *D-Lib Magazine* (julio-agosto, 1997). http://www.dlib.org/dlib/july97/07guthrie.html

Morrison, Alan, Michael Popham y Karen Wikander. "Creating and Documenting Electronic Texts: A Guide to Good Practice." *AHDS Guides to Good Practice.* Arts and Humanities Data Service, 2000. http://ota.ahds.ac.uk/documents/creating/

Text Encoding Initiative (TEI). "The TEI Consortium Homepage." http://www.ctei-c.org/

University of Virginia Library Electronic Text Center. The Electronic Text Center: On-Line Helpsheets.
http://etext.lib.virginia.edu/helpsheets/sgmlscan.html

2. El trabajo con fotografías

Franziska Frey
Image Permanence Institute

¿Por qué son diferentes las fotografías?

Varios aspectos separan las fotografías del resto de los documentos que se van a escanear.

La permanencia

Los materiales que componen las fotografías no son químicamente estables. Entre estos materiales están la plata o los colorantes como materiales que forman la imagen; el papel, el celuloide, u otros plásticos como materiales de soporte; y además, la gelatina, la albúmina, o el colodión como aglutinantes. La influencia de los factores ambientales como la luz, los agentes químicos, el calor, la humedad y las condiciones de almacenamiento afectan y destruyen los materiales fotográficos. El único método confiable para preservarlas por un largo período de tiempo es almacenarlas en un lugar oscuro con una temperatura baja y poca humedad.

Los administradores de las colecciones, que se enfrentan al deterioro de las fotografías en forma de desvanecimiento del color, el síndrome del vinagre en las películas de acetato, y a las películas de nitrato deterioradas e inflamables, debaten si es mejor invertir en el mejoramiento de las colecciones de almacenaje o en la reconversión de formato.

La complejidad

Muchos proyectos de digitalización de fotografías surgieron a partir de proyectos de digitalización de textos. Este enfoque puede conducir a problemas porque es necesario tratar las imágenes de forma muy diferente cuando las mismas se van a digitalizar.

El principal objetivo cuando se digitalizan colecciones de documentos de textos es la legibilidad. Sin embargo, existen muchos aspectos de calidad diferentes a tener en cuenta cuando se digitalizan imágenes. Además, los medios de búsqueda para las imágenes son muy complejos. Todavía se llevan a cabo investigaciones para determinar cómo facilitar mejor las búsquedas efectivas.

La evaluación de las colecciones

La catalogación, si no se ha realizado aún, limitará el avance del proyecto más que la falta de tecnología.

Antes de que comience el proyecto de digitalización, debe evaluarse la colección muy cuidadosamente. No sólo las imágenes sino también el sistema de catalogación debe evaluarse. A la larga, los inadecuados métodos actuales de descripción de imágenes y la enorme cantidad de catalogación que todavía está por realizar con las colecciones de imágenes serán factores que limitarán el avance hacia el futuro digital más que la falta de una tecnología de imaginología adecuada.

Los tipos de fotografías

Las fotografías se pueden clasificar en dos grupos, en dependencia de si la imagen se puede ver por medio de luz reflejada o luz transmitida. La diferencia más importante entre las dos es el rango dinámico-- es decir, la diferencia entre las áreas más claras y las más oscuras de la imagen. Las fotografías impresas de reflexión de cualquier tipo, por lo general, tienen un rango dinámico menor que el de los negativos. Las transparencias de color tienen el mayor rango dinámico.

Negativos Las colecciones de negativos se benefician de manera especial con la digitalización, ya que las hace fácilmente accesibles. Millones de negativos nunca se usan simplemente porque el contenido de la imagen no está disponible al usuario. Se necesita un proceso de impresión para obtener una imagen. Por tanto, no sólo el público, sino también, y con frecuencia, hasta los mismos administradores de las colecciones no saben lo que contiene una colección de negativos. Ya se ha demostrado que, tan pronto como se escanean los negativos y se puede ver una imagen positiva, su uso se incrementa enormemente. Una inmensa cantidad de negativos antiguos es de negativos sobre placa de vidrio. La decisión de digitalizarlos reduce el riesgo de pérdida por rotura, puesto que sólo es necesario manipularlos una vez.

Tipo de Color Otra forma de clasificar las imágenes es por el tipo de color. En dependencia del tipo de color, las imágenes se escanean en blanco y negro o en color.

Color Total-- La mayoría de las fotografías actuales se hacen a todo color. Sin embargo, esta tendencia sólo data de mediados de la década de 1960. Esto quiere decir que la mayoría de las colecciones no van a incluir demasiadas fotografías a color, cosa que cambiará cuando más

fotografías a color se incorporen a las colecciones de bibliotecas y archivos.

Color monocromático -- Una gran cantidad de las fotografías a escanear serán monocromáticas (Reilly). Muchos procesos de impresión fotográfica del siglo XIX poseen colores característicos, *Ej.*, el carmelita violáceo de las impresiones a la albúmina y el color azul de los cianotipos. Esos colores ayudan a la interpretación académica ya que transmiten información acerca de los procesos y proporcionan detalles acerca del grado de deterioro que pueden haber sufrido las fotografías. Es importante guardar esta información sobre el color en el archivo digital, ya que es una característica inherente de la imagen.

Blanco y negro -- Las fotografías en blanco y negro tomadas con película monocromática son o de color negro neutral o no poseen información visual significativa que transmitan las imágenes a través del color. Éstas son, fundamentalmente, negativos o impresiones modernas de revelado por gelatina de plata hechas en el siglo XX.

Fotografía electrónica Cada vez más, las colecciones incluyen imágenes que nunca tuvieron un original en película. El cuidado de los originales en formato electrónico exige que los administradores de las colecciones presten atención a nuevas especialidades como por ejemplo el formato de los archivos, las leyes de propiedad intelectual, tecnología de transferencia de datos a alta velocidad, y gestión de bases de datos.

Formatos

Con frecuencia, las colecciones de imágenes tendrán una variedad de formatos, aunque pueden predominar determinados formatos (*Ej.*, negativos, impresiones). Esta variedad exige el uso de equipos de escaneo versátiles.

Estado de conservación de la colección

Llevar a cabo un estudio diagnóstico de evaluación de la colección antes de escanear ayudará a tomar decisiones con respecto a qué se debe seleccionar. Puede también dar lugar a la aplicación de un plan para controlar las condiciones de conservación de las colecciones en el futuro, por ejemplo proporcionando mejores condiciones de almacenamiento o materiales de envoltura. También muchas veces la preparación de la colección para escanearla incluye el mejoramiento de las condiciones físicas de conservación de la colección.

El tamaño de la colección

El tamaño de la colección también influye en el método de escaneo y en los parámetros. Si la colección es muy pequeña, puede elegir un método que consuma mucho tiempo. Un buen ejemplo es la National Gallery de Londres, que escaneó cada una de las pinturas de la colección con una cámara especial de espectro múltiple. Como la colección sólo consta de 3,000 pinturas, se pudo escanear todo varias veces. Esta no es una solución posible en el caso de una colección que conste de varios miles de imágenes, que no es probable que se vuelvan a escanear en el futuro próximo. Además, en el caso de las colecciones grandes, es necesario planificar con mucho cuidado el flujo de trabajo.

Los objetivos de la digitalización

Como no es probable que se acometa la digitalización de colecciones de gran tamaño más de una vez en una generación debido a los costos, se imponen decisiones inteligentes con respecto a los procesos para escanear y archivar. El término *de archivo* implica que todas las imágenes digitalizadas se optimizan, no sólo para la corriente de flujo de trabajo y dispositivos de digitalización de imágenes, sino para que sigan siendo utilizables en el futuro con sistemas de entrega y salida todavía desconocidos (Frey & Süsstrunk, 1996; Frey, 1997; Frey & Süsstrunk, 1997).

Uno de los grandes problemas que deben tener en cuenta las instituciones antes de comenzar a ejecutar un proyecto de digitalización es el uso que se anticipa tendrán esas colecciones de imágenes digitalizadas. Existe consenso dentro de la comunidad de preservación en cuanto a que es necesario crear una cierta cantidad de archivos de imágenes de cada una de las fotografías para satisfacer un rango de usos. En primer lugar, hay que crear una imagen master o de archivo. Un fichero master o de archivo debe representar la más alta calidad que una institución pueda obtener y pagar. No deberá tratarse para ningún propósito específico de salida y deberá mantenerse sin compactar o compactarse sin pérdida. También necesitará una revisión intensiva de calidad. A partir de este fichero de archivo, se pueden calcular varios derivados. Estos archivos derivados se crean con el propósito de ser utilizados. Las principales cuestiones a tener en cuenta cuando se crean estos archivos derivados son la velocidad de acceso y de transmisión y lo adecuados que sean para ciertos propósitos.

¿Escanear el duplicado o el original?

Hay que tomar una decisión con respecto a si se escanea del original directamente o si se escanea el duplicado. En cada uno de los casos hay ventajas y desventajas. Debido a que cada generación de copia fotográfica conlleva una cierta pérdida de calidad, el uso de intermediarios inmediatamente implica alguna disminución en la calidad.

Los intermediarios también pueden servir, sin embargo, otros propósitos: por ejemplo, pueden servir como masters para copias fotográficas de referencia, o como sustitutos de preservación.

Esto nos lleva a la cuestión de si se debe utilizar el negativo o la impresión para digitalizar, en el supuesto caso de que se disponga de ambos. La calidad siempre será mejor si se utiliza la primera generación de una imagen (*es decir,* el negativo). No obstante, puede haber grandes diferencias entre el negativo y la impresión, fundamentalmente en el reino de la fotografía de arte. Con frecuencia el artista pasa mucho tiempo en el cuarto oscuro creando la fotografía impresa. Los resultados de todo este trabajo se pierden si se escanea el negativo, en lugar de la fotografía impresa. El resultado de la digitalización puede ser bastante desalentador.

Control de la calidad

Inspección visual subjetiva

El mejor tratamiento hacia control de la calidad en la digitalización de imágenes incluye, por una parte, la inspección visual subjetiva y, por otra, medidas objetivas realizadas en el software y directamente en los archivos digitales. Deben realizarse esfuerzos para normalizar los procedimientos y equipos para la evaluación subjetiva.

En la mayoría de los casos, la primera evaluación de una imagen escaneada será visualizarla en un monitor. El que la ve, decidirá si la imagen en el monitor cumple los objetivos establecidos al principio del proyecto de escaneo. Este aspecto es importante porque el juicio de un humano será el que determine la aceptabilidad final de una imagen. Visualizar imágenes y juzgar la calidad de las mismas siempre ha sido una tarea compleja. La persona tiene que conocer qué es lo que está buscando. Debe insistirse en que el control subjetivo de la calidad debe ser ejecutado en equipos calibrados y en un ambiente de visualización normalizado y adecuado.

Cuando se visualiza una imagen en el monitor, se pueden detectar defectos como suciedad, imágenes incompletas, oblicuidad, imágenes invertidas y falta de precisión. En algunos casos puede que sea necesario ir atrás y volver a escanear.

Cómo evaluar los archivos de imágenes digitales

Por otra parte, deben emplearse parámetros objetivos de calidad de la imagen. Esto se puede lograr escaneando micromiras especiales y evaluándolas en softwares especializados (Gann, 1999; Holm, "Survey," 1996).

Las micromiras y el software de evaluación no son sólo para comprobación por parte del proveedor -- también sirven para garantizar la utilidad a largo plazo de los archivos digitales y para proteger las inversiones de las instituciones.

Marco para la calidad de la imagen

Cuando se analiza la calidad de la imagen, es necesario analizar toda la cadena del proceso de la digitalización (Holm, "Factors," 1996). Además del sistema de escanear, también es necesario tomar en cuenta la compactación, los formatos de los archivos, el procesamiento de la imagen para los diversos usos, y la calibración del sistema. La calidad de la imagen se afecta por la secuencia de aplicación de los diferentes pasos de procesamiento de la imagen, entre los que se incluye la compactación. Procesar las imágenes antes de almacenar las imágenes puede afectar la calidad del procesamiento futuro. Por ejemplo, se recomienda no perfeccionar el master de archivo antes de almacenar.

Cada uno de los parámetros fundamentales de calidad de la imagen necesita micromiras especiales para las diferentes formas de imágenes (*Ej.,* impresos, transparencias). Las micromiras deben ser del mismo material que los materiales que se van a escanear – película fotográfica y papel.

Estas micromiras son una parte vital del marco de calidad de la imagen. Después que se escanean las micromiras, se evalúan con un programa de computación (software). Existen algunos componentes de software que funcionan desde como plug-ins hasta examinadores de imágenes totales, otros son programas en si mismos. Sin embargo, debe aclararse que todavía no se dispone de manera comercial de algunas de las micromiras y del software para evaluarlas.

Las micromiras se pueden incorporar en el flujo de trabajo de varias maneras. Versiones completas de las micromiras se podrían escanear cada cierto número de imágenes, cada varios cientos de imágenes y después vincularlas a lotes específicos de archivos de producción, o se podrían incluir con cada imagen versiones más pequeñas de las micromiras. Según más instituciones acometen proyectos de digitalización, será cada vez más importante contar con una herramienta objetiva para comparar los diferentes dispositivos de escaneo.

Reproducción de la tonalidad

La reproducción de la tonalidad es el parámetro más importante para determinar la calidad de una imagen. Si la reproducción de la tonalidad de una imagen es la correcta, en general los usuarios encontrarán que la

imagen es aceptable, incluso si algunos de los otros parámetros no son óptimos. Para que exista el concepto de reproducción de la tonalidad, debe producirse la captura y la presentación de la imagen. Esto quiere decir, que deben hacerse ciertas presuposiciones en lo referente al dispositivo visor final. Tres atributos mutuamente dependientes afectan la reproducción de la tonalidad: la función de conversión opto-electrónica (OECF), el rango dinámico, y el destello. La OECF muestra la relación entre las densidades ópticas de un original y los correspondientes valores digitales del archivo. El rango dinámico se refiere a la capacidad del escáner para capturar variaciones de densidad extremas. El rango dinámico del escáner debe coincidir con el rango dinámico del original o exceder al mismo. El destello se genera por una luz intermitente en un sistema óptico. El destello reduce el rango dinámico de un escáner.

Reproducción del color

En una imagen digital, se pueden aplicar varios intentos de reproducción del color. Intento perceptual, intento colorimétrico relativo e intento colorimétrico absoluto son los términos que con más frecuencia se asocian con el Consorcio Internacional del Color (ICC). Intento perceptual es crear una imagen agradable en un determinado soporte bajo ciertas condiciones de visualización. Intento colorimétrico relativo es hacer coincidir, lo más posible, los colores de la reproducción con los colores del original, tomando en cuenta el soporte de salida y las condiciones de visualización. Intento colorimétrico absoluto es reproducir los colores lo más exacto posible, sin tener en cuenta el soporte de salida y las condiciones de visualización.

La mayoría de las soluciones de que se dispone para medir y controlar la reproducción del color están orientadas hacia la industria de pre impresión. Sin embargo, cuando se escanea una imagen para propósitos de archivo, el uso futuro de la imagen no se conoce todavía. Las decisiones del operador con respecto al color y el contraste no se pueden revertir en un sistema de color RGB a 24-BIT. Es necesario tener en cuenta cualquier mapeo de salida diferente al espacio de color y gamma de la imagen archivada. No obstante, salvar los datos del escáner en bruto puede crear problemas si las características del escáner no se conocen bien y están bien especificadas.

Una de las decisiones a tomar es cuál espacio de color utilizar. Un espacio de color es una representación geométrica de colores en el espacio, casi siempre de tres dimensiones. La razón para las tres dimensiones es el sistema visual humano, que tiene tres receptores independientes y es por tanto un sistema tridimensional. El atributo más importante de un espacio de color en un ambiente de archivo es que esté bien definido. Además, es necesario recordar que existe más de una

solución posible para el problema. El espacio de color correcto depende del propósito y del uso de las imágenes digitales (Süsstrunk, Buckley & Sven, 1999).

Resolución

Una revisión de proyectos digitales anteriores nos muestra que las personas se preocupan más por la resolución espacial. No es para sorprenderse, porque de todos los eslabones débiles en la captura de la imagen digital, la resolución espacial es el que la mayoría de las personas comprende mejor. Sin embargo, la tecnología ha evolucionado, y en la actualidad lograr una resolución especial razonable no es ni extremadamente costoso ni tampoco cuesta demasiado almacenar grandes archivos de datos. La resolución especial es el parámetro para definir lo detalle y los bordes de la reproducción en una imagen. Los detalles pueden ser, por ejemplo, un cabello aislado en un retrato. Una buena reproducción es crítica para la precisión visual de una imagen. La resolución espacial de una imagen digital, *es decir,* el número de detalles que contiene la imagen, se define por lo general por el número de pixels por pulgada (ppi). Mientras mayor sea el número de pixels por pulgada, más detalles se podrán transferir de la imagen original al archivo digital.

Hallar el número equivalente de pixels que describe el contenido de información de una emulsión fotográfica específica no es un proceso directo. Hay que tener en cuenta el formato del original, el grano de la película, la resolución de la película, la exposición y las técnicas de procesamiento para determinar el verdadero contenido de información de una imagen fotográfica específica.

La mejor medida de los detalles y la resolución es la Función de Transferencia de la Modulación (MTF). La MTF es una representación gráfica de la calidad de una imagen que elimina la necesidad de que el observador tome decisiones.

Ruido

Ruido se refiere a variaciones aleatorias de señales asociadas con sistemas de detección y reproducción. En la fotografía convencional, el ruido en una imagen, es una granulación que se puede percibir. Se puede apreciar con más facilidad en áreas de densidad más uniforme. El ruido es un atributo importante de los sistemas electrónicos de digitalización de imágenes. Contar con un sistema de normas ayudará a los usuarios y los fabricantes a determinar la calidad de las imágenes que se producen mediante esos sistemas. Los resultados de los tests para ruido tienen dos propósitos. Primero, se puede determinar el nivel de ruido del sistema, indicando cuántos niveles de BIT de los datos de la imagen son útiles en

realidad. Segundo, para las consideraciones sobre calidad de la imagen, es importante conocer la relación señal-ruido (S/N). Las evaluaciones S/N muestran el efecto del ruido aleatorio en la calidad de lo escaneado. El ruido aleatorio, más que la profundidad de BIT, es el factor primario limitante de la resolución tonal del escáner. El test puede consistir en escanear dos veces una micromira de la escala de los grises. Se sustraen los valores obtenidos y se analiza la desviación estándar del resultado. La sustracción debe eliminar todos los componentes de no ruido (esta es la información de la imagen) y la desviación estándar es una buena medida del ruido aleatorio (Gann, 1999).

Costos

Elementos del presupuesto

Existe una gran variedad de costos a tener en cuenta (Puglia, 1999):

- Selección
- Preparación
- Catalogación/Descripción/Indización
- Preservación/Conservación
- Producción de Intermedios
- Digitalización
- Control de calidad de imágenes y Metadatos
- Infraestructura de red
- Costos continuos de mantenimiento de imágenes y datos

Los costos iniciales

La conversión digital representa aproximadamente la tercera parte de los costos iniciales. Otros costos, fundamentalmente los relacionados con la catalogación, la administración y el control de a calidad, representan las otras dos terceras partes.

Los costos durante el proyecto

Con frecuencia, la planificación y el presupuesto se detienen después de la creación de los valores digitales. Sin embargo, una parte importante del presupuesto incluye los costos de refrescamiento y migración y de apoyo a los sistemas. Todo esto se puede incluir dentro del gran marco de la *gestión de los valores digitales.* Es difícil, en este momento, decir cifras exactas para este proceso. Sin embargo, dado que tanto la comunidad archivística como la industria gráfica están enfocando este aspecto, muy pronto dispondremos de cifras cada vez más reales. En la actualidad se calcula que entre el 5% y el 10% del presupuesto de los costos iniciales por imagen deben calcularse por año para el mantenimiento de las imágenes en el futuro, aunque la migración y la conversión de archivos no se realice anualmente.

La ejecución del proyecto en la propia institución vs. la ejecución con recursos externos

Se han utilizado muchos proyectos pilotos de digitalización de colecciones de imágenes para construir instalaciones propias de las instituciones. Aunque esto es posible lograrlo cuando se trata de proyectos de poca envergadura, en muchos casos será mejor y también necesario establecer una buena relación con un proveedor y una fuente externa para todo el proceso de digitalización. Incluso este proceso exige un buen conocimiento del proceso de digitalización, porque todos los parámetros de digitalización y de construcción del sistema tendrán que establecerlos las instituciones por si mismas. Es muy importante, tal y como se enfatiza en el capítulo sobre las relaciones con los proveedores, establecer una buena relación con el proveedor.

Conclusiones

Muchos de los problemas que surgen de la necesidad de escanear para el futuro desconocido no se han resuelto todavía, y hay mucha incertidumbre en cuanto a cómo proceder. Los responsables de algunos de los más importantes proyectos de reconversión a formato digital han enfrentado el mismo problema. Los rápidos cambios en la tecnología hacen que sea más difícil escoger el mejor momento para crear una política de reconversión que mañana no sea obsoleta.

Si las instituciones no comunican sus necesidades, no obtendrán herramientas para aplicaciones especiales.

La falta de comunicación entre el campo de la técnica y las instituciones sigue siendo un obstáculo formidable. Tanto las instituciones como la industria están interesadas en el diálogo, pero no hay un lenguaje común. No podemos dejar de insistir y de hacer énfasis en que si las instituciones no comunican sus necesidades a las industrias de hardware y de software, no obtendrán las herramientas que necesitan para sus aplicaciones especiales. Los archivos y las bibliotecas deben saber que ellos participan en la creación de las nuevas normas. En la actualidad, parece que cualquiera que sea el primero en el mercado con un nuevo producto, está creando una norma *de facto* para los competidores. Además, el tiempo que tenemos para crear nuevas normas es muy corto, la industria no esperará años para introducir un nuevo producto por la sencilla razón de

que las personas no se ponen de acuerdo con respecto a determinada cuestión.

No se puede ver un proyecto digital como un proceso lineal en el que una tarea va a continuación de la otra. Un proyecto digital debe verse más bien como una estructura compleja de tareas interrelacionadas en la cual cada decisión tiene influencia en otra decisión. El primer paso para penetrar esta compleja estructura es comprender muy bien cada uno de los pasos y encontrar medidas para cuantificarlos. Una vez hecho esto, las entidades independientes se pueden unir en un mismo contexto. Todavía estamos en el primer asalto del proceso, pero con los beneficios que reportan las experiencias adquiridas en otros proyectos digitales, estamos llegando al punto en el que podemos analizar el proceso como un todo.

Bibliografía

Arms, Caroline, ed. *Enabling Access in Digital Libraries*. Washington, DC: Consejo para las Bibliotecas y Recursos de la Información, 1999.

Ester, Michael. *Digital Image Collections: Issues and Practice*. Washington, DC: Comisión para Preservación y Acceso, 1996.

Frey, Franziska. "Digital Imaging for Photographic Collections: Foundations for Technical Standards," *RLG DigiNews* (diciembre 1997). [Online] www.rlg.org/preserv/diginews

Frey, Franziska y James Reilly. *Digital Imaging for Photographic Collections: Foundations for Technical Standards* (noviembre 1999). [Online] www.rit.edu/~661www1/sub_pages/frameset2.html

Frey, Franziska y Sabine Süsstrunk. "Color Issues to Consider in Pictorial Image Data Bases." *Proceedings IS&T's Fifth Color Imaging Conference*, pp. 112-15. Scottsdale, AZ, noviembre 17-20, 1997.

------. "Image Quality Issues for the Digitization of Photographic Collections." *Proceedings IS&T's 49th Annual Conference*, pp. 349-53. Minneapolis, MN, mayo 19-24, 1996.

Gann, Robert G. *Desktop Scanners*. Upper Saddle River, NJ: Prentice Hall, 1999.

Holm, Jack. "Factors to Consider in Pictorial Digital Image Processing," *Proceedings IS&T's 49th Annual Conference*, pp. 298-304. Minneapolis, MN, mayo 19-24, 1996.

------. "Survey of Developing Electronic Photography Standards." *Standards for Electronic Imaging Technologies, Devices, and Systems,*

SPIE, Critical Reviews of Optical Science and Technology Series 61 (1996): 120-52.

Puglia, Steve. "The Costs of Digital Imaging," *RLG DigiNews* (octubre 1999). [Online] http://www.rlg.org/preserv/diginews

Reilly, James. *Care and Identification of 19th-Century Photographic Prints*. Rochester, NY: Eastman Kodak Publication, 1986.

Stephenson, Christie y Patricia McClung, eds. *Delivering Digital Images--Cultural Heritage Resources for Education*. Los Angeles, CA: The Getty Information Institute, 1998.

Susstrunk, Sabine; Robert Buckley y Steve Sven. "Standard RGB Color Spaces," *Proceedings IS&T's Seventh Color Imaging Conference* (noviembre 1999), Scottsdale, AZ.

3. Un Estudio de Caso de OCR

Eileen Gifford Fenton
JSTOR, University of Michigan

¿Qué es OCR?

Reconocimiento de caracteres ópticos, o OCR, es el proceso que convierte el texto de una página impresa en un archivo digital. Esto se logra mediante el uso de un paquete o software de OCR para procesar la imagen digital que se ha obtenido de una página impresa. El software analiza primero la disposición del texto sobre la página y divide el texto en *zonas* que casi siempre corresponden aproximadamente a párrafos. Después se determina el orden de los párrafos y a continuación comienza el análisis de los caracteres. La mayoría de las aplicaciones de OCR funcionan buscando grupos de caracteres, *es decir,* palabras, y comparándolas en un diccionario que se incluye en la aplicación. Cuando se encuentra una correspondencia, el software imprime la palabra correspondiente al archivo de texto; cuando no se encuentra una correspondencia confiable, el software hace una suposición razonable y señala la palabra como una salida de baja confiabilidad. Cuando no se puede leer una palabra o caracter de ninguna manera, se inserta un caracter sustituto para texto ilegible en el lugar.

La precisión de los paquetes de OCR varía ampliamente.

El OCR es un medio efectivo para leer tipos impresos modernos capturados en imágenes de alta calidad. Aunque los softwares de OCR han mejorado de manera significativa en la última década, los mismos todavía no son eficaces cuando se trata de caracteres no arábicos o no modernos y con frecuencia tienen problemas para tràducir los tipos pequeños, algunas fuentes y páginas con disposiciones complejas. La precisión de los paquetes de OCR varía ampliamente de una aplicación a otra y también de una fuente original a otra.

JSTOR y los OCR

JSTOR, organización no lucrativa independiente radicada en la ciudad de Nueva York, NY, tiene la misión a gran escala de convertir y mantener versiones digitales de números atrasados de revistas científicas y desarrollar herramientas de acceso que permitan la búsqueda tanto de texto pleno como de componentes indexados dentro de cada número. Hasta la fecha, JSTOR ha convertido más de 4 millones de páginas de

más de 100 títulos. Más de 500 bibliotecas académicas en América del Norte y otros países se han sumado al proyecto como participantes institucionales.

JSTOR comenzó digitalizando ediciones anteriores de revistas científicas en el otoño de 1994 sólo con un mínimo de personal dedicado a las actividades de producción. Desde aquellos primeros días, tanto los niveles de productividad como la cantidad de personal han aumentado. En la actualidad, JSTOR prepara unas 200,000-250,000 páginas nuevas para digitalizar cada mes. El personal de producción de JSTOR ha crecido hasta llegar a un grupo de 20 personas distribuidas en las actividades de operaciones en las Universidades de Michigan y la de Princeton. Varias otras unidades de JSTOR incluidas las Actividades Interbibliotecarias, Relaciones con los Editores, Servicios al Usuario, Apoyo Tecnológico y Desarrollo, así como un grupo administrativo complementan el trabajo del grupo de producción.

Cada página digitalizada por JSTOR se procesa mediante una aplicación de OCR, y los archivos de texto resultantes se utilizan para respaldar la búsqueda de texto pleno que se ofrece a los usuarios de JSTOR. Para garantizar que los resultados de las búsquedas sean lo más exactos posibles, cada archivo de texto OCR se revisa de forma manual y se corrige hasta un nivel de precisión extremo antes de incluirlo en la base de datos. Eliminar esta revisión manual podría reducir los costos de producción. Sin embargo, se ha demostrado que es un paso esencial para garantizar tanto la calidad general de la base de datos como la precisión de las búsquedas de los académicos.

Aspectos clave a la hora de elegir un OCR

Los proyectos de digitalización varían mucho en cuanto a contenido, propósito y escala y puede que el OCR no sea la mejor solución para todos. Cuando se analiza la posibilidad de utilizar OCR, es útil analizar los siguientes aspectos.

1) Seleccione una tecnología que mejore la habilidad de cumplir los objetivos del proyecto.

Se ha demostrado que la revisión manual es esencial.

Si los objetivos del proyecto incluyen la entrega al usuario de archivos de texto, es conveniente pensar con mucho cuidado acerca del uso de OCR. Ningún producto de OCR es perfecto. Habrá errores de texto presentes en los archivos que se presentan al usuario. Como resultado de esto, debe determinarse con mucho cuidado también el nivel de precisión de OCR

necesario para satisfacer objetivos particulares. Si se utilizan archivos de texto sólo para apoyar la búsqueda, y no se van a presentar al usuario, se pudiera tolerar menos precisión. Las decisiones con respecto a la precisión deben tener en cuenta las características del material original. Las aplicaciones de ORC no traducen bien los textos que no están en idioma inglés, los símbolos matemáticos o químicos, y otros caracteres especiales, y por lo tanto, la presencia de los mismos debe tenerse en cuenta en la decisión que se tome.

2) La escala importa -- mucho.

El enfoque apropiado para generar archivos de texto se afecta de forma drástica si se pasa de un proyecto de 20,000 páginas a un proyecto de 200,000 páginas a un proyecto de 2,000,000 páginas, aún cuando lo objetivos de todos los proyectos sean los mismos. De igual manera, los costos que genera la producción de archivos de texto también cambia de manera drástica según el tamaño o escala del proyecto.

3) No hay una respuesta correcta única.

Las soluciones estarán determinadas por el objetivo del proyecto. No obstante, es difícil generalizar a partir de un proyecto a otro proyecto aún cuando los objetivos del proyecto sean similares. Características muy específicas, como por ejemplo la naturaleza y calidad de los materiales originales, el presupuesto de que se dispone, y el tiempo asignado para la realización del proyecto tendrán un impacto significativo en las decisiones.

4) Los costos serán mucho más elevados de lo que uno espera.

Todos los proyectos, hasta los que se hayan planeado con el mayor cuidado, incluido el OCR nos depararán sorpresas. Puede que el software que se seleccionó de inicio no se comporte con los datos reales como se comportó en las pruebas. Se pueden enfrentar limitaciones de procesamiento en la fase de producción que quedaron ocultas durante la fase piloto. La expansión de un diccionario de la aplicación para que incluya términos especializados puede resultar más difícil que lo que se anticipó de inicio. Muchos acontecimientos inesperados pueden impactar los marcos de tiempo de producción y por lo tanto los presupuestos. Sería muy conveniente poder contar con un margen que se incluyera desde el inicio del proyecto para enfrentar estos posibles acontecimientos inesperados.

5) La respuesta que es correcta para hoy, puede no ser la correcta para el futuro.

Las capacidades de softwares de OCR se han desarrollado de manera significativa en estos últimos cinco a diez años, y se siguen logrando avances. La naturaleza dinámica de esta tecnología quiere decir que los proyectos que sean de algo más que unos meses de duración se podrán beneficiar de la evaluación continua de los nuevos productos según se disponga de los mismos para determinar si se han desarrollado mejores posibilidades de beneficios de costos.

Bibliografía

Rice, Stephen V., George Nagy, y Thomas A. Nartker. *Optical Character Recognition: An Illustrated Guide to the Frontier.* Kluwer Academic Press: parte de Kluwer International Series in Engineering and Computer Science Secs 501. 1999.

Hasta 1997, el Information Science Research Institute de la Universidad de Nevada, Las Vegas llevó a cabo evaluaciones anuales de productos seleccionados de OCR. Se puede obtener información sobre el programa Technology Assessment Program en el sitio http://www.isri.unlv.edu/info/technology/

Website: www.jstor.org. Los lectores podrán encontrar información sobre la misión y la historia de JSTOR, una descripción del contenido de la base de datos, e información sobre las instituciones que participan en el trabajo de JSTOR. También se dispone de una descripción de nuestro proceso de producción, información técnica de interés general, y un vínculo con una demostración de la base de datos.

4. Digitalización de mapas y otros documentos de gran tamaño

Janet Gertz
Columbia University Libraries

Los mapas de papel (y otros documentos de gran tamaño, como los planos arquitectónicos) contienen una riqueza infinita de detalles compuestos por elementos gráficos y textuales. Los mismos incluyen:

- El dibujo de la localidad o locación
- El uso de entidades gráficas como líneas de elevación o símbolos para ciudades de diferentes tamaños
- Nombres de países impresos y otras características
- Color, que transmite información a través de diversos patrones e intensidades.

Cuando son grandes, los mapas con muchos detalles presentan dificultades especiales para la digitalización. Puede haber una gran disparidad entre el tamaño del documento y el tamaño del elemento con significación más pequeño que hay que hacer visible online o en impresos. Los detalles muy pequeños exigen escaneo de alta resolución y el resultado es un archivo de tamaño muy grande. La manipulación de los archivos, el almacenamiento, la entrega, y la presentación se hacen mucho más complicado.

Incluso se afectan los mecanismos de escaneo. Muchos escáners planos tienen limitaciones de tamaño y no pueden manipular grandes mapas. Para escanear estos materiales puede ser necesario trabajar con películas como intermediario, como por ejemplo transparencias de 4x5 o microfichas de un solo cuadro, en las que el original llena el cuerpo de la microficha. Las diapositivas de 35mm son demasiado pequeñas para capturar en plenitud todos los detalles que tienen los mapas grandes. Cuando los originales no sólo son de gran tamaño, sino que además están muy quebradizos, trabajar con una película como intermediario permitirá que se ejerza menos tensión sobre el original frágil. Se producirá una cierta pérdida de calidad debido a que la versión en película es una generación nueva con respecto al original. Sin embargo, se pueden producir imágenes completamente legibles a partir de intermediarios en película, siempre y cuando, por supuesto, la transparencia o la microficha se realicen con gran cuidado y minuciosidad y después se escaneen con suficiente resolución y tonalidad apropiada.

Parámetros de escaneo

La determinación de los parámetros de captura sigue las mismas reglas que para el resto de los documentos.

- Tomar una decisión con respecto a la tonalidad apropiada, casi siempre la escala de los grises o color. El color en la mayoría de los mapas impresos es importante como recurso de codificación, y no por su apariencia específica, como en las obras de arte. De todas maneras, deberá incluirse una barra de colores estándar durante el proceso de escaneo, aunque no sea un requisito el manejo sofisticado del color. (Para un mayor análisis, véase Ester, 1996.)

- Identificar el elemento significativo más pequeño, con frecuencia una línea muy fina.

- Determinar cuántos pixels se necesitan para capturar el elemento legible más pequeño.

- Calcular la resolución necesaria, con frecuencia 200-300 puntos por pulgada cuando se escanea a 24-BIT en color. Para un análisis sobre la resolución y otros aspectos relacionados, véase Gertz *et al.* (1996) y Allen (1998).

- Decidir si el mapa se escanea directamente o a través de un intermediario en película, incluir siempre una regla en la imagen para que las dimensiones y las distancias no queden ambiguas.

Como ejemplo, considere un mapa hipotético de 60cm de ancho y 90cm de largo.

- Los elementos textuales más pequeños son números de menos de 1mm de altura que marcan las elevaciones.

- Los elementos significativos más pequeños son las líneas finas que se utilizan para indicar elevación.

- Diez colores diferentes sirven como códigos, en patrones de puntos, líneas paralelas verticales y horizontales, y otros recursos gráficos.

Una versión escaneada de calidad aceptable permitiría a los usuarios trabajar en la pantalla y con impresiones para:

- Leer el texto de 1mm

- Poder ver las líneas de elevación sin interrupción

- Distinguir perfectamente todos los patrones de códigos de color.

Suponga que se necesita un mínimo de 200 dpi y color a 24-BIT para lograr legibilidad del texto de 1mm y de las líneas y códigos de color.

Para un mapa de 36" ancho, 200 dpi se multiplica hasta 7,200 puntos a través de la superficie del mapa. Si se utiliza película intermediaria, la resolución efectiva debe calcularse también. Resolución efectiva se refiere a la resolución relativa al tamaño del documento original. De todas maneras, una transparencia necesita 7,200 puntos a través del mapa para capturar el mismo grado de detalle. El mapa en la transparencia quizás sólo tiene 4" de ancho. Hay que escanearlo a 1,800 dpi para obtener el mismo nivel de detalle.

Para calcular el tamaño del archivo, use la fórmula que aparece en Kenney y Chapman (1996), p. 20.

Fórmula:	(	altura	x	ancho	x	Profundidad de BIT	x	dpi²	)' 3	
Mapa original:	(	24"	x	36"	x	24	x	200²	)' 3	= 103,680,000
Transparencia:	(	2.667"	x	4"	x	24	x	1800²	)' 3	= 103,680,000

El resultado de digitalizar documentos de gran tamaño evidentemente es una serie de archivos muy grandes. Esto tiene determinadas implicaciones para el creador de la imagen en términos de almacenamiento, acceso y representación. Los archivos grandes ocupan una gran cantidad de espacio de almacenamiento. Debe disponerse de memoria suficiente para cargar y manipular las imágenes. También los archivos de respaldo, la creación de derivados y la transmisión de los archivos absorben una cantidad de tiempo significativa y de soportes de almacenamiento.

Problemas de acceso a los grandes mapas escaneados

La naturaleza misma de esos archivos también se traduce en problemas para los usuarios que tratan de acceder y navegar en las imágenes digitales de grandes mapas.

- La imagen de alta resolución en la que son visibles todos los detalles es demasiado grande para que los usuarios accedan a ella o la manipulen con facilidad si tenemos en cuenta los actuales mecanismos de entrega y la capacidad de las computadoras comunes.

- Cuando se proporcionan derivados de los archivos originales de alta resolución para el acceso, con frecuencia son versiones JPEG cuya resolución se ha reducido de forma considerable. Si la resolución es lo suficientemente baja como para que sea fácil acceder a los archivos, puede ser que los detalles más pequeños sean ilegibles.

- Sólo una parte del mapa cabe en la pantalla. Cuando se utiliza el documento en soporte de papel, los lectores se orientan por las características más sobresalientes a través de la visión periférica a la vez

que se concentran en los detalles más pequeños. En la pantalla es más fácil desorientarse porque la mayor parte de la imagen no está visible todo el tiempo.

- Con un mapa en soporte de papel, de una sola mirada se pueden seguir hitos o características importantes, como por ejemplo carreteras o fronteras desde un extremo a otro, mientras que en la pantalla se necesita correr la pantalla continuamente. La comparación de detalles que están muy separados se hace incómoda en el mejor de los casos si todos no están visibles a la mima vez.

Los beneficios de escanear grandes mapas

A pesar de estas dificultades, existe una variedad de formas para que escanear grandes mapas se convierta en un beneficio.

- El uso de imágenes de alta resolución para producir versiones impresas de gran calidad que reemplacen lo originales muy deteriorados y quebradizos.
- Derivar versiones de menor resolución a partir de las imágenes master que sirvan como imágenes de referencia para la calidad y reducir la manipulación innecesaria de los originales deteriorados.
- Crear CDs con las imágenes y visionarlas directamente en lugar de entregarlas a través de una red. La estación de trabajo deberá ser capaz de manipular los grandes archivos.
- Escanear grande mapas por secciones para generar un grupo de archivos de alta resolución y de un tamaño manipulable. Esto implica el uso de paquetes de software para manipular los archivos por separado y concatenarlos según el usuario pasa de una imagen a la siguiente. Esto también puede complicar la creación de versiones impresas de alta resolución.
- Investigar algunos de los nuevos software de compactación que permitan al usuario acceder a una imagen de menor resolución y después ampliarla a una de mayor resolución sin manipular de manera local todo el archivo de alta resolución. Uno de esos productos es Lizardtech's Multi-Resolution Seamless Image Database (véase http://www.lizardtech.com); también hay otros paquetes de software disponibles.

En conclusión

- El tamaño del original, en proporción al tamaño del elemento significativo más pequeño, determina la resolución necesaria.
- El tamaño del archivo rige la habilidad de almacenar, acceder y presentar la imagen.

- Imágenes excelentes no satisfacerán a los usuarios si no se puede acceder a ellas o si el equipo y el software no son muy adecuados para manipular los archivos. La velocidad y la suavidad para correr la imagen y ampliarla también son importantes.

- Planificar la interfase del usuario tiene que ser parte del diseño inicial del proyecto.

Para estudiar los posibles enfoques para el tratamiento de mapas escaneados, véase:

- Proyecto Memoria Americana: http://memory.loc.gov/ammem/gmdhtml/gmddigit.html

- British Columbia Archives and Records Service: http://www.bcars.gs.gov.bc.ca/cartogr/general/maps/html

- Library of Virginia: http://image.vtls.com/BPW

- Universidad de Connecticut: http://magic.lib.uconn.edu/magic/exhibits/

- Atlantic Neptune: http://mercator.cogs.nscc.ns.ca/neptune.html

- National Oceanic and Atmospheric Administration: http://chartmaker.ncd.noaa.gov/ocs/text/MAP-COLL.htm/

- David Rumsey Associates: http://www.davidramsey.com

Bibliografía

Allen, David. "Creating and Distributing High Resolution Cartographic Images," *RLG DigiNews* 2:4, 1998. http://lyra.rlg.org/preserv/diginews/diginews2-4.html#feature

Ester, Michael. *Digital Image Collections: Issues and Practice.* Washington, DC: Comisión para Preservación y Acceso, 1996.

Gertz, Janet, Robert Cartolano, y Susan Klimley. *Oversize Color Images Project,* Columbia University, 1996. http://www.columbia.edu/dlc/nysmb/

Kenney, Anne, y Stephen Chapman. *Digital Imaging for Libraries and Archives.* Ithaca, NY: Departamento de preservación y conservación, Cornell University Library, 1996.

5. El trabajo con Microfilm

Paul Conway
Yale University Library

El microfilme de preservación puede ser un soporte-fuente excelente para los proyectos de conversión digital si se tienen en cuenta ciertas cuestiones. Esta sección describe lo que necesitan saber los archivistas y bibliotecarios acerca del trabajo con un microfilme que ya existe para producir imágenes digitales de alta calidad que se pueden exhibir o presentar como imágenes y/o procesar con un software de conversión OCR.

Antecedentes – El Proyecto "Libro Abierto"

El microfilm se ha utilizado como soporte para preservación y acceso desde la década de 1930. Ya a mediados de los años 1980, se contaba con normas internacionales que definían claramente la calidad de archivo del microfilm de preservación (Fox, 1996). El Grupo de Bibliotecas de Investigación (RLG), en estrecha colaboración con la American Library Association (ALA), estableció procedimientos para crear películas que cumplan o excedan las normas de archivo (Elkington, 1992). A finales de 1999, la National Endowment for the Humanities había proporcionado apoyo parcial para la preservación en soporte de microfilm de más de 800,000 volúmenes quebradizos y deteriorados. La colección de microfilm de preservación del país es la primera y una de las mayores *bibliotecas virtuales* en el mundo (Conway, Selecting,1996).

A principios de la década de 1990, Don Willis, uno de los más prominentes expertos en la creación de microfilm de preservación, propuso que era posible tanto desde el punto de vista tecnológico como económico, crear imágenes digitales de alta calidad à partir del microfilm (Willis, 1992). En el momento en que él escribía eso, pocas personas fuera del sector comercial – y ningún archivista o bibliotecario de Estados Unidos – estaba en posición de probar la hipótesis propuesta por Willis. La conversión de microfilm quedó confinada en su mayor parte a las corporaciones que necesitaban convertir los llamados *archivos legales* a partir de microfilm (casi todos, expedientes legales de casos judiciales y documentos normales de oficinas) y eso de manera muy selectiva. Lo que se necesitaba era una exploración sistemática de los problemas asociados con el trabajo de verter el contenido de cientos de miles de libros, periódicos y otras publicaciones periódicas en estado de gran deterioro y muy quebradizos que se encontraban preservados en microfilm de 35 milímetros y que ahora estaban depositados en bibliotecas de

investigación y archivos en todo el país. Si se demostraba que era factible obtener imágenes de alta calidad a un costo razonable a partir del corpus de película de preservación del país, entonces se podría añadir este material a lo que entonces se esperaba que podría constituir una fuente nacional de imágenes digitales.

Yale University Library, con la ayuda de la Comisión para Preservación y Acceso y de la National Endowment for the Humanities, aceptó el trabajo de desarrollar una secuencia de proyectos, que se denominaron de manera colectiva el Proyecto Libro Abierto, para demostrar la hipótesis de Don Willis (Waters, 1991). Yale diseñó y ejecutó el Proyecto Libro Abierto en estrecha colaboración con la Universidad de Cornell, que en ese momento estaba también trabajando muy duro en el proyecto de digitalización de R&D, utilizando libros como fuente principal de conversión. Yale adoptó las recomendaciones de Cornell para establecer una base de calidad de la imagen y después continuó y llevó a cabo un complejo estudio de costo para demostrar las presuposiciones económicas subyacentes en el proceso de digitalización de imágenes. El Proyecto Libro Abierto definió la relación entre calidad y costo. También estableció reglas básicas para elevar al máximo la calidad y establecer una base de estimados de costo para el proceso de conversión de microfilm (Conway, "Yale," 1996).

Después que el proyecto de Yale se concluyó, nuevos proyectos han contribuido al conocimiento general básico sobre cómo escanear microfilm. Además, varias empresas o burós de servicios comenzaron a ofrecer servicios de conversión a bibliotecas o archivos. Estas organizaciones comerciales son capaces de cumplir o exceder las expectativas de calidad sobre la base de un determinado costo por imagen que, aunque no es tan bajo como los que identificó y estableció el proyecto de Yale, son bastante eficaces. En 1999, los principales investigadores de los proyectos de Cornell y Yale reunieron sus conocimientos acerca del enfoque híbrido y desarrollaron una serie de recomendaciones para convertir libros quebradizos y deteriorados a partir del film o directamente del objeto original (Chapman, Conway & Kenney, 1999). En conjunto, estos acontecimientos hicieron posible recomendar las mejores prácticas para ciertos tipos de materiales en soporte fílmico e identificar cuándo el microfilm no es la mejor fuente.

Algunas consideraciones sobre la calidad de la imagen

La calidad de la imagen es el interés primordial. El microfilm de 35-mm de alto contraste, producido según las especificaciones ANSI/AIIM a un nivel de Índice de Calidad de 5 (en una escala de 1 a 8) tiene una resolución digital equivalente de al menos 800 puntos por pulgada (dpi). Todavía no es posible (y puede que no sea necesario) alcanzar este nivel

de escaneo a todo el ancho del cuadro de microfilm de 35 mm. Escanear a alta resolución a partir del microfilm varía de 300 a 600 dpi. La profundidad de BIT varía desde bitonal (1 BIT por píxel) hasta el gris completo (8 bits por píxel). No se dispone en la red comercial de escáners para película de rollo a color (ejemplares muy raros en las bibliotecas), aunque esa tecnología es una parte importante de la industria del cine (Kenney & Chapman, 1996).

Debido al alto riesgo de que se dañen, los negativos master de microfilm (1N) no deben utilizarse nunca como fuente para escanearlos. Las investigaciones realizadas en Yale y en Alemania han demostrado que el mismo nivel de calidad de imagen se puede obtener de un negativo duplicado (2N) sin poner en riesgo el negativo master (Weber, 1997). Si sólo se dispone de una copia positivo (3P), es posible obtener una imagen digital legible, aunque debe esperarse alguna disminución apreciable en la calidad de la imagen.

Las características del documento original y las características del soporte de microfilm ejercen una fuerte influencia en la calidad de las imágenes individuales y en el producto total de la imagen. A continuación algunos aspectos importantes.

Características de la fuente original (*Ej.*, libro, documento, impreso, mapa)

- Alto contraste entre el texto (tinta) y la superficie (papel) arroja mejores resultados
- Los bordes descoloridos, dañados y desiguales en el papel complican todo el proceso de escanear
- Sangramiento en el texto limita las opciones
- Moteado, hongos, manchas, y daños producidos por incendios y agua se acentúan cuando se escanea
- Márgenes muy estrechos en volúmenes encuadernados distorsionan la película y la imagen digitalizada a menos que se corrija el problema
- Los dobleces y los insertos de gran tamaño puede que no queden representados con mucha precisión en la versión digital como quedarían en un documento (se requieren modificaciones en la línea de producción de escaneo

Características del Microfilm

Calidad de la imagen

- Polaridad: el microfilm negativo arroja imágenes de mayor calidad que la película positiva
- Densidad: película negativa de contraste medio (dMax ca. .90) a alto contraste (dMax ca. 1.30) da como resultado imágenes de mayor calidad que la película de bajo contraste

(dMax ca. .80). Lineamiento RLG dMin (< .25).

- Tasa de reducción: menos es mejor; es de importancia crucial registrar la tasa de reducción para la reproducción a tamaño original
- Sesgo o inclinación: minimizar o eliminar – no más de 5 grados de la paralela

Calidad del producto

- Colocación consistente: minimice o elimine *la desviación de la línea central*
- Duplicar imágenes: Duplicar imágenes con ilustraciones aparte tiene poco impacto
- Cortes o divisiones: elimine las divisiones o separaciones en un mismo volumen o rollo
- Dimensiones del original: registre con exactitud en las metas bibliográficas
- Cuadros en blanco: elimine o reduzca la cantidad siempre que sea posible
- Orientación: La posición A2 ofrece un producto más consistente con algunos escáners. por lo general es preferible un cuadro por imagen.
- Cuadros de prueba: incorpore RIT Alphanumeric Test Chart como rutina en el trabajo con el escáner

El mínimo de calidad en el escaneado bitonal

- La naturaleza, la calidad y el valor de las ilustraciones complejas determinan si es adecuado el escaneado bitonal; si las ilustraciones son vitales y complejas, el escaneado bitonal puede no ser el apropiado.
- La consistencia precisa del texto (impreso o manuscrito) es esencial para la legibilidad de la imagen digital.
- No ocurren mejoras apreciables en la calidad de la imagen con película de tono continuo en el escaneado bitonal

Cuestiones de conversión de costos

Los costos de la digitalización de imágenes están controlados por la estructura de precio de los escáners, los costos de la mano de obra y la velocidad general del sistema de conversión. La velocidad de salida de un determinado escáner es producto de al menos tres factores:

- La resolución de la imagen (a menos resolución más velocidad de salida)

- La ingeniería eléctrica (una rápida tasa de refrescamiento de la disposición de CCD y una rápida tasa de transferencia de los datos es igual a una salida rápida)

- La ingeniería mecánica (mecanismos más rigurosos de transportación de la película proporcionan también una salida más rápida).

Resulta algo difícil comparar las velocidades de los escáners analizando las especificaciones de los fabricantes.

En su complejo estudio, el Proyecto Libro Abierto analizó los costos del proceso de digitalización en términos de equipamiento y recursos humanos (Conway, *D-Lib Magazine,* 1996). El modelo de costos tenía en cuenta los costos reales de hardware, software, soporte de integración, y medios ópticos y de almacenamiento y después convertía estos costos a costo por libro y costo por imagen. Y lo más importante, el estudio realizado por Yale evaluaba los costos según cada uno de los pasos de procesamiento en el proceso de conversión.

El estudio de Yale identificaba un número de factores que contribuyen a la variación de los costos, entre los que se incluyen los siguientes:

- El impacto de las características del original y el microfilm varía en los pasos del proceso.

- Los pasos de conversión que consumen más tiempo (escanear en modo continuo, indizar, establecer los parámetros del escáner y transferir los archivos) no están muy influenciados por los factores del original o por el microfilm.

- Las características del original influyen más en los costos que las características del microfilm.

- Las características de la fuente u original y del microfilm combinadas, tienen un impacto dramático en la calidad, pero sólo un impacto marginal en los costos.

- La inspección previa al escaneo del microfilm (un paso relativamente poco costoso en el proceso) constituye un mecanismo importante para predecir los retos del control de la calidad, pero no es suficiente para identificar complejidades significativas de escaneo e indización que puedan surgir durante el proceso de conversión.

Características de la fuente original (*Ej.*, libro, documento, impreso, mapa)

- Las características del original que tienen un gran impacto en la calidad (*Ej.*, texto descolorido, sangramiento) tienen poco impacto en el costo de la conversión digital.
- El número de páginas en el cuerpo del material que se escanea tiene un impacto financiero significativo sobre todos los procesos de conversión.
- Los libros que no tienen índice o tabla de contenidos o números de página imponen significativos retos de indización (y por tanto mayores costos), pero además también complican la inspección previa al escaneo y todos los aspectos de control de la calidad.
- La presencia de ilustraciones es sólo uno de los muchos factores que se combinan para explicar las variaciones de los costos de los pasos del proceso que consumen más tiempo.
- Los costos de los procesos de control de la calidad que se realizan durante el escaneo, la indización y la aceptación final están muy influenciados por las características del original (*Ej.*, márgenes muy estrechos y bisagras muy apretadas, texto cortado, ilustraciones).
- La preparación de un volumen encuadernado antes de la microfilmación (*Ej.*, desmontar la encuadernación, corte cuidadoso) puede reducir de manera significativa el costo de crear parámetros de escaneo.

Características del Microfilm

- La tasa de reducción es la característica más importante del microfilm que influye en los costos. Mientras menor es la tasa de reducción, menor es el costo de conversión.
- Las imágenes microfilmadas torcidas, factor demasiado común, aumentan el costo de escaneo, control de la calidad e inspección.
- Los cortes o divisiones dentro de un mismo volumen influyen en el costo de varios pasos importantes, pero son muy poco frecuentes para que se consideren demasiado.
- El costo por equipo o escáner no se afecta por ninguna de las características del microfilm.
- Las variaciones de densidad no tienen impacto sobre el costo de conversión.
- Invertir en microfilmes de mejor calidad sólo beneficia de forma marginal la reducción de los costos.

Empresas de servicios

Los proveedores pueden hacer el trabajo pesado. No es necesario comprar hardware y software para escanear microfilm cuando se ejecuta un proyecto dentro de la misma institución para lograr la conversión de un microfilm. En Estados Unidos existe una serie de compañías que prestan servicios de conversión, entre las que se incluyen:

- Preservation Resources of Bethlehem, PA <http://www.oclc.org/oclc/presres/index.htm>
- Northern Micrographics en La Crosse, WI <http://normicro.com>, y
- microMedia Imaging Systems, Inc. en Lake Success, NY.
- Las fuentes de información para este tipo de servicios son:
- *Imaging Magazine* <http://www.imagingmagazine.com> y la Association for Information and Image Management (AIIM) <http://www.aiim.org>. Es necesario ser miembro ($125 por individuo) para poder utilizar los servicios de la excelente biblioteca y servicios de referencia de AIIM.

Es muy importante probar los productos (*disponibles*) de un buró o empresa de servicios antes de finalizar un contrato. La mayoría de estos burós de servicios realizarán pruebas de escaneo de manera gratuita o por un precio módico como parte del proceso de licitación competitivo. Es responsabilidad suya entonces, especificar el nivel de calidad de las imágenes digitales en términos de resolución, rango dinámico (profundidad de BIT), y procesamiento posterior al escaneado (*Ej.,* eliminar el torcido, desmoteado y ajuste de tono). Es también responsabilidad suya especificar si es aceptable que el proveedor utilice un equipo que usa herramientas de resolución sintética para reducir las limitaciones de resolución del equipo. Por último, es también responsabilidad suya especificar las características de los archivos de salida en cuanto a formato del archivo, convenciones para nombrar los mismos y estructura de directorios, así como en cuanto a mecanismos de entrega (*Ej.,* CD-ROM, servidor FTP, cinta magnética).

Opciones de equipamiento

Si uno trabaja con un contratista para lograr los objetivos de digitalización, no será necesario adquirir los equipos para escanear. No obstante, se puede, y se debe, aprender lo más que se pueda acerca de las capacidades del equipo de escaneo y eso se puede hacer contactando a los fabricantes de sistemas de hardware y software.

Es necesario comprender las capacidades de hardware / software para desarrollar especificaciones de calidad y de costo, independientemente de si el proyecto se ejecuta en la propia biblioteca o institución. Los resultados del escaneo variarán en dependencia de la máquina utilizada, sin embargo, en dependencia de la forma en que el software para una máquina determinada defina los umbrales o límites (análogos a los parámetros de contraste en una fotocopiadora), se programan las diferentes opciones de los filtros, y aplican los diversos algoritmos que interpretan y ajustan la codificación de los píxels. Mientras más se conozca acerca de la forma en que el escáner interpreta y codifica lo que ve, mejores serán las imágenes resultantes . En última instancia, las especificaciones de calidad, las capacidades tecnológicas y las características visuales del original se combinan para determinar la calidad y los costos del producto imagen.

Las cinco compañías que aparecen a continuación fabrican o revenden cuatro sistemas para escanear microfilm en Estados Unidos. En general, el hardware y el software están unidos en un solo paquete. El nivel de personalización que puede especificar el comprador para el hardware o el software varía desde cero (Minolta) hasta extenso (Amitech). El nivel de control de usuario final sobre el equipo también varía mucho. Es importante ver y probar el equipo en condiciones reales antes de comprar equipo alguno. La mejor manera de acometer estas pruebas es pedir a las compañías de hardware una lista breve de referencias de clientes anteriores en el área y contactar esos clientes directamente.

Amitech Corporation <http://www.amitech.com>
5501 Backlick Road
Suite 200
Springfield, VA 22151
Teléfono: 703-256-2020 Fax: 703-256-9153
Amitech revende tres de los cuatro escáners de microfilm (Mekel, SunRise, Wicks & Wilson) que existen en el mercado y también ofrece una variedad de paquetes de software (personalizables) que controlan la operación del escáner y llevan a cabo varias tareas de manejo de datos posterior al escaneado (*Ej.,* eliminar el torcido, desmoteado y compactación).

Mekel Engineering, Inc. <http://www.mekel.com>
2800 Saturn Street, Suite B
Brea, CA 92821-6201
teléfono: 714-996-5600 Fax: 714-996-5696
El Mekel M500 es el mejor producto de conversión de microfilm de alta velocidad. Es capaz de manipular película de 35mm o de 16mm. El Mekel M560 es el hardware asociado para escanear fichas.

Minolta Corporation <www.minolta.com>
101 Williams Drive
Ramsey, NJ 07446
teléfono: 800-964-6658

Minolta fabrica el MS 3000 Microform Scanner, que puede manipular toda una serie de formatos si se cambia el mecanismo de transporte. El escáner es altamente automatizado y ofrece flexibilidad limitada al operador.

SunRise Imaging, Inc. <http://www.sunriseimg.com>
1250 N. Tustin
Anaheim, CA 92807
Teléfono: 714-632-2160 Fax: 714-632-2161
El SunRise ProScan III es el más complejo y completo escáner de microfilm en el mercado. Es capaz de convertir en modo bitonal y en escala de grises y puede manipular una serie de formatos en dependencia de la configuración de los mecanismos de apoyo de la película.

Wicks & Wilson, Inc. <http://www.amitech.com>
Morse Road Basingstoke
Hampshire RG226PQ England
Teléfono: 011441256842211
Las Wicks & Wilson 4000 y la 4001 Scanstations son los equipos llegados más recientemente al mercado de Estados Unidos. Se fabrican en Inglaterra por una compañía que se especializan en aplicaciones de digitalización de alta tecnología, como por ejemplo realidad virtual. En el momento de esta publicación, las máquinas WW sólo se pueden adquirir a través de Amitech. El fabricante afirma que las características clave son un escaneado de alta resolución y facilidad de uso.

Otras investigaciones necesarias

Es necesario investigar para certificar la factibilidad de convertir materiales que no son libros, especialmente manuscritos y periódicos. Además, los retos de trabajar con microfilm que no ha sido creado según normas rigurosas de archivo no se comprenden bien, entre los que se incluyen:

- Película vieja
- Película arañada o dañada
- Película de 16 mm
- Película de tono continuo
- Película de polaridad positiva
- Película de tercera generación.

Conclusión

En la pasada década, la tecnología de escaneado de microfilm ha madurado hasta el punto en que se dispone de opciones en lo referente a capacidades de hardware y software, así como opciones con respecto a la calidad del producto final y el costo de la tecnología. La calidad va en

aumento, los costos por imagen van en descenso. Hay que tener confianza en que la conversión a imagen digital de materiales fundamentalmente de textos a partir de microfilm de preservación es factible, tanto desde el punto de vista técnico como económico, y competitiva con otras tecnologías de conversión.

Bibliografía

Chapman, Stephen, Paul Conway, y Anne R. Kenney. *Digital Imaging and Preservation Microfilm: The Future of the Hybrid Approach for the Preservation of Brittle Books.* Washington, DC: Consejo de Bibliotecas y Recursos para la Información, 1999. [Online] Accesible en: http://www.clir.org/programs/cpa/hybridintro.html#description

Conway, Paul. "Selecting Microfilm for Digital Preservation: A Case Study from Project Open Book." *Library Resources & Technical Services* 40 (enero 1996): 67-77.

-----. "Yale University Library's Project Open Book: Preliminary Research Findings," *D-Lib Magazine* (febrero1996) [Online]. Accesible en: http://www.dlib.org/magazine.html

------. *Conversion of Microfilm to Digital Imagery: A Demonstration Project.* New Haven, CT: Yale University Library, 1996.

Conway, Paul y Shari Weaver. *The Setup Phase of Project Open Book.* Washington, DC: Comisión para preservación y acceso, junio 1994. [Online]. Accesible en: http://www.clir.org/pubs/reports/conway/conway.html.

Elkington, Nancy, ed. *RLG Preservation Microfilming Handbook.* Mountain View, CA: Grupo de Bibliotecas de Investigación, Inc., 1992.

Fox, Lisa, ed. *Preservation Microfilming: A Guide for Librarians & Archivists,* 2nd ed. Chicago: American Library Association, 1996.

Kenney, Anne R. y Stephen Chapman. *Digital Imaging for Libraries and Archives.* Ithaca, NY: Cornell University Library, 1996.

Watters, Donald J. *From Microfilm to Digital Imagery.* Washington, DC: Comisión para preservación y acceso, junio 1991.

Waters, Donald J. y Shari Weaver. *The Organizational Phase of Project Open Book.* Washington, DC: Comisión para preservación y acceso, septiembre 1992. http://www.clir.org/pubs/reports/openbook/openbook.html/openbook.html

Weber, Hartmut y Mariano Dora. *Digitization as a Method of Preservation? Final Report of a Working Group of the Deutsche Forschungsgemeinschaft.* Washington, DC y Amsterdam: Comisión para preservación y acceso y Comisión europea para preservación y acceso, 1997.

Willis, Don. *A Hybrid Systems Approach to Preservation of Printed Materials.* Washington, DC: Comisión para preservación y acceso, 1992. http://www.clir.org/pubs/reports/willis/index.html

6. La digitalización cooperada: escanear bien en cooperación con otros

Steven D. Smith
Amigos Library Services, Inc.

La tecnología de digitalización de imágenes puede ayudar a las bibliotecas, archivos y museos a alcanzar un alto nivel de cooperación que nunca antes fue posible. Las instituciones tradicionalmente han cooperado entre si para llenar vacíos en las colecciones locales – microfilmando archivos y ofreciéndolos en venta, proporcionando números faltantes de publicaciones periódicas o científicas, y lo más evidente, participando en préstamos ínter bibliotecarios. Sin embargo, la digitalización de imágenes ofrece la habilidad de crear colecciones virtuales de objetos que se encuentran en una serie de instituciones ubicadas en lugares geográficamente distantes. También permite la constitución de una red interfase única que permite a los investigadores el acceso a los materiales sin necesidad de preocuparse por la locación física de los mismos. Los proyectos cooperados con el uso de digitalización de imágenes pueden también vincular fuentes primarias con recursos secundarios para ofrecer a los usuarios una fuerte colección capaz de satisfacer las exigencias de todos excepto las investigaciones más profundas.

¿Qué es la Digitalización cooperada de imágenes?

La digitalización cooperada puede asumir una variedad de formas. En esencia, los proyectos cooperados han sido fundamentalmente varias instituciones que unen sus recursos para comprar una estación o estaciones de trabajo de digitalización para el uso de todos los participantes, o para utilizar ese poder adquisitivo agregado con el objetivo de garantizar menores costos de conversión por imagen al contratar los burós de servicios. Otra posibilidad es que las instituciones escaneen y coloquen las imágenes en la red de manera independiente, pero proporcionando un único punto de acceso para todas las colecciones (un ejemplo a gran escala es la Base de Datos de Imágenes Digitalizadas de la Asociación de Bibliotecas de Investigación o Association of Research Libraries [ARL] Digital Image Database http://www.arl.org/did/).

El tipo de cooperación que con más frecuencia se asocia con imágenes digitalizadas crea las colecciones virtuales descritas anteriormente. Entre los ejemplos se incluyen el proyecto Estudio en Escarlata del Grupo de

Bibliotecas de Investigación (http://www.rlg.org/scarlet/sis.html) y el Proyecto Memoria Americana de la Biblioteca del Congreso (http://memory.loc.gov/ammem/). En ambos casos, organizaciones importantes proporcionan la dirección y los lineamientos (e incluso financiamiento parcial), y se abrió a las bibliotecas y archivos de todo el país la contribución a las colecciones. Aunque estos ejemplos representan los esfuerzos de grandes bibliotecas de investigación, la actividad está abierta a bibliotecas, museos y sociedades históricas con colecciones de todos tipos y tamaños. De hecho las ventajas de la cooperación para las instituciones más pequeñas pueden ser mayores que para las grandes bibliotecas de investigación.

¿Para qué cooperar?

La razón principal y la más evidente para la cooperación es proporcionar a los usuarios mejor acceso a las colecciones. Pero hay otras razones adicionales que benefician a las instituciones mismas. La cooperación ofrece oportunidades para:

- Compartir el conocimiento experto
- Ahorrar en los costos de conversión
- Aumentar las oportunidades de financiamiento
- Incrementar la visibilidad de las colecciones al vincularlas con colecciones similares y con otras instituciones.

Quizás el aspecto que mejor vende los proyectos para las instituciones más pequeñas es la posibilidad y la habilidad de compartir las experiencias. Varias instituciones pueden trabajar de conjunto para resolver los problemas de convertir colecciones en soporte de papel y en soporte de película a formato digital y hacer una red de esas colecciones, junto con los problemas de catalogación y creación de metadatos.

¿En qué se diferencia la cooperación en proyectos de digitalización de imágenes de los proyectos cooperados de microfilmación?

La mayor diferencia entre los proyectos de microfilmación y los de digitalización es la complejidad. Además, hay procedimientos y normas establecidas para la microfilmación, mientras que todavía estamos aprendiendo acerca de los métodos óptimos para la digitalización (esa es

la razón de ser de esta publicación). El microfilme de preservación, aunque requiere la participación de seleccionadores y catalogadores, es en gran medida una tarea del personal que se encarga de la reconversión para la preservación. El conocimiento experto del seleccionador y el catalogador, así como la participación de los mismos es por supuesto necesaria, pero a estos tipos no se les pide que hagan nada fuera de lo corriente. La digitalización de imágenes requiere un nivel mayor de participación de las instituciones, e implica que se involucre más variedad de personal, especialmente la inclusión de personal de sistemas.

Además, los proyectos de digitalización no son tan fijos como los de microfilmación, en el que el producto final en esencia sólo se cataloga y se coloca en un estante. Los proyectos de digitalización no se terminan con la creación de las imágenes digitales y los metadatos asociados (cuestión compleja en si misma). Las cuestiones técnicas y administrativas de creación de una red y proporcionar acceso son una legión de cosas, y es necesario tenerlas en cuenta y resolverlas antes de que la primera página llegue al escáner.

- ¿Se podrá acceder a las imágenes por vía de Internet?
- ¿Cómo se va a manejar la cuestión de los derechos?
- ¿Se proporcionará además un texto buscable por computadora junto con las imágenes de objetos de texto?
- ¿Quién es el responsable de mantener el acceso a las imágenes?
- ¿Quién es el propietario de la colección agregada de imágenes digitales?

Pero cada proyecto de digitalización de imágenes puede resultar un producto final más útil que uno de microfilmación – un proyecto que permita acceso simultáneo a las colecciones a muchos usuarios.

Además de los beneficios arriba mencionados, los proyectos cooperados incrementan las oportunidades de obtener financiamiento externo, ya que muchas agencias que otorgan financiamientos han demostrado que sienten preferencia por proyectos coordinados con la participación de varias instituciones. En realidad, los proyectos cooperados pueden resultar menos costosos (*es decir,* de costos más eficaces) por imagen, ya que muchos de los costos relacionados con la digitalización de las imágenes no dependen tanto del número de imágenes o de participantes y, si el financiamiento es externo, el costo de conversión por imagen puede ser mucho menor.

Desde el punto de vista del usuario, es más probable que un proyecto cooperado produzca un producto final mejor, tanto en términos de contenido (utilizando los objetos más relevantes de varias colecciones)

como de forma (se beneficia por la experiencia compartida, el diseño de las bases de datos, el acceso intelectual, la interfase de la web, y así sucesivamente). Y esto se cumple más aún para las instituciones más pequeñas, en las que aunando los recursos o trabajando con instituciones más grandes, pueden lograr que sus colecciones sean más útiles para los investigadores.

Algunas consideraciones a modo de conclusión -- o Cómo NO cooperar

Existen muchos ejemplos de proyectos cooperados de digitalización de imágenes: Estudio en Escarlata, Memoria Americana, el Proyecto de Digitalización de Colorado (http://coloradodigital.coalliance.org/), y los diversos Making of America (véase, por ejemplo, http://moa.cit.cornell.edu/MOA/moa-mission.html y http://sunsite.berkeley.edu/moa2/). Mucho se puede aprender de estos ejemplos, pero también vale la pena tener en cuenta aquellos proyectos que no han logrado despegar – o pasar de la planificación a la ejecución.

Las causas más comunes del fracaso incluyen la resistencia a comprometer recursos (especialmente personal, y sobre todo personal con experiencia técnica), el deseo de esperar a que aparezcan normas industriales antes de seguir adelante, y el fracaso al definir los objetivos del proyecto. La digitalización de imágenes es una actividad de recursos intensivos y no se puede acometer sin comprometer el personal. Esperar a que aparezcan las normas no es razón para dudar. Aunque existen pocas normas referentes a la digitalización, debe reconocerse que al menos están apareciendo las mejores prácticas y otros lineamientos.

Especialmente difícil en los proyectos cooperados es el último aspecto: desarrollar firmes objetivos de proyecto. Muchas instituciones están interesadas en acometer la digitalización porque es una actividad de moda. Sólo son capaces de mencionar los objetivos del proyecto de manera ambigua y vaga y en términos no mensurables relacionados con el mejoramiento del acceso. Para que un proyecto tenga éxito, tiene que tener objetivos firmes y cuantificables.

En un nivel más amplio, los proyectos cooperados de digitalización fracasan porque los planificadores tienen que establecer de manera independiente qué papel desempeñará la digitalización dentro de su propia institución. Las instituciones necesitan confrontar las complejidades de las miles de cuestiones que plantea la digitalización de imágenes – la creación de una red, las aplicaciones de metadatos, la creación y el mantenimiento de la base de datos, derechos, servicios de

referencia para los usuarios de la red – antes de que se planifique un proyecto. Aunque es poco probable que dichas cuestiones se puedan resolver antes de que pueda comenzar un proyecto, es necesario que todas las partes las entiendan antes de continuar adelante.

Cuando todo está dicho y hecho, un proyecto cooperado puede parecer que complica una actividad ya de por si complicada. Pero la cooperación ofrece la considerable ventaja de reunir un gran número de expertos con conocimientos y experiencias mayores y más variados que una sola institución pudiera nunca reunir, lo que aumenta las oportunidades de éxito.

Un recordatorio final: aunque los ejemplos arriba citados involucran algunas de las más grandes e importantes bibliotecas del país, la digitalización cooperada está abierta a instituciones de cualquier tamaño. De hecho, las instituciones más pequeñas tienen mucho que ganar en la cooperación y también mucho que ofrecer. En la microfilmación de libros deteriorados y quebradizos, con frecuencia la mejor copia, o la única se encuentra fuera de las bibliotecas de investigación participantes en el proyecto. Lo mismo sucede con los proyectos de digitalización. Las bibliotecas pequeñas o de mediano tamaño, así como las bibliotecas especializadas y también las bibliotecas de museos o las bibliotecas médicas, tienen mucho que contribuir cuando se trata de archivos, manuscritos, fotografías, publicaciones seriadas y otros materiales interesantes. Se puede decidir la participación por proximidad geográfica, por tipo de biblioteca, o simplemente por la unión con otras instituciones con colecciones similares o afines, pero todas las instituciones pueden participar en una actividad que une las colecciones hasta un punto que nunca antes había sido posible.

Resumen de aspectos clave

- Defina el alcance del proyecto, incluyendo las colecciones más apropiadas y el nivel de indización.
- Defina los papeles respectivos de las instituciones participantes.
- Defina las áreas de responsabilidad de cada institución.
- Establezca objetivos mensurables para evaluar el éxito del proyecto a su culminación.
- Llegue a un acuerdo en cuanto al mantenimiento a largo plazo de las imágenes digitales y los metadatos asociados.

Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Telephone: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

VIII

Relaciones con los proveedores

Janet Gertz
Columbia University Libraries

Introducción

La tarea de conversión a formato digital se puede realizar en la propia institución o contratándola a proveedores del servicio. Independientemente de si la digitalización se realiza con objetivos de preservación, es una pérdida de tiempo y de dinero realizar una tarea de mala calidad. La versión digital de un documento debe ser plenamente funcional. Si lo que se digitaliza es ilegible o tan mal indizado que los usuarios finales no pueden encontrar lo que necesitan o leerlo cuando lo encuentran, el proyecto es un fracaso pues no se proporciona ni preservación ni acceso.

Para que la digitalización sea un éxito, es esencial que la institución comprenda bien claro cuál es el objetivo de la digitalización y qué tipo de producto final servirá ese objetivo. Comprender por qué se acomete un proyecto nos dará las pautas para tomar decisiones, no sólo en cuanto a la calidad de la imagen y las interfaces de usuario, sino también en cuanto a qué parte del trabajo deberá realizarse en la propia institución y qué parte se puede contratar con seguridad. Es importante:

- Involucrar a todos los participantes que deben estarlo (curadores, expertos técnicos, funcionarios de preservación) para determinar los objetivos del proyecto y tomar las decisiones que le darán forma

- Mantener un registro cuidadoso de todas las decisiones que se adopten y por qué, para evitar reinventar la rueda cuando surjan problemas
- Documentar totalmente cómo y por qué se llevó a cabo y se logró el trabajo, tanto dentro de la propia institución como por el proveedor, con el objetivo de facilitar la preservación en el futuro de los recursos digitales como tal.

Siempre habrá en toda operación digital un componente que involucra a la institución. La institución que posee el material o los materiales que se van a digitalizar debe hacerse responsable de:

- Seleccionar los materiales que se van a convertir a formato digital
- Determinar el propósito de la digitalización y la naturaleza del producto que se desea obtener
- Establecer los niveles de calidad necesarios
- Verificar la calidad del trabajo terminado.

Hay opiniones y argumentos a favor de que el trabajo se realice todo en la propia institución, y argumentos a favor de la contratación de empresas de servicios. La diferencia radica en:

- El grado de control inmediato sobre el trabajo
- La variedad de actividades que se pueden realizar
- La eficiencia
- El factor económico.

Tanto los bibliotecarios como los archivistas están aprendiendo todavía cuáles deben ser los parámetros del escaneo de alta calidad. Existen pocos lineamientos, el vocabulario todavía no es del dominio ni del conocimiento de toda la industria, ni tampoco todos los archivistas y bibliotecarios están familiarizados con el mismo. Las instituciones tienen por fuerza que experimentar, según van avanzando por el camino, y la curva de aprendizaje, tanto para las instituciones como para los potenciales proveedores del servicio es bastante pendiente. Relativamente pocos proveedores de servicios de escaneo trabajan con archivos y bibliotecas. La mayoría trabajan casi siempre con organizaciones que desean obtener un servicio de escaneo rápido y barato y conocen muy poco, si algo, de preservación y de la importancia de imágenes de alta calidad, alta resolución o de metadatos. Por fortuna, la situación mejora cada día, y ya existe una serie de proveedores de servicios con experiencia considerable que están dispuestos a compartir su conocimiento experto técnico.

Las mejores prácticas que existen, así como las recomendaciones sobre preservación, calidad, resolución y tonalidad, basadas en el tamaño del menor elemento significativo del documento que se va a escanear, ya se han explicado en capítulos anteriores. También en esos capítulos se ha puesto bien en claro que cada cambio de género, formato y soporte a digitalizar introduce una serie de decisiones en cuanto a la resolución, la tonalidad, los metadatos, el soporte de almacenamiento y la presentación al usuario. No hay una respuesta simple ni una respuesta única, y las interacciones entre los diferentes factores pueden tener repercusiones imprevistas durante la ejecución del proyecto. Puede que sea difícil determinar cuál es el mejor resultado para los propósitos de preservación y acceso, y qué serie de procedimientos y tecnologías nos proporcionarán el resultado final. La ausencia de normas y especificaciones claras, que se acepten en el ámbito nacional también pueden hacer más difícil explicar a los proveedores cuál es el nivel de calidad necesario para lograr resultados aceptables. Sin embargo, en el trabajo con proveedores, la institución debe ser capaz de establecer y explicar cuáles son sus requisitos mediante instrucciones claras, cuantificables y verificables, y debe ser capaz de reconocer si el producto que recibe del proveedor es lo que ha solicitado.

Para qué digitalizar en la propia institución

El argumento fundamental para justificar la digitalización en la propia institución es que ésta mantiene un estrecho control sobre todos los procedimientos, la manipulación de los materiales y la calidad de los productos. No hay necesidad de enviar materiales valiosos o frágiles a otro lugar y no hay que preocuparse por trabajar con un proveedor que resulta ser incompetente, y que al final nos entrega un producto que no fue el que se le solicitó o sencillamente deja de trabajar en la actividad por alguna razón.

Trabajar dentro de la propia institución es una buena manera de aprender la parte técnica de la digitalización a profundidad – útil incluso para el caso en que la mayor parte del trabajo que se realizará en el futuro se enviará a proveedores del servicio. Si no tenemos bien claro qué tipo de producto será mejor, trabajar dentro de la propia institución nos proporciona la vía para experimentar a pequeña escala sin necesidad de pasar por la redacción de contratos y especificaciones técnicas. Un pequeño proyecto piloto dentro de la propia institución puede servir

como preludio para contratar el grueso del trabajo. El trabajo en un proyecto piloto en la institución para probar una variedad de enfoques puede que no sea eficiente, pero es con frecuencia un paso necesario en el proceso de aprendizaje.

Trabajar en la institución es probablemente más exitoso cuando:

- El proyecto es relativamente de menor escala y se puede completar fácilmente dentro de límites de tiempo, o se puede dividir en segmentos más pequeños;
- La institución cuenta con personal calificado con un verdadero interés e incentivo para aprender y con el apoyo de la administración para que se produzca la capacitación;
- La institución ya cuenta con el equipamiento adecuado o con financiamiento para adquirirlo. (Pero, recuerde que tanto el equipamiento como el software se vuelven obsoletos con una rapidez terrible.)

Para qué utilizar proveedores

Los beneficios fundamentales del trabajo con los proveedores del servicio son financieros y técnicos.

- La institución no tiene que dedicar espacio para realizar el escaneo, ni tampoco tiene necesidad de convertir un espacio (que probablemente incluya trabajos de construcción) para resolver los requisitos técnicos eléctricos y de otro tipo.
- La institución no tiene que comprar constantemente los últimos equipos o softwares. El proveedor es responsable de mantenerse actualizado.
- La institución no tiene que ocuparse de contratar personal o adiestrar personal en técnicas sofisticadas, ni tampoco tiene que ocuparse del manejo del personal.
- El proveedor y no la institución es quien tiene que lidiar con las roturas y los desperfectos de los equipos, el tiempo sin trabajar, y la corrección de los errores.
- La institución se beneficia con la economía de escala y de alta productividad que el proveedor se ve obligado a poner en práctica.
- Por último, se establece un precio de antemano.

Muchas instituciones no poseen equipos de escaneo, y muy pocas cuentan con presupuesto suficiente para mantenerse al día con lo último en equipos y software que sale al mercado. Además, con frecuencia, no cuenta con personal que tenga mucha experiencia sobre escaneo y nivel de producción, y, (desafortunadamente), aprender a través de los errores para obtener esa experiencia lleva tiempo. Al menos en teoría, se puede esperar que los proveedores se mantengan actualizados en cuanto a lo último en hardware y software, y también deben tener personal altamente especializado y capacitado. También se supone que tengan una buena idea de qué servicios pueden ofrecer y de lo que cuesta proporcionar esos servicios. La otra cara de la moneda es que la institución está alejada del trabajo. Y se hace esencial que la institución mantenga un cuidadoso control de la calidad. Nadie debe dar por descontado que el trabajo se está realizando tal y como se especificó sin verificar mediante una inspección detallada las imágenes y archivos escaneados por el proveedor. Aunque el trabajo con los proveedores casi siempre implica el envío de los materiales – o de películas intermediarias – a un proveedor de servicios ubicado a cierta distancia de la institución, algunos proveedores pueden estar dispuestos a trasladar los equipos a la institución y a realizar el trabajo en la misma. Esto ofrece algunos de los beneficios del trabajo dentro de la institución, *por ejemplo,* no es necesario trasladar los originales, más control o supervisión sobre el trabajo del proveedor. Pero también incluye algunas de sus deficiencias, *por ejemplo,* la necesidad de proporcionar un área de trabajo adecuada, cuestiones de programación del trabajo diario, seguridad, y seguros.

Cómo seleccionar los servicios y los proveedores

Decididamente, no todos los proveedores son iguales. Incluso si utilizan los mismos equipos y los mismos métodos, algunos proveedores producen un producto mucho mejor que otros. Los buenos proveedores están interesados en aprender qué es importante para la satisfacción de las necesidades de la institución, mientras que otros pueden pensar que un proyecto es algo sin importancia y que no vale la pena esforzarse mucho. Y hay otros que quieren vender su propio sistema en lugar de actuar como servicio de conversión.

Identificar y seleccionar a un proveedor no es un proceso fácil ni rápido, sobre todo cuando se trata de proyectos grandes y complejos. Quizás las instituciones no necesiten seguir uno por uno todos los pasos que se

describen a continuación, pero sí todas tendrán que trabajar sobre una base.

- Desarrollar un concepto inicial del proyecto y sus objetivos.
- Identificar los proveedores potenciales.
- Enviar una SDI (solicitud de información) para explicar con claridad los objetivos del proyecto y para saber cuáles son los proveedores que están interesados y qué ideas tienen con respecto a la gestión del proyecto.
- Establecer la metodología del proyecto y los requisitos de calidad.
- Preparar una corta lista de proveedores del servicio.
- Preparar una SDP (solicitud de propuesta) y enviarla a todos los proveedores de la lista con muestras de los materiales a escanear.
- Comunicarse con los proveedores mientras ellos trabajan en la respuesta a las solicitudes, lo que incluirá visitas a sus instalaciones y entrevistas si fuera posible.
- Evaluar y comparar las propuestas de los proveedores y seleccionar la mejor.
- Preparar y firmar el contrato.
- Trabajar con el proveedor durante la ejecución del proyecto.

¿Cuáles son los objetivos del proyecto?

La institución deberá decidir qué quiere hacer y cómo quiere que se haga. Por ejemplo, ¿el objetivo de la digitalización es un índice visual de una colección de manuscritos, un servicio de lectura de reserva, una reproducción detallada de libros en malas condiciones por su fragilidad, o una combinación de objetivos de acceso y preservación? Una institución puede conocer qué quiere lograr en última instancia sin saber cómo lograrlo. En ese caso, la institución puede describir a los proveedores el producto final que desea obtener y pedirles que le propongan cómo lograrlo, dentro de líneas generales. Otras instituciones pueden saber al detalle los requisitos específicos del producto final en términos de resolución y tonalidad, tipos de archivos y soportes de almacenamiento, metadatos a registrar y posiblemente hasta el tipo de

equipo y de software que quisiera que se utilizara. En ese caso, sólo necesitan localizar a un proveedor capaz de satisfacer esos requisitos.

En todo caso, la institución debe ser capaz de expresar claramente al proveedor lo que necesita. Y esto puede ser un proceso interactivo:

- Una SDI que ofrece una descripción muy general de los objetivos a un número relativamente grande de proveedores
- Respuestas de algunos de los proveedores o de todos que proponen varios posibles enfoques
- Selección de uno o dos enfoques que al parecer son los más apropiados
- Una SDP en la que se ofrecen los enfoques aceptables en mucho más detalle que se envía a unos pocos proveedores de una lista
- Respuestas de los proveedores con información detallada acerca del procedimiento y cotización de precios.

La SDI – Solicitud de Información

El propósito de una SDI es ganar en ideas generales acerca de los posibles enfoques e identificar proveedores potenciales. Si la institución sabe lo que quiere como resultado final, pero no está segura de cuál es la mejor metodología / tecnología / software / metadatos, puede utilizar el proceso de la SDI para tener una idea de las posibilidades. Esto funciona especialmente bien cuando la institución contacta proveedores que tienen una experiencia significativa.

La SDI consta de:

- Una breve descripción del proyecto propuesto, que incluye cantidades, período de tiempo y resultados esperados
- Descripción de cualquier metodología que la institución tenga en mente
- Solicitud a los proveedores para que den sus opiniones acerca de la metodología
- Una invitación a los proveedores para que ofrezcan alternativas que conduzcan a los mismos resultados
- Solicitud a los proveedores para que indiquen si están interesados en optar por el proyecto.

Tiene sentido que con la SDI, se haga contacto con todos los proveedores que sea posible contactar y que podamos pensar de manera razonable que puedan estar interesados en el proyecto. Las respectivas

respuestas permitirán hacer una comparación entre los diversos enfoques, una evaluación de la calidad de escaneo en diferentes circunstancias, evaluar las presuposiciones iniciales acerca del producto contra los resultados actuales, así como dejar un margen de tiempo para hacer una revisión y para rediseñar el proyecto si fuera necesario.

Existen muchas maneras de identificar a los potenciales proveedores. Una de las mejores es pedir recomendaciones a colegas conocedores que hayan realizado algún proyecto o trabajo similar al que la institución pretende hacer. Informes sobre proyectos en ejecución o ya terminados pueden darnos algunas ideas. Asistir a Conferencias en las que los colegas analizan sus proyectos y los proveedores presentan sus servicios; buscar en Internet las páginas web de instituciones como la Biblioteca del Congreso, el Archivo Nacional, el Grupo de Bibliotecas de Investigación (RLG) y otras instituciones y organizaciones relacionadas con la preservación.

La Solicitud de Propuesta (SDP)

La solicitud de propuesta (SDP), a diferencia de la SDI, está diseñada para explicar en detalle a los proveedores potenciales cuáles son los requisitos y especificaciones del proyecto, los criterios que se utilizarán para evaluar sus propuestas y las especificidades de la forma en que deben presentar sus propuestas. Redactar una SDP y evaluar las respuestas puede ser algo complejo y que lleva tiempo. Los documentos de muestra que nos presentan el RLG y la Biblioteca del Congreso son bastante largos y complicados (más de 200 páginas en un caso). Aunque esto puede desalentar a las instituciones de que ni siquiera analicen la posibilidad de realizar el proyecto, vale la pena el esfuerzo de leer los documentos de muestra. Las instituciones pueden extraer de estos ejemplos los principios básicos necesarios para construir su propio documento más simple y les permite adaptar el lenguaje para que se ajusten a sus necesidades y requisitos en lugar de estar inventando de la nada. De todas maneras, es importante que se tenga en cuenta un margen de tiempo suficiente para redactar la SDP, para que los proveedores estudien el proyecto cuidadosamente y para que evalúen sus respuestas. Todo el proceso puede demorar varios meses. Lo bueno es que la SDP se puede transferir casi directamente al contrato.

Una vez que la institución ya tiene una buena idea de las especificaciones de la metodología del proyecto, ha llegado el momento de redactar una SDP y de enviarla a una pequeña lista de proveedores que son los que aparentemente es más probable que puedan realizar el proyecto. El objetivo de la misma es proporcionarle a los proveedores toda la

información para que ellos comprendan qué se desea hacer y puedan determinar un estimado razonable de los costos del trabajo. El valor del proceso de una SDP es que deriva respuestas explícitas que se pueden evaluar objetivamente para elegir al mejor proveedor. Además, especificaciones bien explícitas pueden ayudar a que la institución se evite tener que aceptar al menor postor si esa propuesta no satisface las especificaciones.

La SDP debe ser clara y explícita, con una descripción técnica especifica del producto a entregar e incluir cómo se va a evaluar el cumplimiento. La SDP debe ser lo suficientemente abarcadora como para que los diferentes proveedores propongan metodologías alternativas en los casos en que sea apropiado, pero a la vez lo suficientemente específica como para garantizar que los proveedores comprendan las normas y los niveles que tienen que satisfacer y cumplir. Es conveniente dividir la SDP en secciones que tratan de los requisitos técnicos, los requisitos de gestión, precios y referencias. Siempre que sea posible, pida a los proveedores que redacten sus respuestas en un formato estándar, para facilitar la comparación entre las diferentes propuestas.

El contenido de la SDP incluye:

- Una descripción del proyecto en cuanto a los objetivos en última instancia
- Una descripción de los objetos a escanear con el mayor detalle posible para ayudar al proveedor a hacer una propuesta inteligente
- Cantidad y naturaleza desde el punto de vista físico de los materiales y las dimensiones de los mismos
- Un estudio de las diferentes dimensiones de los materiales a escanear. Los materiales, ¿son bastante uniformes en general? Si en el proyecto se incluyen muchos géneros, describa cada grupo por separado.
- Una descripción de las proporciones. ¿Cuáles proporciones son fáciles y cuáles son difíciles para trabajar?
- Un estudio del contenido de los materiales. ¿Los materiales tienen una misma estructura intelectual que es necesario mantener? ¿Podrán los proveedores constituir lotes con los mismos tipos de materiales o es necesario manipular los materiales en un orden que mezcle diferentes tamaños y tipos?
- Tener en cuenta el idioma. ¿Qué idiomas es necesario que el proveedor conozca o pueda leer (texto o números de páginas) para poder ejecutar el proyecto?

- Instrucciones detalladas con respecto a la metodología que se prefiere utilizar, incluida la resolución, la tonalidad, la profundidad de BIT, formato de los archivos, compactación, plataforma y soporte de almacenamiento

- Instrucciones para producir imágenes derivadas así como imágenes master

- Definición del nivel de exactitud requerido y cómo lo evaluará la institución

- Instrucciones sobre cómo nombrar los archivos y los metadatos

- Cómo conformar los nombres de los archivos

- Si se incluirán números de identificación que ya existen u otra información

- Qué información sobre el procesamiento y los equipos hay que informar (*por ejemplo*, tipo de escáner utilizado, sus parámetros, definiciones de color utilizadas, fecha de captura de la imagen, descripción del tipo y lote, etc., de la película si se va a escanear a partir de intermediarios en película)

- Requisitos sobre cómo se deben codificar los datos y cómo deben presentarse, y qué tipo de hoja de cálculo o base de datos se debe utilizar

- Cronograma de entregas semanales / mensuales, fechas topes y cambios en el cronograma

- Manipulación (y niveles de iluminación si fueran cuestiones a tomar en cuenta), requisitos de seguridad, seguro y traslado de los originales

- Nombre de la persona en la institución que el proveedor deberá contactar en caso de que hubiera alguna pregunta y a quién se le debe enviar la oferta de servicios y en cuántas copias.

Pídale a los proveedores que:

- Le entregue una lista del hardware y el software que se utilizaría (¿Se aceptan archivos y datos en sistemas patentados?)

- Especifique los procedimientos de control de la calidad que utiliza

- Describa la capacidad de producción y que documente por escrito que pueden realizar y terminar el trabajo con la calidad especificada y dentro del marco de tiempo establecido

- Explique cómo se va a realizar la entrega de los materiales y los archivos (si el proveedor los recoge, si habrá un emisario o correo, UPS, u otra forma)

- Describa los controles ambientales en la instalación si eso constituye en problema para los originales.
- Informe el nombre y las calificaciones del administrador del proyecto
- Suministre referencias sobre trabajos similares realizados con otras bibliotecas, archivos o museos
- Escanee una muestra representativa que represente un corte adecuado de los materiales, que incluya objetos fáciles y objetos difíciles. (Si los originales están frágiles o son valiosos, la muestra debe constar de objetos razonablemente similares que sean menos valiosos o que se puedan sacrificar.)
- Su respuesta incluya una propuesta de precios
- Indique los precios en una unidad de medida específica, por ejemplo, por página, por imagen, o según sea adecuado
- Incluya los costos por introducción de datos, costo por soporte de almacenamiento, traslado, seguros y cualesquiera otros cargos adicionales
- Determine si los precios serán estables y firmes por la duración del proyecto
- Ofrezca sugerencias de métodos alternativos por los que se pueda lograr el proyecto con el mismo nivel de calidad.

La comunicación con los proveedores durante el proceso de SDP

Espere preguntas de los proveedores según ellos analizan la SDP. Algunas preguntas pueden refinar mejor el plan de trabajo.

En dependencia de las circunstancias, y sobre todo si los proveedores van a escanear materiales originales, invítelos a asistir a una reunión en la institución para que vean los materiales y a que participen en una sesión de preguntas y respuestas antes de que respondan a la solicitud.

Si es posible, visite las instalaciones de los proveedores para comprobar si cuentan con la capacidad y el personal para realizar el trabajo y si las instalaciones están limpias y bien administradas.

Evaluación de las respuestas de los proveedores

A la vez que redacta la SDP, la institución deberá preparar un plan para evaluar las respuestas y redactar los criterios por medio de los cuales se realizará una comparación precisa y objetiva de las capacidades de los proveedores para satisfacer los requisitos del proyecto. Algunos abogan por establecer un sistema numérico de categorización, en el que se le daría más peso a los aspectos más importantes de la propuesta. Se le podría dar una escala de 1-3 a cada factor a evaluar, y los factores de más importancia se multiplicarían por un coeficiente. De todas maneras, si se asigna un valor o categoría numérica o no, los factores más importantes (los que algunas veces se denominan *factores críticos para el éxito*) son los verdaderamente necesarios para que el proyecto tenga éxito. Por ejemplo, se puede necesitar la capacidad de proporcionar una base de datos para los metadatos (*crítica*), mientras que por otra parte se podría aceptar cualquier software de servicios.

Escriba cuáles son los criterios y las pautas para la evaluación de la calidad de la imagen y la exactitud de los metadatos. Determine cómo va a efectuar la evaluación de la calidad de la imagen, determine qué software, qué monitor y que impresora va a utilizar para ver la imagen; y decida quién va a realizar la evaluación.

Los criterios de valoración de las propuestas incluyen:

- Calidad de los productos y metodología técnica del proveedor
- Enfoque técnico general apropiado
- Habilidad del proveedor para producir trabajos de muestra que cumplan o excedan las especificaciones de la SDP
- Familiaridad con los lineamientos y mejores prácticas existentes
- Procedimientos de control de la calidad satisfactorios
- Capacidad de identificación de objetos inusuales, y de juicio para pedir instrucciones al respecto a la institución
- Habilidad para manipular materiales originales con seguridad y guardarlos también con seguridad
- Demostrada comprensión del alcance del trabajo y de los requisitos
- Pruebas claras de que el proveedor comprende el proyecto

- Habilidad para responder todas las cuestiones planteadas en la SDP en los términos requeridos. (Si no son capaces de seguir instrucciones, ¿serán capaces de realizar el trabajo de manera adecuada?
- Evidencias de su habilidad para llevar a cabo todo el proyecto
- Tamaño de la organización (¿La compañía es lo suficientemente grande como para realizar el proyecto o necesitará contratar personal nuevo y sin experiencia?)
- Habilidad para completar el proyecto dentro del marco de tiempo establecido
- Experiencia previa exitosa con trabajos similares
- Personal e instalaciones
- Una propuesta que sea profesional, tanto en su contenido como en su presentación
- Personal con actitud profesional de cooperación y orientada hacia el servicio
- Equipos modernos e instalaciones limpias y bien organizadas
- Dedicación al trabajo, compromiso con el mismo y con la relación a largo plazo que será necesario establecer para lograr que el proyecto sea un éxito
- Una base financiera sólida para la compañía
- Una propuesta cuyos costos sean razonables.

Haga la evaluación de la calidad y la capacidad y adecuación técnicas de las respuestas por separado a partir de la comparación de las propuestas de costos. Determine cuál de los proveedores satisface los criterios antes de determinar cuál ofrece los mejores precios. Elimine cualquier proveedor que no pueda cumplir los criterios técnicos y de calidad, independientemente del precio.

Compare las SDP y las muestras con todo cuidado utilizando los criterios establecidos. Discuta cualesquiera cuestiones inesperadas o insatisfactorias con el proveedor. Piense en la posibilidad de ofrecer una segunda oportunidad si algún resultado no satisfactorio se debió a un malentendido de los requisitos.

Llame a las referencias y conduzca discusiones profundas sobre la calidad, la actitud ante el servicio, el cumplimiento de los términos y otros factores del proveedor. Antes de llamar, prepare una serie de

preguntas para asegurarse de que no ha omitido ninguna cuestión importante.

Haga una lista por categoría o rango de los proveedores cuya calidad, flujo de trabajo, tecnología, e instalaciones satisfagan los criterios y necesidades del proyecto.

Compare las propuestas de precios para identificar al proveedor según la mejor combinación de alto éxito que haya alcanzado en la evaluación de las propuestas y muestras y bajos precios. Tenga cuidado de cualquier propuesta cuyo precio no sea realista, ya sea muy bajo o muy alto. Los proveedores conocen muy bien los precios de los otros proveedores y sus propios márgenes de ganancias. Los proveedores que proponen trabajos similares por lo general caen dentro de un rango de precios (aunque las compañías que son muy pequeñas o que son parte de una gran corporación pueden caer o bien hacia un extremo o hacia el otro). Si el precio de cualquier proveedor es significativamente bajo, sospeche. Una propuesta muy baja puede indicar un proveedor que machetea el trabajo o que no ha comprendido bien qué se necesita realmente para lograr el producto final. Precios muy bajos pueden ser un indicio de baja calidad a menos que ese proveedor sea mucho mayor y tenga mucha más experiencia que los otros, esté ubicado en una zona donde se pagan precios muy bajos o esté proponiendo una solución completamente distinta. Por otra parte, un precio muy elevado no es necesariamente garantía de mucha calidad – puede indicar simplemente que estén tratando de cobrar en exceso.

Si la propuesta más baja proviene de un proveedor que no es necesariamente el preferido, determine si ese mejor postor puede satisfacer las especificaciones basándose en muestras, referencias y cotizaciones de otros proyectos. Si la respuesta es negativa, tiene la base para rechazar la propuesta. Si ningún proveedor combina todos los factores deseados, analice posibles cambios con el proveedor(s) preferido para acercarse lo más posible a la desiderata.

El Contrato

En dependencia de la institución y de la complejidad del trabajo a realizar, se puede necesitar desde una simple carta de acuerdo hasta un contrato completo antes de que pueda comenzar el trabajo del proveedor en un proyecto. Las especificaciones planteadas en la SDP sirven como cuerpo del acuerdo, a lo que se añaden detalles logísticos, legales y

financieros. De más está decir que se recomienda la participación de las oficinas legales y financieras de la institución desde las primeras etapas del proceso.

Los contratos comienzan normalmente con secciones que abarcan las obligaciones legales de las dos partes y con una descripción del trabajo que cubre el contrato. Otros detalles, como por ejemplo, los procedimientos, se pueden anexar como apéndices. El contrato debe:

- Decir cuál es el objetivo del proyecto o qué se supone que sea el producto
- Describir brevemente qué cosas debe enviar la institución, o de qué es responsable, entre los que se incluyen los objetos originales, intermediarios en película, lista de nombres que se le van a dar a los archivos o un sistema para nombrar los archivos
- Describir brevemente qué acuerda hacer el proveedor, por ejemplo, producir imágenes digitales, llevar a cabo OCR, crear archivos de metadatos
- Establecer los términos legales que cubre el contrato
- Especificar los términos de aceptación del producto mediante la definición del nivel mínimo aceptable de exactitud
- Especificar cómo se definirán y rectificarán los errores, cuáles se rectificarán de manera gratuita y por cuáles pagará la institución
- Especificar en qué forma se van a transportar y manipular los materiales y los archivos digitales
- Especificar las cuestiones de seguros, seguridad y ambiente de almacenamiento mientras los materiales se encuentran en tránsito en las instalaciones del proveedor
- Especificar el nombre de la persona que será el contacto en primera instancia, y ajustar visitas periódicas a la instalación así como visitas del proveedor a la institución
- Declarar que los materiales y cualesquiera productos obtenidos de los mismos son propiedad de la institución y no se podrán utilizar o distribuir para ningún propósito sin la autorización oficial por escrito de la institución
- Especificar la frecuencia de los informes y de las facturas y qué información deberán contener las mismas
- Especificar las fechas tope y las penalidades por incumplimientos

- Definir qué constituirá incumplimiento, cómo disolver el contrato de manera amigable o sobre bases de acuerdo mutuo, y como se manejará el arbitraje

- Declarar los precios y las garantías de que los precios serán firmes por la duración del contrato o aumentarán sólo según condiciones específicas.

La segunda parte del contrato trata de las especificaciones técnicas, entre las que se incluye:

- Equipamiento y software que se va a utilizar

- Soporte de almacenamiento que se va a utilizar

- Especificaciones a seguir en cuanto a la resolución, la tonalidad, el formato de los archivos, la compactación, etc.

- La forma en que se van a introducir los nombres de los archivos y los metadatos

- Cómo llevará el proveedor el control de la calidad.

Los apéndices abarcan los detalles de:

- Descripciones profundas y detalladas de los materiales

- Cronograma, fecha tope de entrega y marcas

- Corrección de errores, manipulación de originales, traslado

- Muestras de modelos de trabajo

- Otras informaciones útiles.

Una vez que el contrato está redactado de manera satisfactoria para la institución, el procedimiento regular es enviar dos copias al proveedor. Probablemente, el proveedor sugerirá algunos cambios. Si como resultado de las negociaciones se producen cambios significativos, será necesario redactar una nueva versión. Si los cambios son sencillos y pequeños, se pueden rectificar en el mismo contrato y tanto la institución como el proveedor firman sus iniciales en el mismo texto. Una vez que el contrato está totalmente acordado, el proveedor firma ambas copias y las devuelve a la institución. La institución firma ambas copias y devuelve una copia al proveedor.

El trabajo y la comunicación con los proveedores

Además de una cuidadosa planificación de antemano, las claves para lograr un proyecto de éxito son la flexibilidad y la comunicación constante con el proveedor durante el proyecto. Mientras mejor sea la comunicación entre la institución y el proveedor, es más probable que el proyecto salga mejor. Es importante asignar la responsabilidad de la comunicación diaria a una persona que dentro de la institución esté muy estrechamente relacionada con el proyecto, incluso si esa persona se ve obligada a referir algunas cuestiones o preguntas a otras personas dentro de la institución. De la misma manera, es conveniente pedirle al proveedor que nombre a una persona para que sirva como contacto con la institución para el proyecto. Inste al proveedor a comunicarse con la institución de inmediato ya sea por teléfono o por fax o por e-mail cuando surja cualquier problema, y responda rápidamente. Los retrasos en la solución de un problema de menor envergadura pueden retrasar todo el proyecto. Si es posible, visite al proveedor durante la ejecución del proyecto e invite al proveedor también a visitar la institución.

La esencia de las buenas relaciones con el proveedor es ser justo con el mismo. Manténgase dentro del cronograma, o si surgen retrasos inevitables, informe al proveedor lo más pronto posible y esté preparado para cambiar todo el cronograma del proyecto. Los proveedores programan el trabajo que ellos contratan en un cronograma por lo general muy apretado. Si se producen retrasos por parte de la institución que desajustan su programa y sacan ese proyecto específico de su marco de tiempo asignado, el proveedor no puede retrasar el trabajo de otra institución para recuperar el tiempo perdido.

Manténgase dentro de los niveles de productividad ajustados en lo que respecta a las entregas al proveedor y evite cambios no advertidos en la naturaleza y calidad de los materiales. Por ejemplo, si el proveedor ha hecho su propuesta sobre la base de una muestra uniforme de materiales legibles y que no están en malas condiciones de conservación, por ejemplo quebradizos, la inclusión de una cantidad significativa de artículos en mal estado, quebradizos o de poco contraste puede afectar de manera muy seria el flujo de trabajo y la productividad anticipados. Un aumento o una disminución significativas de lo que envía la institución puede desestabilizar el flujo de trabajo del proveedor y debe negociarse. Asegúrese de que todos los materiales están preparados en la forma acordada antes de ser enviados al proveedor. Por último, identifique

todos los materiales de manera bien clara y consecuente para que el proveedor pueda identificar fácilmente y determinar cuál es cuál.

El trabajo con los proveedores: Control de la calidad y Correcciones de Manipulación

Otros capítulos describen cómo realizar el control de la calidad de las imágenes. La precisión de los nombres de los archivos y de los metadatos proporcionados por el proveedor también hay que verificarla, ya que metadatos o nombres de archivos erróneos significan que en esencia, la imagen se ha perdido. Examinar imagen por imagen con cuidado en busca de fallos es una operación que consume mucho tiempo, pero es esencial para garantizar que el producto del proveedor cumpla las especificaciones.

Cuando se trabaja con un proveedor con el que no estamos familiarizados, es de importancia especial llevar a cabo un control de la calidad meticuloso si es posible en las primeras etapas del proyecto. También la premura en la devolución de los errores es importante, ya que:

- Evita que el proveedor continúe repitiendo los mismos errores
- Alerta al proveedor de que existen problemas en los procedimientos o de que el personal no está bien entrenado
- Normalmente hay una fecha en la que se hace un corte (casi siempre después de varios meses) después de la cual el proveedor no acepta errores para rectificación gratuita; después de esa fecha será necesario volver a pagar el trabajo.

Cuando se está realizando un proyecto nuevo, siempre debe esperarse que se produzca algún malentendido. Surgirán situaciones inesperadas cuando la institución y el proveedor comiencen a trabajar con materiales que varían mucho más de lo anticipado. Espere que los primeros meses sean un período de adaptación o acomodamiento. Probablemente, en los primeros envíos siempre habrá algunos trabajos que habrá que rehacer. Pero una vez que se hayan ajustado los procedimientos y se hayan resuelto los problemas iniciales, debe esperarse que el proveedor satisfaga las especificaciones de la institución como cosa de rutina.

Si el proyecto es pequeño, puede ser posible examinar cada imagen, pero la mayoría de los proyectos son demasiado grandes. En ese caso, se recomiendan los siguientes procedimientos.

- Prepare un primer envío que sea más manejable y que deberá devolverse en las primeras etapas del proyecto.
- Lleve a cabo un control de la calidad meticuloso al 100% de las imágenes y metadatos del primer envío que le sea devuelto.
- Registre todos los errores en detalle en hojas de cálculo para el control de la calidad.
- Evalúe los errores para determinar cuáles son responsabilidad de la institución debido a errores en los procedimientos de la misma, a información que la institución no entregó al proveedor, o a variaciones en los materiales que el proveedor escaneó y de las que no se le advirtió. La institución tendrá que pagar para que se rectifiquen esos errores, y es evidente que tendrá que revisar los procedimientos para evitar nuevos errores en el futuro.
- Determine cuáles errores son responsabilidad del proveedor. Si el por ciento de errores es superior al acordado, devuelva todo el envío con una explicación detallada de los errores y exija al proveedor que comience de nuevo y produzca un nuevo lote. Según sea necesario, discuta los cambios que el proveedor tendrá que hacer en los procedimientos de captura de imagen y control de la calidad.
- Repita la inspección al 100%. Si el por ciento de errores sigue siendo muy elevado, devuélvalo de nuevo para que lo hagan todo otra vez. Si el proveedor no logra hacerlo correctamente en la tercera vez, puede que haya llegado el momento de renegociar todo el proyecto.
- Si el trabajo realizado por el proveedor cumple el por ciento de errores acordado o es menor, devuelva sólo los casos individuales para la corrección
- Continúe con la inspección del 100% en los dos o tres primeros envíos.
- Cuando está claro que el proveedor devuelve de manera regular un producto que está por debajo de la tasa o por ciento de errores aceptable, inspeccione un por ciento menor de cada envío (casi siempre el 10%). Esto quiere decir que algunos errores se irán sin detectar hasta que algún día algún usuario trate de acceder a esas imágenes.
- Si el por ciento de errores comienza a ascender, vuelva a inspeccionar el 100% hasta que se identifique y se resuelva el problema.

Manténgase en las definiciones acordadas para el control de la calidad. Si la institución decide que no le gusta la calidad del producto, pero el proveedor está cumpliendo las especificaciones acordadas, el cambio hacia niveles de calidad más elevado es una cuestión a negociar.

Bibliografía

Lineamientos de Solicitud de Propuesta para Digitalización Electrónica, ANSI/AIIM TR27-1996.

Muestra de documentos de SDI y SDP se pueden encontrar en: The Library of Congress. El sitio The American Memory Project Background Papers and Technical Information incluye tres SDP de bibliotecas digitales de la National Digital Library para servicios de escaneo y conversión de formato. Estos documentos son largos y complejos que incluyen todos los detalles y requisitos que exigen las agencias federales. http://memory.loc.gov/ammem/ftpfiles.html

Research Libraries Group (RLG), Preservation Program Tools for Digital Imaging, que ofrece una serie de documentos producidos por el departamento de Preservación para Bibliotecas de Investigación de la Universidad de Cornell. Incluye estimados de costos, lineamientos para crear una SDI y una SDP, y modelos de SDI y SDP. Mucho más simples que los de la Biblioteca del Congreso, los modelos son muy detallados. http://lyra.rlg.org/preserv/RLGtools.html

Buscar Contenidos Sitio del NEDCC

Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Telephone: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.

IX

Longevidad Digital

Howard Besser
University of California, Los Angeles
School of Education & Information Studies

Cuando contamos con una gran cantidad de recursos ya comprometidos a la reconversión a formato digital, necesitamos entonces analizar cómo podemos garantizar que la información digital continúe siendo accesible durante un período de tiempo prolongado. Este capítulo presenta primero el problema de cómo evitar que la información digital desaparezca. Una vez presentado ese problema, analiza detenidamente cinco factores clave que constituyen problemas para la longevidad digital. Por último dedicamos la atención a una serie de sugerencias con las cuales es muy probable que se pueda alargar la longevidad de la información en formato digital, concentrándonos en primer lugar en los metadatos. Este capítulo está dirigido a las comunidades de bibliotecas, museos y archivos. Sin embargo, las observaciones del mismo son útiles para todos aquellos que tengan interés en garantizar la longevidad de cualquier tipo de información digital.

La corta vida de la información digital

Aunque el advenimiento del almacenamiento por medios electrónicos es relativamente algo reciente, hay una gran cantidad de información almacenada en formato electrónico se ha deteriorado y ha desaparecido. Por ejemplo, los archivos en cinta de video o de audio, como las entrevistas relativamente recientes diseñadas para capturar los últimos restos culturales de los ancianos de las tribus Navajo, puede que ya no sean salvables (Sanders, 1997).

La mayoría de las personas tiende a pensar que (a diferencia de la información en formato analógico) la información digital será eterna, y

no comprenden la fragilidad de las obras en formato digital. Muchísima información en formato digital (por ejemplo una parte considerable de la misión Viking a Marte) se ha perdido debido al deterioro de las cintas magnéticas en las que esa información es residente. Pero el problema del deterioro de los soportes de almacenamiento palidece si se le compara con el problema que representa el rápido cambio de los dispositivos de almacenamiento y los cambios de formato de los ficheros o archivos. Hoy en día es casi imposible leer archivos de discos floppy de 8 pulgadas que eran tan populares hace 20 años, y tratar de decodificar los archivos en WordStar que se utilizaba hace sólo diez o doce años puede convertirse en una pesadilla. Grandes cantidades de información en soporte digital de hace sólo veinte años están, desde todo punto de vista práctico, perdidas.

Para evitar más pérdidas, necesitamos abordar y llegar a algún tipo de solución con el problema de la longevidad en el mundo digital. Necesitamos entender en qué forma difiere la preservación en el mundo digital de la preservación a la que estamos acostumbrados en el mundo analógico. En el mundo analógico, todos nuestros esfuerzos para preservar una obra se concentran en esa obra como artefacto. Pero según nos vamos metiendo en el tema de la preservación en formato digital, es necesario que demos un salto conceptual de lo que es preservar un objeto físico a preservar un contenido informativo que puede diferir completamente de cualquier artefacto físico.

Las siguientes secciones abordan cinco factores claves que presentan los problemas de la longevidad digital: el Problema de la Visualización, el Problema de la Decodificación, el Problema de la Inter-Relación, el Problema de la Custodia, y el Problema de la Traducción.

El problema de la visualización

La información digital creada en el pasado requiere el mantenimiento de una infraestructura y una base de conocimientos para poder visualizarla. Por ejemplo, para visualizar un archivo en un procesador de texto antiguo, se necesita un software que comprenda los sistemas de codificación del software original y pueda mostrar el archivo de manera adecuada en la pantalla. Sin esto, lo único que lograremos ver será cascarilla. Pero para poder mantener vivos estos ficheros con el tiempo, también necesitamos mantener el software para correrlo o conocimientos sobre el esquema de codificación, y debemos poder producir el software que utiliza el esquema de codificación para poder mostrar de manera adecuada los archivos digitales en la pantalla.

El problema de la información digital no es sobrevivir...

En el mundo analógico, los formatos anteriores persistieron con el tiempo. Las tablillas de escritura cuneiforme, los papiros y los libros existieron hasta que alguien o algo (incendios, terremotos) actuó para destruirlos. Pero el defecto de la información digital es que no sobrevive a menos que alguien adopte medidas conscientes para hacerlas persistir. Con frecuencia, en el pasado, hemos encontrado manuscritos antiguos o libros abandonados u ocultos en sótanos o áticos. Los archivos de procesadores de textos de la actualidad que encontremos en sótanos o áticos en el futuro no serán legibles a menos que sus autores tomen medidas concretas para hacerlos persistir. Incluso aunque podamos leer discos floppy que encontremos y descubramos que contienen archivos, es probable que no podamos descifrar esos ficheros y visualizarlos de manera adecuada.

Cuando descubrimos obras antiguas en formato analógico, al menos podemos visualizarlas tanto a ellas como su estructura, incluso si hemos perdido la habilidad de descodificar el lenguaje en que están escritas. Y el descubrimiento posterior de obras como la Piedra Rossetta nos permite descodificar su estructura y significado. De igual manera, cuando descubrimos películas viejas (ya sean imágenes fijas o en movimiento), incluso aunque no tengamos a nuestra disposición el proyector adecuado para ese formato, podemos sostenerlas a la luz y ver qué es lo que hay en ellas.

La información digital requiere una serie elaborada de conocimientos y/o ambiente de computación para poder descifrarla. Por lo general, la información está codificada y para verla se necesitan aplicaciones de software que corran en un sistema operativo en particular y que necesita una plataforma de hardware en particular. Además, la información por lo general se almacena en un soporte, o medio, o dispositivo físico (por ejemplo un disco duro, un disco floppy, o CD-ROM) que necesita un tipo de driver específico conectado a un tipo de computadora especial.

Estamos creando una Torre de Babel en lo que respecta a la proliferación de combinaciones que se necesitan para visualizar un archivo.

Cada pieza de esa infraestructura cambia a una velocidad increíble – de forma tal que permite a la industria de la computación venderle el mismo tipo de producto a la misma persona repetidas veces (porque el

individuo supuestamente necesita una versión más rápida o más actualizada). Los rápidos cambios que se producen en las versiones de hardware y software constituyen un verdadero dolor de cabeza para los interesados en la longevidad digital. Esto incluye problemas con los formatos de los archivos, los dispositivos o soportes de almacenamiento, los sistemas operativos y los hardware.

La mayoría de los procesadores de texto de la actualidad no pueden leer archivos creados con procesadores de texto más viejos. La mayor parte de las organizaciones tienen problemas incluso para abrir ficheros creados con los procesadores de texto más populares de hace sólo diez o doce años (WordStar). De hecho, los procesadores de texto más populares en la actualidad (como el Microsoft Word) no pueden leer archivos creados con versiones anteriores del mismo procesador de texto (y muchas veces sólo pueden leer archivos creados en las dos versiones más recientes). ¿Cómo podemos siquiera pensar que los archivos que creamos hoy puedan ser legibles en nuestros ambientes de información dentro de 100 años?

Cuando los procesadores de texto de la actualidad son capaces de abrir los ficheros de las versiones más recientes, con frecuencia estos ficheros pierden su formato. Las fuentes o tipos de letras, los subrayados, los márgenes y los formatos se pierden o desaparecen. Pero al menos nuestros procesadores de textos más antiguos son en principio textos en código ASCII entremezclado con comandos para formato. Y por lo menos cuando intentamos recuperar o resucitar esos ficheros tenemos alguna esperanza de encontrar palabras o frases contenidas en los mismos. Sin embargo, en el caso de los archivos que no se basan en texto en código ASCII (como los archivos en formato multimedia), hay pocas esperanzas de que los arqueólogos de dentro de cien años puedan descifrar algo en esos archivos o formatos. Los formatos como TIFF, AVI, las diversas versiones de MPEG, y otros presentarán incluso problemas de longevidad mayores y más complejos que los archivos en procesadores de texto.

Los dispositivos de almacenamientos, que también cambian con mucha rapidez, también constituyen problemas para el futuro. En menos de 20 años hemos pasado por diversos dispositivos de almacenamiento transportables, entre los que se incluyen: floppies de 8", floppies de 5.25", floppies de 3.5", CD-ROMs, y DVDs. Con los incrementos en la densidad de almacenamiento, existen pocas esperanzas de que el movimiento hacia nuevos dispositivos de almacenamiento se detenga en algún momento en el futuro próximo. En la actualidad, cuando descubrimos un floppy de 8", primero tenemos que buscar y encontrar el drive o torre de 8" adecuado, conectarlo a una computadora y a un

sistema operativo que tenga el driver adecuado y lo pueda leer, y después de todo eso, todavía enfrentamos los problemas explicados anteriormente para poder descifrar el archivo. Con los cambios de los sistemas operativos (CP/M, MS DOS, Windows, Windows 95, Windows NT, Windows 2000) y los cambios en las plataformas de hardware (8088, 8086, 286, 386, 486, Pentium, Pentium II, Pentium III), estamos creando literalmente una Torre de Babel en lo que respecta a las combinaciones que se necesitan para visualizar un archivo.

Aunque la longevidad digital parezca necesitarlo, ¿cómo podemos siquiera esperar que podamos enfrentar todas esas permutaciones y combinaciones? Piensen en todos los formatos que tendríamos que salvar, de todos los esfuerzos que necesitaremos realizar para descifrar los archivos que existen en la actualidad.

El problema de la decodificación

Para poder resolver los problemas que son el resultado del uso de la tecnología digital a corto plazo, nos hemos involucrado en prácticas que pueden resultar a su vez en un peligro a largo plazo. Dos ejemplos dignos de destacar son la manera en que hemos enfrentado las limitaciones de almacenamiento y el comercio digital.

En el pasado, debido a que el almacenamiento a gran escala era costoso y el ancho de banda relativamente estrecho, muchos depositarios respondieron a estas limitaciones compactando sus imágenes master o multimedia. De acuerdo con el razonamiento que dominaba hasta hace poco, los archivos master compactados ocupan menos espacio de almacenamiento, es más fácil entregarlas a los usuarios a través de conexiones más lentas, y también se manejan internamente de manera más conveniente. En años recientes, una serie de instituciones han comenzado a cuestionarse esta premisa, ya que los costos de almacenamiento han caído en picada y las velocidades de las redes han aumentado de manera dramática. Sin embargo, todavía en muchas instituciones persiste la creencia de que se debe compactar hasta los archivos master.

La compactación crea una serie de problemas. En primer lugar, todavía no comprendemos bien los efectos a largo plazo de la compactación. La compactación puede ser *con pérdidas* o *sin pérdidas*. Por definición, cuando una imagen compactada sin pérdidas se descompacta, esta es idéntica a la imagen antes de que fuera compactada. Pero cuando se descompacta una imagen compactada con pérdidas, la mima es diferente de la imagen original porque se eliminó parte de la información como parte del proceso de compactación. Los formatos comunes de

compactación con pérdidas, por ejemplo el JPEG, en esencia tratan de desechar información que el ojo humano no puede distinguir bien (por ejemplo combina colores que están muy cerca uno de otro; elimina los rangos espectrales fuera del alcance de la percepción humana). Pero todavía no comprendemos si algunos de estos datos eliminados serán útiles para aplicaciones futuras que utilizarán la visión de las máquinas (en lugar de la humana) – aplicaciones que pueden realizar funciones como análisis del color, comparación y superposición de imágenes, por ejemplo. El uso de compactación con pérdida en la actualidad puede impedir el uso de ciertas imágenes en el futuro.

Otra cuestión muy importante es que tanto la compactación con pérdida como la sin pérdidas añaden otro nivel más de complejidad a la codificación de un archivo, haciendo más difícil todavía la labor de los futuros arqueólogos para tratar de descifrar el contenido de los mismos.

De igual manera, una serie de esfuerzos para mejorar el comercio digital pueden representar una amenaza para la longevidad. Los esquemas de codificación cuyo objetivo es inhibir el uso no autorizado, añaden un nivel de complejidad a la codificación de un archivo, lo que a su vez complica aun más el trabajo de los futuros arqueólogos que tratarán de descifrar el contenido de los archivos. Y es difícil creer que todas las piezas de los complejos esquemas de comercio como por ejemplo, la arquitectura de contenido (que se basa tanto en la codificación como en la existencia continuada de una autoridad que pueda aprobar el pago de una transacción y revelar la clave adecuada para descifrar el archivo), sobrevivan lo suficiente como para garantizar el acceso a un archivo digital durante más de una década.

La mayoría de estos esquemas de codificación son patentados, y también la mayoría no se aviene a normas de amplia aceptación. El nivel de complejidad que añade la descodificación hace difícil creer que cualquiera podrá ser capaz de descodificar los archivos codificados hoy, ni siquiera dentro de cincuenta años.

El problema de la Inter-Relación

En el mundo digital, cada vez más la información está interrelacionada con otras informaciones. La World Wide Web es un ejemplo fundamental de cómo cualquier trabajo u obra puede incorporar otra serie de trabajos u obras o dirigirnos hacia ellas. Con frecuencia una obra determinada puede estar integrada o compuesta en realidad por más de un archivo bien diferenciable que puede estar presentado o no como si fuera un sólo archivo (por ejemplo, cuando un usuario visualiza lo que parece ser una sola imagen que en realidad está compuesta por una

imagen digital residente en un archivo, y cuyo título y otros metadatos descriptivos residen en un archivo por separado).

En la actualidad se estimula a los diseñadores de la Web a que se involucren en las mejores prácticas, aprovechen las ventajas de los aspectos de hipertexto de la World Wide Web y desglosen los documentos en pedazos pequeñitos, cada uno almacenado en un archivo diferente. Estos pedacitos se pueden volver a ensamblar en el momento de la visualización para que vuelvan a ser el archivo original completo, o las diversas piezas del rompecabezas se pueden recontextualizar en diferentes formas para diferentes propósitos. Esto quiere decir que incluso las obras más simples pueden estar integradas por varios archivos y que cualquier archivo puede ser parte de más de una obra.

En la Web de la actualidad es difícil luchar para lograr que nuestras obras persistan cuando las mismas señalan hacia otras obras que son propiedad de otros o forman parte de ellas. Debido a que el esquema actual de referencias de los archivos de la Web (el URL) se basa en la ubicación de un archivo, cada vez que cambia la ubicación del archivo, los enlaces se rompen y los usuarios sufren el error de mensaje más común que se produce en la Web ("404 No se encuentra"). Por lo general este problema es provocado por cierta simple reorganización en el sitio Web al que nos han dirigido (el cambio de nombre de un archivo o de una carpeta / directorio en alguna parte más arriba en la jerarquía de almacenamiento, o el cambio de nombre de un servidor). Pero este acto común de manejo de archivo / sitio crea el caos en cualquier trabajo u obra que incorpore a apunte hacia archivos de ese sitio.

Otro aspecto subordinado dentro del problema de la interrelación es la cuestión de determinar los límites o fronteras de un conjunto de informaciones (o incluso de un objeto digital). Hoy las fronteras de una obra digital no están confinadas a un archivo único. Con frecuencia, una página Web incorpora imágenes, gráficos y botones que están almacenados en archivos por separado (y algunas veces incluso en servidores por separado administrados por instituciones diferentes). Incluso obras tradicionales, como un artículo periodístico, o un ensayo están desglosados en varios archivos por separado que se ensamblan en el momento de la visualización, o permanecen como archivos independientes con vínculos en el cual el usuario debe marcar (con el propósito estilístico de no presentar al usuario con vistas que excedan dos pantallas completas).

Si queremos tomar medidas para preservar una de estas complejas obras, necesitamos desarrollar lineamientos en los cuales deberán basarse los límites o fronteras de las obras. Si una obra incorpora partes o piezas que son propiedad de otra organización o están administradas por otra

institución (íconos, logos, imágenes, textos), ¿el hecho de salvar una copia de alguna de esas piezas implica cuestiones de propiedad intelectual? Si queremos poder mostrar a los futuros investigadores qué tipo de información organizaba y distribuía una organización de estos tiempos, ¿debemos tratar de salvar la página principal de esa organización y todas y cada una de las páginas con las que esa página principal vinculaba? ¿Y entonces qué haríamos con las páginas vinculadas a esas otras páginas? ¿Dónde están las fronteras? Este problema no es muy diferente del problema que enfrentan los conferencistas de hoy día cuando quieren demostrar su sitio Web en una sala de conferencias que no tiene conexión a Internet; se ven obligados a decidir cuántos niveles o capas de archivos inter-relacionados tendrán que descargar en la computadora de la demostración.

Términos que definen la longevidad digital

Los enfoques técnicos claves para mantener la información digital viva en el tiempo fueron expuestos por primera vez en 1996 en un informe a la Comisión para Preservación y Acceso (Fuerza de Choque 1996).

- *El refrescamiento* implica trasladar periódicamente un fichero de un soporte de almacenamiento a otro para evitar el deterioro físico o la obsolencia de ese soporte. Debido a que los dispositivos de almacenamiento físicos (incluso los CD-ROMs) se deterioran, y debido a que los cambios tecnológicos hacen que los dispositivos de almacenamiento más antiguos (por ejemplo los floppy de 8") sean inaccesibles para las nuevas computadoras, es muy probable que durante muchos más años se haga necesario algún tipo de refrescamiento constante.

- *La migración* es un enfoque que implica trasladar los archivos periódicamente de un formato de codificación de archivos a otro formato que sea utilizable en el ambiente de las computadoras modernas. (Un ejemplo sería trasladar un archivo en WordStar a WordPerfect, después a Word 3.0, después a Word 5.0, y luego a Word 97.) La migración pretende limitar el problema de los archivos codificados en una amplia variedad de formatos de archivo que han existido en el tiempo convirtiendo gradualmente todos los formatos anteriores a un número limitado de formatos contemporáneos.

- *La emulación* trata de resolver un problema similar al que aborda la migración, pero su enfoque es concentrarse en las aplicaciones de software más que en los archivos que contienen la información. Los que están a favor de la emulación quieren construir un software que imite cada uno de los

tipos de aplicación que se haya escrito para cada tipo de formato de archivo y hacerlo correr en cualquiera que sea el ambiente de computación de que se disponga. (Así, con los emuladores adecuados, aplicaciones como WordStar y Word 3.0 podrían correr perfectamente en las máquinas actuales.)

Tanto el enfoque de la migración como el de la emulación requieren refrescamiento.

El problema de la custodia

Aunque se han desarrollado una serie de tradiciones en lo que respecta a qué organizaciones deben asumir la responsabilidad de preservar y mantener los diversos tipos de materiales analógicos (correspondencia, manuscritos, material impreso), no existe todavía esa tradición con respecto al material digital. Como resultado de esto, gran parte del material que se origina en formato digital en la actualidad se va por esas grietas y no es muy probable que esté accesible para las generaciones futuras.

Por ejemplo, los bibliotecarios a cargo de colecciones especiales que persiguen de forma agresiva el desarrollo de colecciones basado en material impreso en sus áreas particulares de especialidad reclaman que debería ser responsabilidad del personal de computación de las instituciones el desarrollo de colecciones de materiales originados en formato digital ("Collecting at the Margins. . . ," 1999). Sin embargo, ese mismo personal de computación dice que debe ser responsabilidad de los especialistas en la materia o en el tema desarrollar las colecciones de materiales en formato digital. Y mientras tanto, gran parte de estos frágiles materiales simplemente no se colecciona.

Otro ejemplo es la correspondencia, que en el mundo analógico deja un rastro en papel. La mayoría de las organizaciones sigue lineamientos para salvaguardar cantidades significativas de correspondencia en soporte de

papel. Pocas organizaciones han desarrollado lineamientos similares para salvaguardar la correspondencia electrónica, y poco individuos tienen idea de cómo pueden salvar su propia correspondencia incluso si tuvieran interés en hacerlo. Este problema se hace cada vez más agudo según se origina más y más correspondencia en formato digital.

Un último ejemplo pertenece al dominio de la creación literaria. En el mundo analógico, los autores solían dejar el rastro de su importante proceso de creación en forma de numerosas versiones o borradores, manuscritos marcados, y correspondencia. En la actualidad utilizan procesadores de texto y correo electrónico tanto para los borradores como para la correspondencia. Con frecuencia, sólo salvan muy pocos de los borradores y ninguna correspondencia.

Una cuestión de mayor importancia que enfrentamos en la actualidad es: ¿Quién deberá ser el responsable de salvar los materiales en formato electrónico? ¿Deberán cargar con esta responsabilidad los propios individuos? ¿O deberán las entidades sociales (como empresas, bibliotecas, archivos y asociaciones de profesionales) intervenir de manera agresiva para salvar el material? ¿Y cómo van a decidir qué es lo que se debe salvar?

Otro problema crítico es: ¿Cómo se deben salvar? Nuestro campo profesional todavía necesita desarrollar lineamientos y mejores prácticas de manera que las organizaciones y los individuos que quieren realizar un esfuerzo y tratar de que la información digital persista sepan cómo hacerlo.

Una función clave de los archivos es garantizar la autenticidad de una obra. Esto lo hacen recogiendo *evidencias o pruebas* y manteniendo una *cadena de custodia*. Pero cuando las obras están sometidas a repetidos actos de refrescamiento tal y como proponen la mayoría de los enfoques encaminados a la longevidad digital (véase Sidebar), se interrumpen y se pierden estas formas tradicionales de garantizar la autenticidad. Los archivos que han sido copiados repetidas veces a un nuevo estrato enfrentan la probabilidad de que se introduzcan cambios, y sabemos poco acerca de cómo controlar la mutabilidad a través de repetidos refrescamientos.

El problema de la traducción

Cuando el contenido se traduce a nuevos dispositivos de entrega (por ejemplo, formatos digitales), el cambio de forma con frecuencia sirve para cambiar parte del significado. Las conversiones de un formato analógico a otro formato analógico enfrentan este mismo problema, al

igual que las conversiones de formato analógico a formato digital (una fotografía de una pintura no es igual que la pintura, y una representación digital de un objeto no es igual que el objeto) (Besser 1987).

Debido a que podemos hacer copias de los archivos digitales, algunas personas creen, equivocadamente, que la conversión digital-a-digital no enfrentará los mismos problemas de traducción que enfrentan las conversiones analógico-a-digital. Y esto no es cierto porque, aunque los bits en el contenido del archivo pueden ser idénticos, el ambiente de aplicaciones utilizado para visualizar el archivo será, casi seguro, diferente. En realidad, la razón misma de convertir el archivo es que somos incapaces de sostener de manera exitosa un mismo ambiente de aplicación en el tiempo.

Muchas personas han experimentado esto en las ocasiones en que su procesador de texto importa "exitosamente" un documento creado en una versión anterior del mismo procesador de texto, porque se pierde el formato (cambian los márgenes, el centrado, los subrayados, la fuente) o la puntuación (se pierden los apóstrofes o las comillas). Esto también se puede producir en ambientes de emulación, porque los creadores de estos ambientes deciden y seleccionan cuáles aspectos del ambiente van a emular, y no pueden simplemente emular todos y cada uno de los aspectos. (Por ejemplo, al creador [Jaron Lanier] de *Moon Dust*, uno de los primeros juegos de computadora, se le mostró una reciente emulación del juego, y él manifestó que el juego era completamente diferente al que él había diseñado, porque el ritmo era diferente.)

Cuando se salva un trabajo, es de crítica importancia que salvemos partes del ambiente de esa obra que puede que no sean obvias de inmediato. Por ejemplo, es probable que cualquiera reconozca que debemos salvar la imagen de cada página en un libro digitalizado. Pero para que el libro sea utilizable, debemos también salvar importantes comportamientos del libro, como los metadatos y los comportamientos acompañantes que permitirán a los futuros usuarios pasar las páginas, pasar directamente del índice a un capítulo en particular, o simplemente hacia adelante o hacia atrás en el cuerpo principal del texto del libro o por entre las citas o las notas al pie (Making of America II . . ., 1998). Salvar sólo las imágenes de páginas de un libro sin sus comportamientos sería como salvar un juego de video con las interacciones en un determinado tipo de representación, pero a la vez olvidando una de las funciones más importantes.

Con una obra que comienza en formato digital, necesitamos comprender mejor los aspectos del ambiente original de la obra que son críticos para visualizar la misma, y necesitamos buscar vías para que todas los comportamientos importantes de la obra sean sostenibles según

trasladamos sus contenidos a través de generaciones de migraciones o de emulaciones (Besser & Gilliland-Swetland, 1999). También necesitamos comprender cómo afecta la naturaleza de una obra cada nuevo ambiente de visualización. (Por ejemplo, muchos cineastas argumentarían que su película cambia de manera radical cuando se muestra en una pantalla de video. ¿Cómo se sentirán los creadores actuales de multimedia si sus obras se mostraran en ambientes futuros en los que ya no existen los tubos de rayos catódicos para visualizar?)

Las vías para mejorar o alargar la longevidad digital

Teniendo en cuenta estos formidables problemas, ¿cómo podemos tener la esperanza de garantizar la longevidad de las obras de hoy que queremos preservar? Algunos de estos enfoques se plantearon por primera vez en 1998 (Lyman & Besser, 1998), pero la información que aparece a continuación es resultado de ideas y acontecimientos más recientes.

El enfoque amplio

En primer lugar, es necesario reconocer que sabemos mucho acerca de cómo preservar los bits en el tiempo. Durante más de un cuarto de siglo la comunidad de procesamiento de datos ha movido grandes masas centralizadas de bits de un soporte físico de almacenamiento a otro. Nuestra comunidad necesita estudiar en los departamentos de procesamiento de datos de las corporaciones y las universidades para conocer y aprender sus experiencias y para obtener cifras de costos. Después tenemos que analizar cómo se podrían aplicar estas informaciones y cifras a los cuerpos de información digital menos centralizados de nuestra comunidad.

A la vez que analizamos esta experiencia, también tenemos que tener en cuenta que preservar bits es sólo una pequeña parte del problema. Este problema se minimiza ante los problemas mucho más complejos de garantizar que los formatos de los archivos sean accesibles y de los problemas que incluyen la organización, la política, y el papel y la responsabilidad de cada cual.

En los miles de años que han transcurrido desde que la Biblioteca de Alejandría se destruyó, la redundancia ha sido la clave de la preservación de la información. La existencia de múltiples copias de una obra geográficamente dispersa por varios lugares ha ayudado a preservar obras tanto de desastres naturales como de desastres provocados por el hombre (que van desde los incendios y los terremotos hasta la destrucción accidental de una serie de obras). Cualquier estrategia a largo plazo para

la preservación de información digital debe incorporar relaciones de cooperación entre las locaciones y organizaciones físicamente dispersas. Tenemos que desarrollar proyectos de cooperación internacional en los que las organizaciones estén dispuestas a almacenar y refrescar las copias redundantes de obras que están bajo la custodia de dichas organizaciones.

Las leyes actuales de la propiedad intelectual inhiben a los archivos y las bibliotecas de preservar información en formato digital, sobre todo porque gran cantidad de la información que esas instituciones adquieren está controlada por una licencia en lugar de ser propiedad de la misma. Un estudio reciente sobre derecho de autor realizado por la Academia Nacional de Ciencias (Comité sobre los Derechos de la Propiedad Intelectual. . . , 2000) recomendó enérgicamente que se cambien las leyes de propiedad intelectual para que se permita que estas instituciones puedan preservar legalmente la información en formato digital, y que se asignen fondos para la preservación digital. Es necesario continuar monitoreando los cambios en las leyes de la propiedad intelectual (Besser, *Copyright* sitio web) y presionar para lograr cambios que nos permitan realizar la preservación digital sin tener que enfrentar cargos o sentencias criminales.

Necesitamos más experiencia en las dos estrategias competitivas para la preservación digital – *la emulación y la migración* (véase Sidebar). El enfoque de la emulación es muy experimental, y tenemos que monitorear los dos estudios experimentales en el ámbito internacional que han comenzado recientemente a explorar en esta área: NEDLIB, patrocinado por la Comunidad Europea (Networked European Deposit Library website); y el Proyecto CEDARS (CURL Exemplars in Digital Archives website), patrocinado por el Comité Conjunto de Sistemas de Información de Gran Bretaña y la Fundación Nacional de las Ciencias de Estados Unidos.

Lo que nosotros como comunidad podemos hacer

Aunque nadie ha resuelto todavía la variedad de problemas que rodean la longevidad digital, existe un número particular de acciones que podemos emprender y que aumentarían las probabilidades de que una obra que queremos salvar siga estando accesible durante un período de tiempo prolongado. También hay una serie de acciones con las que nuestra comunidad en su conjunto debe comenzar a lidiar para poder reducir este problema tan inmenso.

Nuestra comunidad necesita insistir en la creación de normas claras y legibles para que un objeto digital auto identifique su formato y las aplicaciones necesarias para visualizarlo. Con una norma para insertar el nombre de la aplicación de visualización en un lugar específico con un

encabezamiento de imagen, los arqueólogos del siglo XXII que descubran los archivos de hoy día podrán al menos descubrir qué aplicaciones necesitarán buscar para poder visualizar ese archivo. El trabajo sobre este aspecto y también sobre una serie de otros problemas relacionados con la longevidad de las imágenes digitales comenzó como parte de una reunión realizada en la primavera de 1999 y patrocinada por la Comisión para Preservación y Acceso, la Organización Nacional de Normas para la Información, y el Grupo Bibliotecas de Investigación (Besser, 1999).

Nuestra comunidad necesita comprender mejor en qué forma se relaciona una información con otra u otras informaciones (Besser & Gilliland-Swetland, 1999). En particular, necesitamos ganar mayor claridad en lo que respecta a las fronteras de los objetos de información. Cuando estamos tratando de salvar algo (en especial un objeto de hipertexto o hipermedia), tenemos que saber qué partes necesitamos salvar en realidad.

Por último, nuestra comunidad necesita desarrollar una serie de lineamientos concretos que puedan utilizar las personas y las organizaciones que desean hacer persistir la información. En cierto modo, este capítulo es un intento más en la lucha por lo que se pudiera incluir en esos lineamientos.

Cuando se toma la decisión de preservar en formato digital una serie de obras, la institución debe comprender, en primer lugar, las necesidades especiales de los tipos de obras contenidos en esa colección. Esto quiere decir comprender en qué forma la reconversión de formato puede afectar la comprensión y la utilidad de esas obras. Significa comprender cuáles son las fronteras de esta obra y cuáles piezas deben salvarse (quizás hasta incluir las piezas contextuales). Como vimos en el ejemplo del libro digitalizado, esto también quiere decir salvar los comportamientos de una obra, y no simplemente sólo sus contenidos.

El papel o la función de los metadatos

A esta altura de las circunstancias, la inclusión de metadatos extensos o amplios es la mejor manera que tenemos de reducir al mínimo los riesgos de que un objeto digital se vuelva inaccesible. Los metadatos, si están bien utilizados, pueden:

- Identificar el nombre de la obra, quién la creó, quién realizó la reconversión de formato, y oras informaciones descriptivas

- Proporcionar informaciones únicas y vínculos con organizaciones, archivos, o bases de datos que poseen metadatos más amplios y descriptivos acerca de esta obra (esto es de especial importancia en el

probable caso de que el archivo digital y sus metadatos externos se separen por alguna razón)

- Explicar el ambiente técnico que se necesita para visualizar la obra, incluidas las aplicaciones y números de versiones que se necesitan, otros archivos con los que hay que vincular, y así sucesivamente.

Varios tipos de metadatos que hoy nos parecen insignificantes o sin importancia pueden ser de importancia crítica para visualizar adecuadamente estos archivos en el futuro. (Por ejemplo, información salvada acerca del perfil de color de un escáner será crítica para sistemas futuros de manejo de color para poder responder por las diferencias de color en el momento de presentar en pantalla o para presentar los colores en un dispositivo en particular.) Una buena práctica es salvar cualesquiera metadatos que sean económicos y fáciles de capturar, o que alguien nos ha indicado que pudieran ser importantes en algún momento.

Bibliografía

Todos aquellos involucrados en la planificación de la longevidad digital deberían leer los textos claves que estudian los problemas que se presentan en nuestro campo: el informe de la Comisión para Preservación y Acceso (Fuerza de Choque, 1996), la conferencia sobre preservación digital patrocinada por el Getty Institute *Time & Bits* (MacLean & Davis, 1998), y otros artículos que aparecen como referencia en la Sunsite Longevity Page (Besser, *Digital Longevity* website). También pueden monitorear de manera continua la página web Sunsite Longevity Page (Besser, *Digital Longevity* website), las publicaciones de la Comisión para Preservación y Acceso (Commission on Preservation and Access website), y el trabajo realizado por el Archivo Internet (Internet Archive website).

Besser, Howard. "The Changing Museum" en Ching-chich Chen, ed., *Information: The Transformation of Society*, pp. 14-19. Actas de la 50a Reunión Anual de la Sociedad Noretamericana para la Información Científica, Medford, NJ: Learned Information, Inc.

-----. *Copyright* (website). http://www.gseis.ucla.edu/~howard/Copyright/

-----. *Digital Longevity* (website). http://sunsite.berkeley.edu/Longevity/

-----. Image Metadata: meeting summary, 1999. http://sunsite.berkeley.edu/ Imaging/Databases/Metadata/niso-4-99-summary/

Besser, Howard y Anne Gilliland-Swetland. *Multimedia: Issues in Using Visual Material in Cultural Heritage Organizations*, primavera 1999 clase y website. http://www.sims.berkeley.edu/impact/s99/

"Collecting at the Margins: Social Protest and Counterculture Materials," Collection Development Librarians of Academic Libraries Discussion, enero. 30, 1999. Conferencia de invierno de la American Library Association, 1999, Philadelphia.

Comisión para Preservación y Acceso (website). http://www.clir.org/programs/cpa/cpa.html

Comité para los derechos de la Propiedad Intelectual y la Infraestructura en Surgimiento de la Información, Consejo Nacional para la Investigación, Academia Nacional de Ciencias. "The Digital Dilemma: Intellectual Property in the Information Age," Washington: National Academy Press, 2000.

Ejemplares de CURL en Digital Archives (website). http://www.leeds.ac.uk/cedars/

Internet Archive (website). http://www.archive.org

Lyman, Peter y Howard Besser. "Defining the Problem of our Vanishing Memory: Background, Current Status, Models for Resolution" en Margaret MacLean y Ben H. Davis, eds. *Time & Bits: Managing Digital Continuity,* pp. 11-20, Los Angeles, CA: J. Paul Getty Trust, 1998.

MacLean, Margaret y Ben H. Davis, eds. *Time & Bits: Managing Digital Continuity,* Los Angeles: J. Paul Getty Trust, 1998.

Making of America II White Paper (1998). http://sunsite.berkeley.edu/moa2/

Networked European Deposit Library (website). http://www.konbib.nl/nedlib/

Sanders, Terry. *Into the Future: Preservation of Information in the Electronic Age* (16 mm film, 60 minutos). Santa Monica, CA: American Film Foundation, 1997.

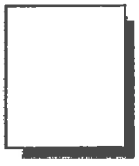
Task Force on Archiving of Digital Information. *Preserving Digital Information,* Comisión para Preservación y Acceso y Grupo de

Bibliotecas de Información, 1996.
http://www.rlg.org/ArchTF/tfadi.index.htm

Northeast Document Conservation
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.



X

Comentario Académico: Un usuario expresa su opinión

Charles Rhyne
Reed College

Hablo como académica y profesora. Junto con mis colegas y estudiantes, dependo de los grandes almacenes de información que representan las bibliotecas, los archivos y los museos. La mayor parte del tiempo, damos por sentado la existencia de esos recursos, pero cuando falta un documento que necesitamos, admitimos la enorme dependencia que tenemos de estos materiales y lo grave que es la pérdida o destrucción de los mismos. En ocasiones, reflexionamos sobre la visión de futuro que tuvieron aquellos que establecieron estas instituciones y aplaudimos el buen juicio de los que adquirieron, organizaron y han cuidado estos materiales. También admiramos la compleja e ignorada actividad que permite que estos materiales estén al alcance de tantos usuarios con tan variados intereses y necesidades.

Hace poco, una estudiante comenzó junto conmigo su tesis sobre el famoso sitio precolombino de Monte Albán, que ella descubrió el año pasado por un programa de televisión y espera poder visitarlo en el futuro próximo. Esperando por ella en la biblioteca hay estantes repletos de textos e ilustraciones, artículos en revistas populares y publicaciones científicas, catálogos de exhibiciones y artículos de periódicos en microfilme y microfichas. Varios cientos de diapositivas originales en color se encuentran depositadas en el Departamento de Historia del Arte, disponibles y accesibles para su uso en clases y para los proyectos de los estudiantes. A quince minutos, en el Museo de Arte de Portland, se exhiben otros objetos relacionados con el tema.

La mayoría de los materiales digitales no son confiables como evidencia

Ahora, por supuesto, ella también dispone de materiales en formato digital para realizar su estudio, en CD-ROMs y en la World Wide Web (Internet). Pero existe una sorprendente diferencia entre estos nuevos materiales digitales y los materiales analógicos tradicionales que hemos utilizado en el pasado. En pocas palabras, sólo unos pocos de esos materiales son confiables como evidencia. El hype comercial que rodea el uso de las computadoras ha producido CD-ROMs de tan pobre contenido y de imágenes de tan mala calidad que muy pocos se hubieran publicado en formato impreso. Cualquier búsqueda en Internet sobre cualquier tema arroja tanta cantidad de sitios que enloquece, gran parte de ellos con un diseño innovador, de aficionados, y fotografías turísticas con muy baja resolución. Afortunadamente, se añaden cada vez más sitios de alta calidad, muchos de ellos basados en materiales originales pertenecientes a las colecciones de los museos, universidades e institutos de investigación. Pero incluso en estos, las imágenes que se encuentran en Internet son, en raras ocasiones, de alta resolución, y cuando es este el caso, se visualizan en monitores cuyas características de tamaño y visualización son impredecibles, casi siempre ajustadas a la velocidad más que a la calidad. Sólo en las páginas Web más avanzadas el usuario puede conocer qué es lo que busca, y hasta esto puede ser alterado o desaparecer sin previo aviso. La inundación en Internet de materiales no confiables y la falta de normas reconocidas para juzgarlos ha dilatado la aceptación de publicaciones digitales en la evaluación de trabajos profesionales de los profesores de las universidades, y estas no estimulan a los académicos serios a hacer ese tipo de trabajo.

Pocos profesores académicos participan

Todos tenemos mucha culpa. Somos testigos de una inmensa transformación en la creación, retención, disponibilidad y uso de nuestro registro cultural, que ya está transformando la manera en que entendemos el mundo. Pudiera esperarse que todos los seres humanos que estén intelectualmente vivos en las instituciones de investigación y de educación reconocerían la importancia única de poder participar en una transformación histórica tan trascendental, y posiblemente contribuir con la misma. Muy pocos de nosotros los profesores nos hemos involucrado profundamente, sin embargo, me atrevería a decir que tampoco se nos ha pedido que lo hagamos.

Nos sentamos a
mirar los toros
desde las gradas...

Muchos de nosotros ahora utilizamos el correo electrónico, quizás incluso hemos desarrollado páginas Web para nuestras clases y, especialmente cuando disponemos de financiamiento puede que hayamos propuesto nuestro propio material de investigación para proyectos de los estudiantes. Pero de lo contrario, nos sentamos a mirar los toros desde las gradas, observado, siguiendo los acontecimientos, algunas veces quejándonos de Internet como "el agujero negro" o del control del mundo digital por parte de los ingenieros. No hemos puesto en claro que diferentes disciplinas tienen diferentes necesidades, que hay ciertas necesidades básicas que deben satisfacerse si se va a utilizar en realidad el material digitalizado, y que las normas y procedimientos que adoptan los especialistas en tecnología e información no necesariamente sirven a todos nuestros estudiantes, profesores o nuestros colegas investigadores. Esto representa una gran pérdida, porque el futuro del mundo digital tiene mucho que ganar de la participación activa de los usuarios en cada etapa del proceso y mucho que perder si lo seguimos a distancia.

Los Prototipos

Algunas áreas de la comunidad de los investigadores — las más notables la de las investigaciones del espacio, las militares y las médicas — bien financiadas por los programas gubernamentales y las fundaciones, han estado al frente del desarrollo digital. En ocasiones, el resto de la comunidad de los académicos y el público en general se benefician con esas innovaciones, pero la lección fundamental que hay que aprender es que los proyectos digitales en estos campos han estado dirigidos a satisfacer las necesidades particulares de sus usuarios finales. En estos campos, se ha buscado el apoyo de los usuarios, quienes han estado profundamente involucrados en cada una de las etapas del proceso.
En el campo de las Ciencias Sociales y Humanísticas, también se han desarrollado algunos prototipos impresionantes con financiamiento de fundaciones, impelidos por el reconocimiento del potencial de la revolución digital. En los principales ejemplos, las bibliotecas y archivos nacionales, las grandes universidades, y los institutos de investigación han comenzado a digitalizar sus materiales raros a partir de normas de calidad muy elevadas, ya que admiten que es peligroso manipular documentos raros o valiosos, que muchas veces es costoso volver a

digitalizar, e imposible predecir las cosas que el público va a preguntar con respecto a esos documentos en el futuro. Una característica notable de los mejores prototipos es que los académicos y curadores de los materiales han aunado fuerzas con los especialistas en computación, para con frecuencia producir no sólo formas innovadoras de registrar y acceder a los resultados digitales, sino también una nueva visión del material, lo que constituye un avance desde el punto de vista académico.

Las imágenes fotográficas

Como especialista en historia del arte, estoy más que consciente de los usos que hacemos de las fotografías de arte y arquitectura, la documentación más esencial en nuestro campo. Por tanto, tengo muy presentes las características que necesitamos en las imágenes digitales si queremos que las fotografías sean útiles como material de estudio y de investigación para los académicos. Permítanme utilizar el mundo que tan rápidamente se expande de las imágenes digitales para describir algunos de los defectos actuales y el inmenso potencial para el futuro.

En el mundo académico, las palabras siempre han tenido prioridad sobre las imágenes.

Con frecuencia, cuando observamos imágenes de obras de arte en el monitor de la computadora, yo comento sobre la falta de detalles, la imposibilidad de ver la forma en que está construido un arco o de distinguir los brochazos en un cuadro. Muchas veces esto se considera una queja innecesaria, una expectativa falta de razón en el caso de una imagen de computadora. Para abordar este problema, yo me refiero algunas veces a las imágenes de mapas. En este caso, aunque mi deseo de apreciar la forma de una elevación o la ubicación de un lago se pudiera desechar, tan pronto como señalo que las palabras que identifican estos accidentes geográficos son ilegibles, todo el mundo está de acuerdo al instante en que la imagen digital tiene que capturarse y presentarse a mayor resolución.

En el mundo académico, las palabras siempre han tenido prioridad sobre las imágenes. Pero en la sociedad, la fotografía ha ido reclamando un lugar cada vez más importante en las primeras páginas de los periódicos y revistas, sin mencionar la televisión, y el mundo académico se está adaptando poco a poco. Las publicaciones profesionales contienen cada vez más y mayores ilustraciones, en colores, que reflejan no sólo la

reducción en el precio de la reproducción, sino también el papel cada vez más importante que desempeñan la fotografía y el cine en muchas disciplinas académicas. La biología, la medicina y la arqueología contemporáneas dependen cada vez más de las imágenes fotográficas, y también otras ciencias como la historia, la antropología y otras ciencias sociales han descubierto un caudal de información no estudiada esperando por ser descubierta en las fotografías. ¿Qué sostiene en la mano esa persona? ¿Qué dice ese letrero en la vidriera de la tienda? ¿Cómo se vestía la gente para asistir a ese acto o esa actividad? Como resultado, la adquisición y preservación de fotografías se ha convertido en una actividad tan importante como la preservación de palabras.

Necesitamos herramientas de búsqueda visual comparables a las de búsqueda de textos.

De igual manera, la búsqueda en computadoras a través de palabras clave y categorías de nombres, tan útil en el caso de los textos, sirve para búsquedas en fotografías sólo cuando la fotografía ya se ha señalado con la palabra clave que el usuario quiere localizar. Como conocen los archivistas fotográficos, este es un proceso que requiere un trabajo intenso y un alto grado de conocimiento experto en la materia; además, es imposible predecir todas las categorías que serán de interés para los futuros académicos. Aparte de las búsquedas verbales, muchos usuarios necesitan también procedimientos de búsquedas visuales. No sólo los artistas que dan clases de diseño básico o arquitectura urbana necesitan herramientas de búsqueda visual, sino también las necesitan los botánicos que buscan determinadas características en la estructura de las hojas de una planta o los antropólogos que buscan variedades de ornamentación del cuerpo. Sería más útil poder escanear miles de fotografías de textiles en todo el mundo para analizar ciertos patrones de tejidos que pretender que cada archivista fotográfico introdujera el nombre correcto de cada patrón de tejido en cada fotografía de textil. Por difícil que sea de lograr, necesitamos herramientas de búsqueda visual comparables en rango y especificidad a las de los textos.

Las normas tradicionales de evidencias

Las normas profesionales de evidencias se desarrollaron en cada disciplina a través de la revisión de otros colegas profesionales y los debates continúan a pesar de las décadas transcurridas. Mientras que por

una parte constantemente aparecen productos de computadoras nuevos y mejorados, y se mejoran cada vez más, por otra los objetivos de investigación y enseñanza evolucionan con lentitud, basados en enfoques exitosos en uso en la actualidad. Es algo que no se produce de manera natural. El reto es estar abiertos a las dramáticas posibilidades que ofrece la tecnología digital sin dejarnos engañar por el entusiasmo sin límites de la nueva tecnología. Esto exige, por encima de todo que estemos bien afincados en nuestras respectivas disciplinas y una visión a largo plazo de lo que queremos lograr en nuestra investigación, enseñanza, y en la diseminación de información e ideas.

Los estudiantes y los académicos utilizarán algunos materiales digitales, porque ya pueden acceder a ellos en las computadoras de sus instituciones, pero no están en las bibliotecas. La habilidad de las computadoras de lograr que tanto los textos como las imágenes se puedan acceder desde cualquier parte del mundo es uno de los grandes logros de la revolución de las computadoras. Sin embargo, a la larga, para los trabajos académicos serios (incluidos los trabajos de curso, las tesis de maestría y doctorado que se escriben en cientos de universidades en todo el mundo cada año), el material digital sólo sustituirá el material impreso cuando cumpla las mismas normas de exactitud y confiabilidad. Hoy por hoy, estamos muy lejos de que eso se cumpla. De todas formas, ya se empiezan a publicar nuevas publicaciones científicas y profesionales en versión electrónica u on-line, revisadas por otros académicos y con las mismas normas de las publicaciones impresas tradicionales; y poco a poco, los estudiantes se entrenan para discriminar las publicaciones on-line de la misma manera que han aprendido a discriminar los materiales impresos. Los bibliotecarios desempeñan un papel esencial en este proceso educacional.

Los supuestos defectos se pueden convertir en logros

Sin embargo, no debemos acomodarnos a normas anteriores, porque esas normas eran defectuosas. Por ejemplo, escuchamos con frecuencia la queja de que las imágenes digitales de fotografías se pueden cambiar con facilidad sin que el usuario se percate de ello – Pero esto también es válido para las fotografías (eso sin contar que muchos de los cambios hechos en las fotografías eran burdos y por lo tanto se detectaban con facilidad). La tecnología digital en realidad ofrece la posibilidad de registrar esas imágenes automáticamente por primera vez. Y esto es, de hecho, lo que se necesita si las imágenes digitales de fotografías se van a utilizar como evidencia. Necesitamos un sistema que proporcione la posibilidad de registrar automáticamente la creación de cada imagen y los cambios realizados en las mismas dentro del record digital de la imagen.

Se puede lograr que las imágenes digitales sean más confiables como evidencia . . .

Quizás esto nos anime a registrar también lo que deberíamos haber hecho todo este tiempo, la forma en que se hace cada fotografía. Ahora contamos con cámaras que registran en la misma película o registro digital la fecha y la hora en que se tomó cada fotografía, al menos si la fecha y la hora se han introducido correctamente. Pero también es necesario conocer el tipo de cámara, el lente y la película utilizada. Ya ahora es de conocimiento general que ninguna fotografía es la imagen exacta de un objeto o un sujeto. Todo depende de la iluminación, el ángulo de visión, la distancia hasta el sujeto, y por supuesto, las características de la película y el lente. Muchas de estas cosas son visibles de forma razonable en la fotografía misma, pero muchas no lo son. Hasta el punto que estas se pueden registrar ahora de manera automática en los datos que constituyen el registro digital de cada imagen, existe la oportunidad de hacer imágenes de fotografías más confiables como evidencia que lo que han sido en el pasado.

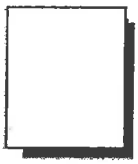
Las imágenes digitales pueden hacer posible otros adelantos significativos. Para dar un ejemplo, la lectura cuidadosa de artículos académicos en revistas científicas profesionales de todas las disciplinas que dependen de la fotografía como evidencia tienen el defecto de lo inadecuadas que son las ilustraciones. Las reproducciones son pocas en cantidad, demasiado pequeñas y muy pocas veces tienen el color adecuado. Ha sido demasiado costoso publicar la gran cantidad de fotografías de las cuales dependen los textos. En muchos casos, los lectores tienen que aceptar los argumentos simplemente por fe o no emitir un juicio. Las fotografías no nos permiten poner en duda las aseveraciones de los artículos. Pero es mucho menos costoso publicar imágenes grandes a todo color de alta calidad en Internet. Ahora sería posible publicar artículos totalmente ilustrados en Internet con reproducciones de alta calidad de todas las fotografías de las que depende el artículo. De igual manera, ahora podemos incluir apéndices de todos los datos con los detalles en los que se basa gran parte de la investigación.

Una sociedad con participación plena

No tengo duda alguna de que la revolución digital es una de las grandes revoluciones de la información en la historia de la humanidad, comparable con la invención de la imprenta y de la fotografía, pero que se produce a velocidad supersónica. Tenemos el privilegio de vivir en una

época que nos da la oportunidad de ser testigos de esta transformación, de experimentarla, e incluso de participar en su desarrollo. En realidad, si todos los elementos de la sociedad participan en su desarrollo, esa transformación podrá lograr su inmenso potencial. Al decidir qué materiales se van a digitalizar, cómo se van a preservar, y como vamos a hacerlos accesibles, tenemos que reconocer el interés tan serio del público en todas las esferas del conocimiento humano, el derecho del público a saber, y los enormes e inagotables recursos que ofrecen estos materiales. La revolución digital nos ofrece la primera oportunidad de ser una sociedad con participación plena.

Buscar Contenidos Sitio del NEDCC



Northeast Document Conservation Center
100 Brickstone Square
Andover, MA 01810-1494
Teléfono: (978) 470-1010
Fax: (978) 475-6021

Última modificación: 4 de diciembre del 2000

Copyright 2000. Northeast Document Conservation Center.