



Academia
de **Ingeniería**
México

EL CASO DE LOS ARCHIVOS JUDICIALES FEDERALES: 1815 - 2003.

ESPECIALIDAD: COMUNICACIONES Y ELECTRÓNICA

MÓNICA MARÍA DEL ROSARIO BARRERA RIVERA

DOCTORADA DE LA UNIVERSIDAD
NACIONAL AUTÓNOMA DE MÉXICO

1. Resumen Ejecutivo	
2. Palabras Clave	
3. Introducción	
4. Desarrollo del tema	
4.1. Marco Jurídico	
4.2. Planteamiento del Problema	
4.3. Terminología Básica de la digitalización	10
4.4. Solución de la problemática planteada	29
5. Conclusiones	35
6. Referencias	39
7. Bibliografía	40
8. Agradecimientos	42
9. Resumen Curricular	43
1. RESUMEN EJECUTIVO	

El Poder Judicial Federal de México cuenta con un archivo judicial histórico, de 1815 a 2003, correspondiente a 98 kilómetros de documentos, es decir, más de 7 millones de expedientes, ubicados en más de 40 localidades de la República Mexicana.

Adicionalmente y en el marco de las modificaciones aprobadas en 2007, al Artículo 6º de la Constitución Política de los Estados Unidos Mexicanos, que obliga a todos los entes públicos a automatizar su información para que esté disponible a la ciudadanía, en el marco de la cultura de la transparencia y rendición de cuentas, surge la imperiosa necesidad en este Poder Federal, de clasificar y depurar los archivos históricos judiciales, que son altamente especializados y guardarlos de manera segura, permitiendo su consulta, al menor costobeneficio posible. La guarda y explotación están implícitas en un programa de digitalización documental, cuyo objetivo es "conservar para difundir", inmerso a su vez en un Sistema de Gestión Electrónica de Documentos (SGED).

Para el desarrollo del tema, se expuso y exploró el marco jurídico específico: El Artículo 6º de nuestra Carta Magna/ la Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental y los Lineamientos emitidos por el Poder Judicial Federal sobre la clasificación documental, incluida la información clasificada como reservada y como confidencial, el 5 de septiembre de 2003.

Con este marco, se hace el planteamiento del problema, su dimensión y clasificación, de acuerdo al análisis y diagnóstico elaborado para esta realidad del Poder Judicial Federal.

Para el planteamiento de un conjunto de soluciones a la problemática presentada, se expone la terminología y aspectos básicos que hay que tomar en cuenta para las soluciones en materia de digitalización: conversión analógico-digital; muestreo; resolución; dimensiones de pixel, profundidad, rango dinámico/ tamaño de archivo, compresión, formatos de archivo; conversión; control de calidad; infraestructura técnica, a saber, cadena de digitalización, creación de imágenes y gestión de archivos; así como entrega de la información (redes, monitores, impresoras).

Llevar a cabo un proyecto de digitalización significa establecer un equilibrio entre las necesidades de los usuarios reales y potenciales, la infraestructura tecnológica utilizada por el proyecto y los recursos financieros y humanos disponibles.

En el proceso de digitalización de documentos, las instituciones pueden optar por realizar la totalidad de los procesos internamente o contratarlo a una empresa especializada externa. Las ventajas y desventajas de este tipo de soluciones también se plantean en el trabajo.

Finalmente, se resaltan los beneficios que puede brindar este programa de digitalización para el país; el Poder judicial de la Federación y el público, en general.

2. PALABRAS CLAVE

- Archivo.
- Digitalización.
- Reconocimiento óptico de Caracteres (OCR).
- Escáner.
- Sistemas de Gestión o Administración Electrónica de documentos (SGED).
- Transparencia y acceso a la información pública. • Conservar para difundir.

3. INTRODUCCION

Ante la problemática de conservar para utilizar y difundir 98 kilómetros de documentos; es decir más de 7 millones de expedientes de sentencias, de archivo histórico judicial de 1815 a 2003, ubicado o disperso en más de 40 localidades en todo el país, el Poder Judicial Federal se plantea la necesidad de clasificar y depurar esta información altamente especializada y guardarla de manera segura, permitiendo su consulta, al menor costo-beneficio posible.

Adicionalmente, ante las modificaciones al Artículo 6^o de la Constitución Política de los Estados Unidos Mexicanos, que obligan a todos los entes públicos a automatizar su información, para que esté disponible a la ciudadanía, en el marco de la cultura de transparencia y rendición de cuentas, en este contexto, surge la necesidad de implantar en esta materia, en el Poder Judicial Federal, un ambicioso proyecto de digitalización del archivo histórico judicial, cuyo objetivo es "conservar para difundir".

Son varias las razones que originan la implementación de un proyecto de digitalización, es decir, la conversión digital de documentos originales no digitales y entre ellas se encuentran las siguientes:

- Incrementar el acceso: ésta es la razón principal y la más obvia, cuando se sabe que hay una demanda por parte de los usuarios y el archivo desea mejorar el acceso a una determinada colección.
- Mejorar los servicios para un grupo creciente de usuarios, proporcionando un acceso de mayor calidad a los recursos de la institución.
- Reducir la manipulación y el uso de materiales originales frágiles y crear una "copia de seguridad" para el material deteriorado como documentos quebradizos.
- Ofrecer a la institución oportunidades para el desarrollo de su infraestructura técnica y para la formación técnica de su personal,
- Impulsar el desarrollo de recursos cooperativos, compartiendo intereses comunes con otras instituciones para incrementar el acceso a nivel internacional.

Cuando una institución inicia un proyecto de digitalización, debe tenerse muy claro el objetivo, hay que analizar con cuidado el proceso, planear con todo detalle la implementación y se deben determinar con precisión los costos del proyecto, ya que la digitalización supone un trabajo intenso y que puede llegar a ser muy caro, si no se planea con cuidado, siendo muy importante capturar una imagen de modo que sea posible utilizarla para satisfacer diferentes necesidades.

Para empezar es importante entender el significado de la terminología básica y establecer los conceptos y aspectos a analizar para poder resolver la problemática que se presenta por lo que de inicio se hace una breve descripción de dicha terminología, conceptos y

aspectos, con el fin de que se pueda desarrollar de la mejor manera posible el proceso de digitalización.

4. DESARROLLO DEL TEMA

4.1. Marco Jurídico

Los principios fundamentales del Buen Gobierno señalados por la Organización para la Cooperación y el Desarrollo Económico (OCDE), el Banco Mundial y el Banco Interamericano del Desarrollo (BID), se fundamentan en la transparencia del quehacer de las instituciones públicas y en la rendición de cuentas respecto del manejo y aprovechamiento cabal de los recursos de que disponen, de su aplicación para el cumplimiento de objetivos y metas institucionales, así como de los resultados alcanzados durante un periodo determinado.

Señalan que dichos principios facilitan la mejora del compromiso gubernamental ante la ciudadanía, ya que se crea y fomenta una cultura de consultas de información - proceso que no debe generar burocracia - y diálogo, refuerzan las políticas públicas en cualquiera de los niveles de gobierno y suponen la obligación de quien ejerce un poder, de responsabilizarse de la labor que tiene a su cargo, de someterse a evaluaciones de su desempeño y a dar a conocer los resultados de esas evaluaciones.

Por tanto, la transparencia y la rendición de cuentas fortalecen la relación entre la autoridad pública y ciudadanía, la cual, cada vez más requiere y exige conocer y participar de las decisiones gubernamentales, así como vigilar y evaluar la actuación de los servidores públicos y de las instituciones (Testigos Sociales y Contraloría Social).

En el contexto del Buen Gobierno y en el marco del artículo 6º de la Constitución Política de los Estados Unidos Mexicanos, la Federación dio cabida a la aplicación de los principios de transparencia y de rendición de cuentas, promulgando en 2002, la Ley Federal de Transparencia y Acceso a la Información Pública Gubernamental, así como creando El Instituto Federal de Acceso a la Información, como órgano autónomo para regular y vigilar el cumplimiento de la citada Ley.

A partir de ese hecho, las administraciones Estatales y los poderes Judicial y Legislativo paulatinamente, han venido asumiendo el citado precepto constitucional, en cuanto a que el derecho a la información será garantizado por el Estado, emitiendo sus respectivos ordenamientos legales en materia de Transparencia y Acceso a la Información Pública.

Con fecha 6 de marzo de 2007, la Cámara de Diputados aprobó en lo general y en lo particular, con 425 votos a favor y 1 abstención, el proyecto de decreto que reforma el Artículo 6º Constitucional, en su primera lectura, exceptuándolo de la segunda y turnado a la Cámara de Senadores para los efectos constitucionales.

El proyecto de decreto fue promovido a través de la "Iniciativa Chihuahua", encabezada por los Estados de Chihuahua, Zacatecas, Aguascalientes, Veracruz y el Distrito Federal, el cual, en el marco del Pacto Federal, resulta trascendental en materia de Transparencia y Rendición de Cuentas, ya que versa sobre la obligación de construir instituciones y leyes coherentes para la Federación, Entidades Federativas y Municipios.

Dicho proyecto da cabida a los principios básicos indispensables en la materia, a fin de que se establezcan a nivel constitucional, las garantías mínimas que implican el ejercicio del derecho al acceso a la información pública por parte de la ciudadanía, a partir de las

cuales los gobiernos estatales y municipales deberán alinear sus respectivos ordenamientos relativos al tema.

La exposición de motivos señala entre otros aspectos, que la transparencia y acceso a la información pública de los Estados, es parte medular de un federalismo más eficiente, con más obligaciones a las que deben corresponder mayores recursos y el ejercicio presupuestal debe estar sellado por la transparencia y acceso a la información; por tanto debe ser evaluado y vigilado por la ciudadanía.

El citado proyecto de decreto refiere la inclusión de un segundo párrafo y VII fracciones en el artículo 6^o Constitucional, así como tres artículos transitorios; se cita:

Artículo 6^o.....

"Para el ejercicio del derecho de acceso a la información, la Federación, los Estados y el Distrito Federal, en el ámbito de sus respectivas competencias, se regirán por los siguientes principios y bases:

1. Toda la información en posesión de cualquier autoridad, entidad, órgano y organismo federal, estatal y municipal, es pública y sólo podrá ser reservada temporalmente por razones de interés público en los términos que fijen las leyes. En la interpretación de este derecho deberá prevalecer el principio de máxima publicidad.
11. La información que se refiere a la vida privada y los datos personales será protegida en los términos y con las excepciones que fijen las leyes.
111. Toda persona, sin necesidad de acreditar interés alguno o justificar su utilización, tendrá acceso gratuito a la información pública, a sus datos personales o a la rectificación de éstos.
- IV. Se establecerán mecanismos de acceso a la información y procedimientos de revisión expeditos. Estos procedimientos se sustanciarán ante órganos u organismos especializados e imparciales, con autonomía operativa, de gestión presupuestaria y de decisión.
- V. Los sujetos obligados deberán preservar sus documentos en archivos administrativos actualizados y publicarán a través de los medios electrónicos disponibles, la información completa y actualizada sobre sus indicadores de gestión y el ejercicio de los recursos públicos.
- VI. Las leyes determinarán la manera en que los sujetos obligados deberán hacer pública la información relativa a los recursos públicos que entreguen a personas físicas o morales.
- VII. La inobservancia a las disposiciones en materia de acceso a la información pública será sancionada en los términos que dispongan las leyes.

TRANSITORIOS

Primero. El presente Decreto entrará en vigor al día siguiente de su publicación en el Diario Oficial de la Federación.

Segundo. La Federación, los Estados y el Distrito Federal, en sus respectivos ámbitos de competencia, deberán expedir las leyes en materia de acceso a la información pública y transparencia, o en su caso, realizar las modificaciones necesarias, a más tardar un año después de la entrada en vigor de este Decreto.

Tercero. La Federación, los Estados y el Distrito Federal deberán contar con sistemas electrónicos para que cualquier persona pueda hacer uso remoto de los mecanismos de acceso a la información y de los procedimientos de revisión a los que se refiere este Decreto, a más tardar en dos años a partir de la entrada en vigor del mismo. Las leyes locales establecerán lo necesario para que los municipios con población superior a setenta mil

habitantes y las demarcaciones territoriales del Distrito Federal cuenten en el mismo plazo con los sistemas electrónicos respectivos".

Por tanto, la propuesta de reforma Constitucional, se centra en que se plasmen las obligaciones mínimas e iguales de transparencia, a ser cumplidas por todos los sectores públicos del país: federal, estatal y municipal; que todo ciudadano tenga el mismo derecho sin importar el propósito de la solicitud; que utilice los mismos mecanismos y medios expeditos para acceder a la información pública en cualquier lugar donde se encuentre dentro del territorio nacional; que los gobiernos dentro del plazo señalado en el proyecto de decreto, cuenten con los medios sistematizados para que la información se solicite y proporcione vía electrónica.

Ahora bien, en este marco, el 5 de septiembre de 2003, se publicó en el Diario Oficial de la Federación (DOF), en el rubro Poder Judicial, Consejo de la Judicatura Federal, los "Lineamientos de la Comisión para la Transparencia y Acceso a la Información del Consejo de la Judicatura Federal, de los Tribunales de Circuito y los Juzgados de Distrito, relativos a los criterios de clasificación y conservación de la información reservada y confidencial, para este órgano del Poder Judicial de la Federación, los Tribunales de Circuito y los Juzgados de Distrito".

En estos lineamientos se establecen:

- Los criterios de clasificación de la información reservada.
- Los criterios de clasificación de la información confidencial.

Todo lo anterior hace evidente la necesidad de digitalizar las sentencias judiciales, para la rendición de cuentas y la utilización de la propia información, en cumplimiento del Artículo 6º Constitucional.

Ajeno a este trabajo está el describir las modificaciones que el Poder Judicial de la Federación ha tenido desde 1995, por las que ahora está conformado por la Suprema Corte, el Consejo de la Judicatura Federal y el Tribunal Federal Electoral. Baste con decir que el manejo del Archivo histórico del Poder, de 1815 en que se estableció el Primer Tribunal Federal, en Ario de Rosales, Michoacán, por José María Morelos y Pavón, hasta 2003, está a cargo de archivística judicial de la Suprema Corte y que los principales

generadores de archivo en tránsito y vivo es el Consejo de la Judicatura Federal con sus más de 600 juzgados y tribunales Federales en todo el país.

Con este marco, a continuación se presentará la problemática relativa a la dimensión del archivo histórico judicial (1815-2003).

4.2. Planteamiento del Problema

El Archivo Judicial bajo resguardo de la Suprema Corte de Justicia de la Nación constituye el patrimonio histórico judicial del país, al resguardar el devenir y evolución del Poder Judicial de la Federación. El reto de mantener organizado dicho acervo, integrado por poco más de 98 kilómetros de documentación (más de 7 millones de expedientes), y actualizarlo permanentemente es evidentemente una tarea altamente compleja.

Actualmente, dicho acervo se encuentra resguardado en los depósitos de la Suprema Corte de Justicia de la Nación en el Distrito Federal, en los dispuestos en las Casas de Cultura Jurídica y en el Centro Archivístico Judicial de Toluca, Estado de México; se compone de expedientes judiciales resueltos desde 1825 (y algunos de años anteriores, prácticamente desde 1815), por el Pleno y las Salas de la Suprema Corte y por los Tribunales Colegiados de Circuito, Tribunales Unitarios de Circuito y Juzgados de Distrito. Esta documentación es consultada principalmente por los miembros del Poder Judicial de la Federación, así como por la sociedad en general, a través de los módulos de acceso a la información, de conformidad con la normativa en la materia, a los que acuden investigadores, abogados, historiadores, archivistas y estudiantes de diversas disciplinas.

En razón de sus dimensiones, y sus rápido crecimiento como resultado de la intensa actividad jurisdiccional de los Tribunales y Juzgados Federales, los procesos para su traslado, resguardo y consulta física desde diversos puntos del país son complejos, lo que provoca que la consulta física sea escasa en proporción a su extensa conformación y distribución a lo largo del país: esto sin contar los problemas de espacio en los depósitos documentales.

Los archivos judiciales que resguarda el Máximo tribunal implican 98 kilómetros de documentos y más de 7 millones de expedientes. De éstos, los Archivos Judiciales de los Juzgados de Distrito representan el 58%, los Archivos Judiciales de los tribunales Colegiados de Circuito el 27%, los Archivos Judiciales de los tribunales Unitarios de Circuito el 9% y los Archivos Judiciales de la Suprema Corte de Justicia de la Nación el 6%.

Los archivos judiciales del propio Alto Tribunal están en la situación siguiente:

Periodo	Metros de expedientes res guardados	Total de expedientes res uardados	Metros de expedientes di italizados	Total de expedientes di italizados	Porcentaje del acervo di italizado
Siglo XIX y anteriores	324	90 000	324	90 000	
Siglo XX y XXI	4,876	760,000	4,000	727,000	

TOTALS	5 200	850 000	4 324	817 000
---------------	-------	---------	-------	---------

Los archivos judiciales de los Tribunales Colegiados de Circuito, de los Tribunales Unitarios de Circuito y de los Juzgados de Distrito están en la siguiente situación:

Periodo	Metros de expedientes res guardados	Metros de expedientes di italizados	Total de expedientes di italizados	Porcentaje del acervo di italizado
1950 y años anteriores	5 000	250	39 000	5%
Si lo XX XXI	88,000		570 000	
TOTALS	93 000	4 250	609 000	

El Máximo Tribunal cuenta con cuarenta y cuatro áreas de depósito documental dispuestas en todo el país:

ARCHIVO JUDICIAL	
Su rema Corte de Justicia de la Nación	2 áreas de depósito ubicadas en su edificio sede en el anexo de Bolivar 30,
Tribunal de Circuito y Juzgados de Distrito en el Primer Circuito Judicial	3 áreas de depósito ubicadas en el edificio de los Tribunales y Juzgados Federales de San Lázaro en el Reclusorio Norte Oriente.
Tribunales de Circuito y Juzgados de Distrito Foráneos	38 áreas de depósito en las Casas de la Cultura Jurídica.
Tribunales de Circuito y Juzgados de distrito de todo el país.	1 Centro Archivístico Judicial en Toluca, Estado de México,

El problema reside en la necesidad de conservar la información con trascendencia documental e histórica, agilizar la gestión de la documentación, así como la óptima utilización de los recursos para su administración y resguardo, de archivos muy especializados como son los archivos judiciales es decir: "Conservar para difundir".

4.3. TERMINOLOGÍA BÁSICA DE LA DIGITALIZACIÓN

Señales analógicas VS Señales digitales

"Señales analógicas: existente en el mundo real, con características que la identifican, tales como amplitud y valores continuos.

Señales digitales: empleadas por los sistemas electrónicos en informáticos, con características distintivas tales como frecuencia y valores discretos.

La necesidad de traducir el mundo real (analógico) para ser tratado por las computadoras (digital) originó métodos de conversión, también conocido como digitalización, es decir, pasar de señales analógicas a señales digitales manipuladas por los sistemas".¹

Una imagen digital puede ser representada mediante una matriz bidimensional de números.

C Los elementos para digitalizar una señal analógica radican en un convertidor analógico-digital que opera a través de muestreo, cuantificación y codificación.

- "Muestreo: Toma muestras periódicas de la señal (samplig).
- Cuantificación: Se mide cada una de las muestra tomadas.
- Codificación: Traduce las muestras a valores discretos. (0 y 1)".²

Imágenes digitales.

Una imagen digital es una "fotografía electrónica" convertida en un conjunto de elementos pictóricos (píxeles) y ordenada de acuerdo con una relación predefinida de columnas y filas. El número de píxeles en una tabla dada define la resolución de la imagen. Cada pixel tiene un determinado valor tonal que depende de la intensidad de la luz reflejada desde el documento 0 original al dispositivo de carga acoplado (CCC), mediante diodos sensibles a la luz. Cuando los documentos se exponen a la luz crean una carga eléctrica proporcional, que genera, a través de una conversión analógica/digital, una serie de señales digitales representadas por un código binario. La unidad más pequeña de información almacenada en un ordenador se denomina bit (acrónimo de binary digit). El número de bits utilizados para representar cada pixel en una imagen determina el número de colores o escala de grises que se pueden o representar en una imagen digital. Esto se llama profundidad del bit.

Calidad de la imagen.

La calidad de la imagen durante la captura depende de la suma de resultados de la resolución 0 aplicada al escaneo, la profundidad del bit de la imagen escaneada, los procesos de mejora y el nivel de compresión aplicada, el dispositivo de escaneo utilizado o técnicas usadas, y la preparación del operador del escáner.

e Resolución

La resolución es la frecuencia espacial a la cual se realiza la muestra de una imagen digital. La medida de la resolución se determina por el número de píxeles leídos en una distancia lineal de una pulgada (2.54 cm) en el documento digitalizado.

Las unidades más utilizadas son "puntos por pulgadas" -ppp- (dots per inch - dpi-), o "píxeles por pulgadas" -ppp (pixels per inch - ppi -). Las resoluciones por digitalización de documentos están normalmente en el rango de 50 a 600 ppp.

Las resoluciones de escaneo más frecuentemente utilizadas en documentos blanco y negro en SGED son 200, 300 y 400 ppp. En la digitalización de imágenes color de alta calidad, son típicas resoluciones de 1,200 a 2,400 ppp.

¹ SAMBARINO, Julia. "Digitalización: Concepto Preliminares" Escuela Superior de Cómputo del Instituto Politécnico Nacional (IPN), México, 2009. P.3. Idem

Dimensiones de/ pixel

Las dimensiones del pixel son las medidas horizontales y verticales de una imagen, expresadas en píxeles. Las dimensiones de pixel se pueden determinar multiplicando tanto el ancho como la altura por el dpi. Una cámara digital también tendrá dimensiones de pixel, expresadas como la cantidad de píxeles en forma horizontal y en forma vertical que definen su resolución (por ejemplo: 2.048 por 3.072). El dpi se calcula dividiendo las dimensiones de un documento por la dimensión de pixel correspondiente respecto de la cual se encuentra alineado.

Ejemplo:

Dimensión de pixel: Un documento de 8 x 10 pulgadas que se escanea a 300 dpi posee dimensiones de pixel de 2400 píxeles (8 pulgadas x 300 dpi) por 3000 píxeles (10 pulgadas x 300 dpi).

Profundidad de bits.

La profundidad del bit es la medida del número de bits utilizados para definir cada pixel.

A mayor profundidad de bits, pueden representarse mayor número de tonos en grises y color.

- escaneo binario, utilizando un bit por pixel para representar el blanco o el negro, • Escaneo en escala de grises, utilizando múltiples bits por pixel para representar las escalas de grises. El mejor nivel de escala de grises es de ocho bits por pixel, y en este nivel la imagen que se muestra puede seleccionar hasta 256 diferentes niveles de grises.
- Escaneo en color utilizando múltiples bits por pixel para representar el color; se considera un buen nivel de color 24 bits por pixel y permite una selección de hasta 16.7 millones de colores.

La elección de la profundidad de los bits afecta a las posibilidades de captura, tanto de la apariencia física del documento original como de su contenido informativo. De este modo, las decisiones sobre la profundidad de los bits deben tener en cuenta si el aspecto físico del documento, o de sus partes, tienen un valor informativo añadido que es necesario reflejar. Este puede ser el caso de los proyectos de digitalización que tienen como objetivo la realización de facsímiles de los documentos originales.

Profundidad de bits: De izquierda a derecha - imagen bitonal de 1 bit, a escala de grises de 8 bits, y a color de 24 bits.

Cálculos binarios para la cantidad de tonos representados por profundidades de bits comunes:

1 bit (2^1) = 2 tonos

2 bits (2^2) = 4 tonos

3 bits (2^3) = 8 tonos

4 bits (2^4) = 16 tonos

8 bits (2^8) = 256 tonos

16 bits (2^{16}) = 65.536 tonos

24 bits (2^{24}) = 16,7 millones de tonos

Rango dinámico.

Rango dinámico es el rango de diferencia tonal entre la parte más clara y la más oscura de una imagen. Cuanto más alto sea el rango dinámico, se pueden potencialmente representar más matices, a pesar de que el rango dinámico no se correlaciona en forma automática con la cantidad de tonos reproducidos. Por ejemplo, el microfilm de alto contraste exhibe un rango dinámico amplio, pero presenta pocos tonos. El rango dinámico también describe la capacidad de un sistema digital de reproducir información tonal. Esta capacidad es más importante en los documentos de tono continuo que exhiben tonos que varían ligeramente, y en el caso de las fotografías puede ser el aspecto más importante de la calidad de imagen.

Tamaño de archivo.

El tamaño del archivo se calcula multiplicando el área de superficie (altura x ancho) de un documento a ser escaneado, por la profundidad de bits y el dpi.

Debido a que el archivo de imagen se representa en bytes, que están formados por 8 bits, se divide esta cifra entre 8.

Fórmula 1 para e/ tamaño de archivo

Tamaño de archivo = (altura x ancho x profundidad de bits x dpi²) / 8

Si se proporcionan las dimensiones de pixel, éstas se multiplican entre sí y por la profundidad de bit para determinar la cantidad de bits presentes en un archivo de imagen. Por ejemplo, si se captura una imagen de 24 bits con una cámara digital con dimensiones de pixel de 2.048 x

3.072, entonces el tamaño de archivo es igual a (2048 x 3072 x 24) / 8, 0 50.331.648 bytes.

Fórmula 2 para e/ tamaño de archivo

Tamaño de archivo = (dimensiones de pixel x profundidad de bits) / 8

Sistema convencional para dar nombres a los archivos según el tamaño de los mismos: Debido a que las imágenes digitales tienen como resultado archivos muy grandes, la cantidad de bytes con frecuencia se representa en incrementos de 2¹⁰ (1.024) o más:

1 Kilobyte (KB) = 1.024 bytes

1 Megabyte (MB) --- 1.024 KB

1 Gigabyte (GB) --- 1.024

MB 1 Terabyte (TB) = 1.024

GB

Compresión.

La compresión se utiliza para reducir el tamaño del archivo de imagen para su almacenamiento, procesamiento y transmisión. El tamaño del archivo para las imágenes digitales puede ser muy grande, complicando las capacidades informáticas y de redes de muchos sistemas. Todas las técnicas de compresión abrevian la cadena de código binario en una imagen sin comprimir, a una forma de abreviatura matemática, basada en algoritmos complejos. Existen técnicas de compresión estándar y otras patentadas. En general es mejor utilizar una técnica de compresión estándar y ampliamente compatible, antes que una patentada, que puede ofrecer compresión más eficiente y/o mejor calidad, pero que puede no prestarse a un uso o a estrategias de preservación digital a largo plazo.

En la comunidad de las bibliotecas y los archivos hay un fuerte debate acerca del uso de la compresión en archivos maestros de imágenes.

Los sistemas de compresión también pueden caracterizarse como sin pérdida o con pérdida. Los sistemas sin pérdida, como ITU-T.6, abrevian el código binario sin desechar información, por lo que, cuando se "descomprime" la imagen, ésta es idéntica bit por bit al original. Los sistemas con pérdida, como JPEG, utilizan una manera de compensar o desechar la información menos importante, basada en un entendimiento de la percepción visual. Sin embargo, puede ser extremadamente difícil detectar los efectos de la compresión con pérdida, y la imagen puede considerarse "sin pérdida visual". La compresión sin pérdida se utiliza con mayor frecuencia en el escaneado bitonal de material de texto. La compresión con pérdida típicamente se utiliza con imágenes tonales, y en particular imágenes de tono continuo en donde la simple abreviatura de información no tendrá como resultado un ahorro de archivo apreciable.

Formatos de/ archivo

Los formatos de archivo consisten tanto en la configuración de bits que comprende la imagen como la información del encabezamiento acerca de cómo leer e interpretar el archivo. Los formatos de archivo varían en términos de resolución, profundidad de bits, capacidades de color, y soporte para compresión y metadatos.

Los principales formatos gráficos de imagen son:

Bitmap (Windows Bitmap File) .bmp

Creado por Microsoft Windows Paint. Fichero sin compresión, típicamente de 8 bits, soportando paletas de hasta 24 bits. Gran tamaño de fichero, lo que le convierte en un formato poco manejable y raramente utilizado en sistemas de gestión electrónica de documentos (SGED).

TIFF (Tagged Image File Format) .tif .tiff

Desarrollado por Aldus Corp. en 1986 para guardar imágenes desde escáneres y tarjetas capturadoras de video. TIFF 3 Y TIFF 4 asume las especificaciones establecidas por CCITT Grupo III y Grupo IV respectivamente para imágenes bitonales (1 bit). Posteriormente TIFF 5 Y TIFF 6 asumen imágenes con escala de grises (4 u 8) bits) y paleta de colores (normalmente 24 bits, ampliable hasta 64 bits). Se ha convertido en un estándar de facto en los SGED. Compresión sin pérdida ITU-T6/LZW y con pérdida JPEG. Precisa de conexión externa para su ejecución en Web.

GIF (Graphic Interchange Format) .gif

Fue desarrollado por el intercambio de imágenes de CompuServe. Bitonal, escala de grises o color de 1 a 8 bits. Compresión sin pérdida LZW, bajo licencia de Unisys. Características añadidas en versión GIF 89^a: carga progresiva, máscara de transparencia de 1 bit y animación simple. Utilización generalizada y estándar en Web con soporte directo a partir de Explorer 3 y Navigator 2.

JPEG (Join Photographic Expert Group) .jpeg .jpg

O Diseñado para este grupo para la compresión de imágenes fotográficas. Formato de 8 bits en escala de grises y 24 bits en color, con alta compresión y pérdida variable. Estándar para la

representación de imágenes bitonales o con pequeño número de colores). JPEG intenta eliminar la información que el ojo humano no es capaz de distinguir. Utilización generalizada y estándar en Web con soporte directo a partir de Explorer 2 y Navigator 2.

JFIF (JPEG File Interchange Format) .jfif .jif

Presentada por C-Cube Microsystems. Formato de archive que permite que los flujos de datos JPEG sean intercambiados entre una amplia variedad de plataformas y aplicaciones. Soporte e Web directo a partir de Explorer 2 y Navigator 2.

Photo CD .pdc

e Presentado por Kodak en 1992. Formato color de 24 bits concebido para ser creado a partir de cualquier negativo de 35 mm (color o b&n) o diapositiva color. Define 6 resoluciones de tamaño entre 192 x 128 pixeles y 6,144 x 4,096 pixeles, con o sin pérdidas. Precisa de aplicación Java para su ejecución Web.

PNG (portable Network Graphics) .png

Fuer diseñado para reemplazar el formato GIF, y ha sido aprobado por W3C (World Wide Web Consortium). Profundidad de bits entre 1 y 48 (hasta 16 en escala de grises). Compresión sin pérdida, un 10% superior a la del formato GIF. Visualización progresiva en 2 direcciones. No permite animación. Utilización directa en Web con soporte a partir de Explorer 4 y Navigator

4.04

PDF (Portable Document Format) .pdf

Lenguaje descriptor de páginas desarrollado por Adobe Corp. que contiene las capacidades gráficas de PostScript (menos flexible, pero más eficiente) y con capacidades hipertextuales. Profundidad de bits variable (4 a 8 bits escala de grises y hasta 64 bits color). Descompresión sin pérdida ITU T6 ó LZW, o con pérdida JPEG. Formato propietario pero estándar de facto. e Precisa de conexión o aplicación externa para su ejecución en Web.

Selección del material a digitalizar.

Esta es una de las partes más importantes del proceso de digitalización y hay que elegir con cuidado los materiales que van a ser digitalizados y para ello se deben de tener en cuenta los siguientes puntos:

Atributos del documento

¿El material se presta para la digitalización? ¿Puede capturarse el contenido informativo de una manera apropiada en forma digital? ¿Los formatos físicos y la condición del material o constituyen impedimentos graves? ¿El material intermedio, tal como microfilm o diapositivas, se encuentra disponible y en buenas condiciones? ¿Cuál es el tamaño y la complejidad de la colección, en lo que respecta a variedad de documentos?

Consideraciones acerca de la preservación

¿se pondrá en peligro el material durante el proceso de digitalización? ¿Los sustitutos digitales reducirían el uso de los originales, de tal modo ofreciéndoles protección contra la manipulación? ¿se considera a la reproducción digital como un medio para reemplazar los originales?

Organización y documentación disponible

¿se encuentra el material en un orden coherente, estructurado de manera lógica? ¿se encuentra paginado o su disposición está indicada por algún otro medio? ¿Está completo? ¿Existe información descriptiva, de navegación o estructural adecuada sobre el material, como por ejemplo registros bibliográficos o asistencia de búsqueda detallada?

Usos previstos

¿Qué clases, nivel y frecuencia de uso se prevén? ¿Existe un entendimiento claro de las necesidades del usuario? ¿Puede la digitalización soportar estos usos? ¿El acceso al material será considerablemente mayor como consecuencia de la digitalización? ¿Puede su institución soportar una variedad de usos, por ejemplo, impresión, navegación, revisión detallada? ¿Existen cuestiones sobre seguridad o acceso que deban tenerse en cuenta (por ejemplo, acceso restringido a determinadas personas o uso de acuerdo con determinadas condiciones)

Incremento de la colección digitalizada

¿Existe un incentivo adicional para digitalizar el material, basado en la disponibilidad de recursos digitales complementarios (incluyendo datos y metadatos)? ¿Existe la oportunidad de obtener una cooperación multi-institucional? ¿La de crear una coherencia temática o una "masa crítica"?

Duplicación del esfuerzo

¿El material ha sido digitalizado anteriormente por otra fuente confiable? Si así fuera, ¿los archivos digitales son de una calidad, documentación, y funcionalidad suficientes para servir a sus fines? ¿Qué condiciones regulan el acceso y uso de esos archivos?

Capacidad institucional

¿La institución posee la infraestructura técnica necesaria para la gestión, entrega, y mantenimiento de los materiales digitalizados? ¿sus principales usuarios poseen recursos de informática y de conectividad apropiados para utilizar de manera eficaz estos materiales?

Recursos financieros

¿se puede determinar el costo total de la adquisición de la imagen (selección, preparación, captura, indexación, y control de calidad)? ¿Este costo se encuentra justificado en base a los beneficios reales o percibidos que devenguen de la digitalización? ¿Existen fondos para respaldar este esfuerzo? ¿Existe un compromiso institucional para gestionar y preservar en forma continuada estos archivos?

Conversión.

La captura de imágenes digitales debe tomar en cuenta los procesos técnicos comprendidos al convertir una representación analógica en digital, así como también los atributos de los documentos fuente en sí mismos: dimensiones físicas y presentación, nivel de detalles, rango tonal, y presencia de color. Los documentos también se pueden caracterizar por el proceso de producción utilizado para crearlos, incluyendo medios manuales, mecánicos, fotográficos, y, últimamente, electrónicos. Además, todos los documentos con formato de papel y película están comprendidos en una de las siguientes cinco categorías, las que afectarán su grabación digital.

Tipos de documentos

- Texto impreso / Dibujos de líneas simples representación en base a bordes definidos, sin variación de tono, como un libro que contiene texto y gráficos de líneas simples.
- Manuscritos — representaciones en base a bordes suaves que se producen a mano o a máquina, pero no exhiben los bordes definidos típicos de los procesos a máquina, como el dibujo de una letra o una línea.
- Media Tinta — reproducción de materiales gráficos o fotográficos representados por una cuadrícula con un esquema de puntos o líneas de diferente tamaño y espaciadas regularmente que, habitualmente se encuentran en un ángulo. También incluye algunos tipos de arte gráfica, como por ejemplo, los grabados.
- Tono Continuo — elementos tales como fotografías, acuarelas y algunos dibujos de líneas finamente grabadas que exhiben tonos que varían suave o sutilmente.
- Combinado — documentos que contienen dos o más de las categorías mencionadas anteriormente, como por ejemplo, los libros ilustrados.

Reconocimiento óptico de caracteres (OCR)

Para obtener una imagen del documento a digitalizar se emplean dispositivos ópticos como los escáneres.

El reconocimiento óptico de caracteres (OCR) es un método para reconocer textualmente una imagen digitalizada. El OCR recibe la imagen digitalizada y esto es un archivo que puede ser editado y usado como tal por cualquier programa o aplicación que lo necesite.

"Todos los algoritmos de reconocimiento óptico de Caracteres (OCR) persiguen la finalidad de poder un texto para poderlo tratar posteriormente.

Sin embargo existen algunas variantes de estos, como son:

- ICR (intelligent carácter recognition) se usa para designar el reconocimiento de caracteres manuscritos.
- OCV (Optical carácter verification) Hace referencia a la verificación de contenidos previamente conocidos.
- OMR (Optical mark recognition) reconocimiento de marcas".²

Los algoritmos de reconocimiento óptico de caracteres se basa en la "binarización" fragmentación de imagen, adelgazamiento de los componentes y compresión de patrones.

En esta última etapa destacan varios métodos para el reconocimiento de caracteres basados en la programación dinámica: estadísticos, estructurales, de redes neuronales, markorianos, de lógica borrosa entre otros.

Control de calidad.

El control de calidad (QC = quality control) es un componente esencial de un programa de digitalización de imágenes y tiene como fin asegurar que se han cumplido las expectativas

² Dbidem p.5

en cuanto a calidad. El mismo abarca procedimientos y técnicas para verificar la calidad, precisión y consistencia de los productos digitales. Las estrategias de control de calidad pueden ser implementadas en diferentes niveles:

- Evaluación inicial

Se utiliza un subconjunto de documentos (a ser convertidos en la empresa o por un proveedor de servicios) para verificar que las decisiones técnicas tomadas durante la evaluación de referencia son las apropiadas. Esta evaluación ocurre con anterioridad a la implementación del proyecto.

- Evaluación continuada

El mismo proceso de garantía de calidad utilizado para confirmar las decisiones de la evaluación de referencia puede ser ampliado y extendido a la totalidad de la colección para asegurar la calidad de todo el programa de digitalización de imágenes.

Los siguientes pasos resumen los puntos principales de un programa de control de calidad.

a. Identificación de los productos

El primer paso es identificar claramente los productos a ser evaluados. Los mismos pueden

O incluir metadatos imágenes complementarios, originales incluyendo y derivadas, texto impresiones, convertido y bases archivos de marcados.datos de imágenes y

b. Desarrollo de un enfoque sistemático

Para medir la calidad y juzgar si los productos son satisfactorios o no, hay que definir claramente las características que permitan clasificar los productos digitales como "aceptables" ó "inaceptables".

c. Determinación de un punto de referencia

¿con qué se está comparando las imágenes al juzgarlas? No siempre es sencillo responder esta pregunta. Por ejemplo, si la conversión está basada en un intermedio, la imagen digital se encuentra a dos "generaciones" de distancia respecto del original. La misma ha sido copiada a película (primera generación), la cual posteriormente es escaneada (segunda generación). ¿cuál debería ser el punto de referencia al valorar una imagen de tales características, el documento original o la transparencia? La inspección de calidad de la imagen, ¿estará centrada en el original o en el derivado (o en ambos)?

d. Definición del alcance y los métodos

Hay que determinar el alcance de la revisión de calidad. ¿se inspeccionarán todas las imágenes, o solamente un subconjunto de prueba (por ejemplo el 20%)? Hay que definir tanto la metodología como el modo en que se realizarán los juicios sobre calidad. Por ejemplo, ¿se evaluarán las imágenes en forma visual a una ampliación del 100% (1:1) en la pantalla y se les comparará con los documentos originales? ¿O la evaluación estará basada sólo en una valoración subjetiva de las imágenes en la pantalla, sin remitirse a los originales?

e. Control del entorno del control de calidad

Por lo general se subestima el impacto que producen las condiciones de visualización de imagen sobre la percepción de la calidad. En un entorno inadecuado, incluso una imagen

de alta calidad puede ser percibida como no satisfactoria. Por ejemplo, una imagen a color de 24 bits puede verse "posterizada" en exceso cuando es utilizada en una computadora configurada de manera incorrecta, que no puede proporcionar una paleta de colores completa.

Factores que afectan la calidad de la imagen en la pantalla son:

Configuración de hardware

Es difícil recomendar una configuración de hardware ideal. La regla general es armar un sistema que pueda satisfacer las necesidades de velocidad, memoria, almacenamiento y calidad de presentación. ¿Qué clase de imágenes se están creando? ¿Qué cantidad? ¿con qué fines? ¿Qué nivel de revisión en pantalla se necesita? Será necesaria una computadora rápida y confiable con una amplia capacidad de procesamiento y memoria para poder recuperar y manipular los grandes archivos que se están creando, especialmente al crear imágenes en color.

Software de recuperación de imágenes

Hay que utilizar un software de recuperación apropiado para las imágenes.

O Condiciones de visualización

Hay que controlar el entorno de visualización. Se debe comprender que el monitor y el documento fuente requieren condiciones de visualización diferentes. Se podrá visualizar mejor el original en un ambiente con mucha luz y el monitor trabaja mejor en un medio con poca luz. Sin embargo, un entorno con poca luz no se equipara a un cuarto oscuro. Vista en la oscuridad, una imagen en pantalla parecería carecer de suficiente contraste.

Características humanas

La evaluación de la calidad de la imagen requiere sofisticación visual, especialmente para las evaluaciones subjetivas. Idealmente, la misma persona debe evaluar todas las imágenes, con el mismo equipo y bajo los mismos parámetros. El personal debe ser entrenado en particular en lo que se refiere a cómo transmitir en forma efectiva la información sobre la apariencia del color. Algunas deficiencias de visión del color están relacionadas con un gen recesivo

defectuoso en el cromosoma X. Siendo que las mujeres tienen dos cromosomas X y los hombres uno, la probabilidad de tener visión del color defectuosa es de 1 en 250 para las mujeres, y de 1 en 12 para los hombres. Aún entre personal experto en evaluaciones visuales, no es raro que se presenten diferencias en los resultados que se deban a variaciones normales del ojo humano. Se puede utilizar una buena prueba de visión del color para evaluar la visión de un individuo.

Calibración de monitor

Las imágenes pueden aparecer de manera diferente en los distintos monitores. La calibración es el proceso de ajuste de la configuración de la conversión de color del monitor a un nivel estándar, de manera que la imagen se presenta de igual manera en diferentes monitores. El método ideal es utilizar hardware de calibración de monitor y el software adjunto. Sin embargo, si no se tiene acceso a estos recursos, se pueden utilizar las herramientas de calibración de los programas de aplicación.

Gestión de color

Uno de los principales desafíos en la digitalización de documentos a color es mantener la apariencia y consistencia del color en toda la cadena de digitalización, incluyendo el escaneado, la visualización, e impresión. La reproducción precisa de los colores es difícil debido a que los dispositivos de entrada y los de salida tratan a los colores de manera diferente. El objetivo del software de sistema de gestión de color (CMS) es asegurar que los colores del original coincidan con la mayor precisión posible con la reproducción digital en la pantalla o impresa.

Cadena de digitalización

La tecnología necesaria para navegar desde un extremo de la cadena de digitalización al otro consta principalmente de: hardware, software y redes.

La cadena de digitalización y la infraestructura técnica que la sostiene se dividen en tres componentes fundamentales: creación, gestión y entrega.

La Creación de imágenes se ocupa de la captura o conversión inicial de un documento u objeto a la forma digital, por lo general con un escáner o cámara digital. A la imagen inicial se le pueden aplicar uno o más pasos de procesamiento de archivo o de imagen, que pueden alterar, agregar o extraer datos. Las clases generales de procesamiento incluyen la edición de la imagen (escalarla, comprimirla, otorgarle nitidez, etc.) y la creación de metadatos.

Manejo de archivos se refiere a la organización, almacenamiento y mantenimiento de imágenes y metadatos relacionados.

La Entrega de la imagen comprende el proceso de hacer llegar las imágenes al usuario y abarca redes, dispositivos de visualización e impresoras.

Las computadoras y sus interconexiones de red son componentes integrales de la cadena de digitalización. Cada eslabón de la cadena comprende una o más computadoras y sus diversos componentes (RAM, CPU, bus interno, tarjetas de expansión, soporte de periféricos, dispositivos de almacenamiento y soporte de red). Los requisitos de configuración cambiarán, dependiendo de las necesidades informáticas específicas de cada componente.

Existe una amplia selección de dispositivos que comienzan la cadena de digitalización, y se utilizará el término escáner para referirse a todos los dispositivos de captura de imágenes, incluyendo las cámaras digitales.

Algunas preguntas que son fundamentales para seleccionar el tipo de escáner son las siguientes:

- ¿Es este escáner compatible con los documentos que van a ser digitalizados? ¿Puede manejar la variedad de tamaños, tipos de documentos (hojas simples, volúmenes encuadernados), medios (reflectivos, transparentes), y la condición de los originales? • ¿Puede este escáner producir la calidad requerida para satisfacer las necesidades? Siempre es posible obtener una imagen de calidad inferior partiendo de una de calidad superior, pero por más magia digital que se utilice, no se podrán restablecer con exactitud detalles que, para empezar, nunca se han capturado. Entre los factores a

considerar se incluyen la resolución óptica (opuesta a la interpolada), la profundidad de bits, el rango dinámico, y la relación señal-ruido.

- ¿Puede este escáner soportar el programa de producción y presupuesto de conversión? (Hay que prestar atención a las afirmaciones de rendimiento con frecuencia un factor de suma importancia en el costo del escáner). ¿Cuáles son las capacidades de manejo de documentos del mismo? ¿su ciclo de servicio, MTBF (Mean Time Between Failure - Tiempo Promedio Entre Fallas) y capacidad de vida útil? ¿Qué tipo de contratos de mantenimiento están disponibles (en las instalaciones, reemplazo durante las 24 horas, servicio de depósito)?

Los escáneres funcionan iluminando el objeto o documento a ser digitalizado y dirigiendo la luz reflejada (por lo general a través de una serie de espejos y lentes) sobre un elemento fotosensible. En la mayoría de los escáneres, el medio sensible es un circuito electrónico integrado sensible a la luz conocido como un dispositivo acoplado cargado (CCC). Los fotositos sensibles a la luz dispuestos a lo largo del CCD convierten los niveles de brillo en señales electrónicas que luego se procesan en una imagen digital.

CCD es, sin la menor duda, la tecnología de sensibilidad a la luz que más comúnmente se utiliza en los escáneres modernos. También existen otras dos tecnologías, CIS (Contact Image Sensor - Sensor de Imagen de Contacto) y PMT (photomultiplier tube tubo fotomultiplicador) que se encuentran en los extremos inferior y superior del mercado de escáneres, respectivamente. CIS es una tecnología más reciente que permite que los escáneres sean más pequeños y livianos, pero sacrifica el rango dinámico, la profundidad de campo y la resolución. Los escáneres de tambor de base PMT producen imágenes de una muy alta calidad, pero tienen una aplicación limitada en el escaneado para bibliotecas y archivos.

Otra tecnología de sensibilidad, CMOS (Complementary Metal Oxide Semiconductor - Semiconductor de óxido de Metal Complementario), aparece principalmente en las cámaras digitales de mano de gama baja, en las cuales su bajo costo, bajo consumo de energía y su fácil integración de componentes permiten diseños más pequeños y económicos.

Hay varios tipos de escáneres.

Los escáneres planos son el tipo de escáner más conocido y vendido. Son versátiles, fáciles de manejar, y con una amplia disponibilidad. Su popularidad para la publicación en la Web abrió un gran mercado, forzando a la baja los precios de las unidades. En el otro extremo, las unidades profesionales para el mercado de gráficos ahora compiten con los escáneres de tambor en cuanto a calidad. Todos utilizan la misma tecnología básica, en la cual un sensor de luz (por lo general un CCD) y una fuente de luz, ambos montados sobre un brazo móvil, pasan sobre el documento, que está fijo sobre una placa de vidrio. Algunos modelos poseen manipuladores de documento automáticos (ADH), que pueden aumentar el rendimiento y disminuir la fatiga del operador en el caso de grupos de documentos uniformes que se encuentran en condiciones razonablemente buenas. Una variante especializada del escáner plano es el escáner de libros de trayectoria aérea, en el cual la fuente de luz, la selección de sensores y la óptica son trasladados a un ensamble de brazo de trayectoria aérea bajo el cual puede colocarse un volumen encuadernado con las hojas hacia arriba, para ser escaneado.

Los escáneres con alimentador de hojas utilizan la misma tecnología básica que los escáneres planos, pero maximizan el rendimiento, por lo general a expensas de la calidad. Diseñados generalmente para entornos de negocios de grandes volúmenes, típicamente escanean en blanco y negro o escala de grises con resoluciones relativamente bajas. Se espera que los documentos sean de un tamaño uniforme y con una solidez suficiente para soportar una manipulación bastante brusca, aunque los mecanismos de transporte de algunos modelos más nuevos reducen la tensión. Ya sea que utilicen un transporte de rodillos, cinta, tambor o de vacío, el sensor de luz y la fuente de luz permanecen fijos mientras que se mueve el documento. Una subclase de escáneres con alimentador de hojas son los modelos de pie específicamente diseñados para los documentos de gran formato, como los mapas y los planos arquitectónicos.

Los escáneres de tambor producen escaneados con la mayor resolución y calidad que cualquier otro tipo de escáner, pero esto tiene su precio. Además de su costo, los escáneres de tambor son lentos, no son indicados para documentos de papel quebradizo y requieren un alto nivel de habilidad por parte del operador. Por eso típicamente se los encuentra en agencias de servicios que satisfacen las necesidades del mercado de pre-impresión a color.

Los escáneres para microfilm son dispositivos altamente especializados para digitalizar películas en rollo, microfichas y tarjetas de apertura. Puede ser difícil obtener una calidad buena y consistente en un escáner para microfilm debido a que los mismos pueden tener un funcionamiento complejo, la calidad y condición de la película puede variar, y debido a que ofrecen capacidad de mejora mínima. Sólo unas pocas empresas fabrican escáneres para microfilm, y la falta de competencia contribuye al alto costo de estos dispositivos.

Procesamiento de imágenes / archivos

Luego del escaneado hay una variedad de pasos de procesamiento. Tales procedimientos pueden ocurrir en cualquier punto de la cadena de digitalización, desde inmediatamente luego del escaneado hasta justo antes de la entrega a los usuarios finales. Los mismos pueden ser modificaciones personalizadas que sólo afectan a ciertos archivos, o procesamiento automatizado masivo de todos los archivos (procesamiento por lotes). Pueden ser operaciones realizadas por única vez o realizadas repetidas veces a medida que se las necesita.

Algunos ejemplos de operaciones de procesamiento de imágenes / archivos son los siguientes:

- Edición, retoque, mejora incluye pasos tales como eliminación de muaré (descreening), eliminación de puntos (despeckling), eliminación de oblicuidad (deskewing), aumento de nitidez (sharpening), utilización de filtros personalizados y ajuste de profundidad de bits. En algunos casos el software de escaneado realiza estos pasos. En otros, se utilizan herramientas de edición de imágenes separadas (por ejemplo: Adobe Photoshop, Corel Photo Paint, Image Magic).
- Compresión - algunas veces llevada a cabo por el firmware dedicado del escáner o hardware dedicado de la computadora. La compresión también puede ser una operación sólo de software, a pesar de que el hardware dedicado es más rápido y

se lo debería considerar cuando se crean archivos muy grandes o grandes cantidades de archivos.

- Conversión de formato de archivo el escaneado original puede no estar en un formato adecuado para todos los usos previstos, por lo que requiere conversión.
- Escala es probable que los escaneados capturados a alta resolución no sean adecuados para la visualización en la pantalla. Con frecuencia se necesita aplicar escala (es decir, reducción de resolución a través de eliminación de bits) para poder crear imágenes para entrega en la Web.
- OCR (reconocimiento óptico de caracteres)- conversión de texto escaneado a texto legible por medio de una máquina, que se puede buscar o indexar.
- Creación de metadatos - agregado de texto que ayuda a describir, rastrear, organizar o mantener una imagen.

Manejo de archivos.

El manejo de archivos consiste en una serie de pasos interrelacionados, diseñados para asegurar la fácil identificación, organización, acceso y mantenimiento de los archivos. Dado que hay fuertes conexiones entre los diversos aspectos de la gestión de archivos, hay que planear muy bien para evitar tomar decisiones que limiten las opciones posteriormente. Es en especial importante mantener las líneas de comunicación abiertas entre el personal técnico y el personal del proyecto durante la etapa de planeación.

Los pasos del manejo de archivos incluyen:

- Seguimiento (consideraciones básicas del sistema de archivos). Otro aspecto del seguimiento está comprendido en Metadatos;
- Bases de datos de imágenes y otras soluciones de gestión de imágenes (software especial para organizar archivos de imágenes);
- Almacenamiento (dispositivos y medios);
- Mantenimiento (copias de seguridad -backup-, migración, preservación y seguridad)

Los sistemas de asignación de nombres de archivos y directorios por omisión son rara vez óptimos para una colección específica. Las decisiones sensatas acerca de los archivos y los directorios pueden ayudar a minimizar el caos, en especial en el caso de colecciones muy e grandes. Hasta cierto punto, la naturaleza del material que se está escaneando sugerirá los principios de organización. Las series con frecuencia se dividen en volúmenes y números, las monografías tienen números de página, las colecciones de manuscritos o de fotografías tienen números de carpeta o de acceso, etc. En la mayoría de los casos, algún aspecto de estos principios de organización física pueden traducirse a organización de sistemas de archivos.

Algunas recomendaciones básicas acerca de los sistemas de archivos son las siguientes:

- Utilizar un sistema de asignación de nombres que sea compatible con cualquier sistema operativo y medio de almacenamiento que se planea utilizar;
- Utilizar extensiones de archivo estándar para los distintos tipos de archivos;
- No sobrecargar los directorios con demasiados archivos;
- Confiar en el software de gestión de almacenamiento para manejar grandes colecciones a través de múltiples unidades físicas de disco;

- Permitir que la colección pueda crecer en grandes cantidades.

Muchas de las primeras iniciativas digitales se basaban en la programación personalizada para manejar grandes colecciones de archivos de imágenes. Las rutinas para el procesamiento por lotes, la organización y entrega de archivos, se escribían utilizando lenguajes de texto de alto nivel, como por ejemplo Perl y Tcl. Hoy en día existen muchos productos ya armados que pueden simplificar radicalmente el proceso de manejar una gran colección de archivos de imágenes. Sin embargo, aún el sistema más simple requiere algún tipo de personalización. Las colecciones más grandes y aquellas con metadatos complejos requieren herramientas más sofisticadas, que a su vez necesitan un mayor grado de mantenimiento y supervisión por parte del personal. Por ende, la experiencia en programación es una habilidad buscada en el personal que maneja bases de datos de imágenes.

Las bases de datos de imágenes varían significativamente en cuanto a la facilidad de uso y al nivel de funcionalidad. Realizan un seguimiento de sus archivos, proporcionan funciones de búsqueda y recuperación, suministran una interfaz de acceso, controlan el nivel y tipo de uso, y proporcionan algo de seguridad al controlar quién tiene acceso a qué. Ninguna herramienta

tiene elegido posibilidades más cuidadosamente de satisfacer necesita todas las ser necesidades, reevaluado e en incluso forma el regular conjunto para de determinar herramientassi aún sigue siendo la mejor elección.

Los criterios generales para evaluar las bases de datos de imágenes incluyen los siguientes:

- Objetivo para el cual se creó la colección digital;
- Tamaño y tasa de crecimiento de la colección digital;
- Complejidad y volatilidad de los metadatos complementarios;
- Nivel de demanda y de rendimiento esperado;
- Infraestructura técnica existente, incluyendo disponibilidad del personal de sistemas capacitado; • Gasto.

Categorías básicas de los sistemas de bases de datos

Las siguientes son las categorías fundamentales:

Las bases de datos de escritorio comunes son medianamente económicas y simples de usar, pero limitadas en cuanto a tamaño y funcionalidad.

Las aplicaciones de bases de datos cliente-servidor son más costosas y sofisticadas que las bases de datos de escritorio, y también son más difíciles de utilizar y mantener.

Los sistemas especializados de gestión de imágenes pueden ofrecer una completa solución ya e armada, con estructuras de datos predefinidas, pero son más costosos y menos flexibles en términos de personalización y compatibilidad.

Típicamente, el componente de la infraestructura técnica al que se le presta mayor atención e es el dispositivo de captura, debido a que interactúa directamente con el objeto tangible que

se digitaliza y tiene la mayor influencia sobre la calidad y fidelidad de la imagen resultante. Se presta mucha menos atención al medio de almacenamiento en el cual residirán los bits capturados. Esto es una pena, dado que las malas elecciones en cuanto a la tecnología de almacenamiento pueden ser perjudiciales para cada etapa de la digitalización y pueden tener como resultado retrasos, entrega ineficiente, costos innecesarios a corto y a largo plazo y corrupción y pérdida de datos.

La reticencia a centrarse en la tecnología de almacenamiento es comprensible. Los aparatos de almacenamiento llevan a cabo una función utilitaria y de rutina dentro de la cadena de digitalización y es fácil no darles el valor que realmente tienen. Además, el almacenamiento masivo es una de las tecnologías informáticas más competitivas y que más rápido avanzan. Como resultado, aún para quienes poseen habilidad técnica, puede ser muy desalentador seguirle el ritmo al panorama de almacenamiento siempre cambiante, y más aún comprender algunos de sus aspectos más complejos. Excepto en el caso de instalaciones relativamente pequeñas, las decisiones respecto de la tecnología de almacenamiento probablemente se tomen previa consulta con el personal de sistemas. Para que esa relación de consulta sea una asociación efectiva, el conocimiento de la terminología y los conceptos básicos son la base para realizar las preguntas correctas.

Los criterios generales para la evaluación incluyen:

- Velocidad (lectura / escritura, transferencia de datos);
- Capacidad;
- Fiabilidad (estabilidad, redundancia);
- Estandarización;
- Costo;
- Aptitud para la tarea.

Los rápidos cambios en la tecnología del almacenamiento han alterado el impacto de estos criterios sobre la planeación de la digitalización. A principios de la década de 1990 el almacenamiento era costoso, lento y tenía una capacidad relativamente limitada. Los proyectos con archivos de imágenes de muchos gigabites experimentaron varias tecnologías de disco óptico nuevas (y con frecuencia patentadas) para encontrar medios asequibles de proteger sus nuevos tesoros digitales, con frecuencia sacrificando velocidad y fiabilidad en el proceso.

En la actualidad, la unidad de disco magnético giratorio es el líder indiscutido del almacenamiento. En casi todos los proyectos más ambiciosos, la fase de producción de la digitalización está basada en unidades ATA paralelas comunes y poco costosas, actualmente disponibles en capacidades de varios terabites por unidad y velocidades de transferencia de interfaz de hasta 133 MB por segundo. Es muy probable que ni la velocidad ni la capacidad creen cuellos de botella en lo que respecta al rendimiento.

Hoy en día, hay más posibilidades de que el desafío del almacenamiento surja en la etapa de entrega, desde los esfuerzos por consolidar distintas colecciones digitales en una gran biblioteca digital que muchas veces puede contener terabites de datos (un terabite son 1000 gigabites). La gestión, entrega y mantenimiento de tales colecciones en forma eficiente no es una tarea trivial y el precio de grandes selecciones de almacenamiento con una alta fiabilidad, excelente rendimiento y facilidades de copia de seguridad (backup)

integradas puede incluso ejercer demasiada presión sobre los presupuestos. Las colecciones más pequeñas que tienen una gran demanda también pueden requerir sistemas de almacenamiento de rendimiento superior.

Dentro de la gama de tecnologías de almacenamiento disponibles, por lo general es más seguro elegir una que esté en su punto máximo de popularidad y aceptación o muy cerca del mismo. Las tecnologías demasiado cercanas a la vanguardia pueden no lograr nunca el apoyo generalizado por parte de los fabricantes o de los usuarios, dejando a quienes las adoptan desde un comienzo con hardware o medios huérfanos y no compatibles. Las tecnologías demasiado cercanas a la retaguardia pueden sufrir disminución de soporte de producto y tener una menor cantidad de métodos de actualización. Asimismo, no hay que comprar una cantidad sustancialmente mayor de almacenamiento de la que se estima que será necesaria hasta dentro de un par de años. El almacenamiento infrutilizado no es redituable, especialmente debido a la rápida disminución en el precio y la expectativa de una vida útil relativamente corta. En la actualidad, la mayoría de los sistemas de almacenamiento están diseñados para ajustarse a un crecimiento en aumento.

Las tecnologías de almacenamiento masivo se pueden clasificar de distintas maneras. El sistema de almacenamiento subyacente (magnético, óptico o magnetóptico), el tipo de unidad (fija o removible), el material del medio (cinta, disco rígido, disco flexible) y la interfaz de hardware (ATA, ATAPI, SCSI, USB, Fireware / IEEE 1394, Canal de Fibra) en forma conjunta definen las características de cada tecnología.

Los sistemas de almacenamiento también se distinguen en almacenamiento de conexión directa o almacenamiento conectado a la red. El almacenamiento de conexión directa incluye unidades de escritorio estándar que se instalan dentro de un gabinete de computadora o se cablean directamente al mismo. El almacenamiento conectado a la red por lo general abarca almacenamiento accesible a múltiples computadoras y que puede estar conectado a un servidor y se puede acceder a él por medio de protocolos de sistema de archivos especiales (por ejemplo: Sistema de Archivo de Red o Sistema de Archivo Común de Internet) o puede ser parte de un sistema de almacenamiento que funciona en forma independiente de cualquier servidor en particular (por ejemplo, una Red SAN Red de Área de Almacenamiento).

Desde la invención del disco rígido en 1952, el avance tecnológico regular y veloz ha dado lugar a mejoras increíbles en cuanto a capacidad, velocidad, fiabilidad y relación precio-rendimiento. La fuerza propulsora de estas mejoras ha sido el crecimiento constante de la cantidad de datos que se pueden almacenar en la misma área (conocida como "densidad de área"). El costo unitario del almacenamiento básico en disco rígido disminuyó aproximadamente un cien por ciento entre 1997 y 2002 y sigue disminuyendo a pasos agigantados. Se prevé que el costo por unidad de almacenamiento continuará disminuyendo en forma abrupta y que la capacidad de unidad continuará aumentando, de modo que es muy poco probable que incluso las colecciones de imágenes digitales más grandes y de crecimiento más rápido tengan problemas de capacidad o asequibilidad en el almacenamiento masivo. Otras formas de almacenamiento masivo, tales como los sistemas de disco óptico y de cinta magnética, están viendo mejoras en cuanto al precio y el rendimiento, pero a una velocidad inferior a la del disco magnético.

El inconveniente de un cambio tecnológico tan veloz es la también rápida obsolescencia. La necesidad de reemplazar los sistemas de almacenamiento a intervalos de tiempo

cortos (quizás cada 3 a 5 años) anulan algunos de los beneficios relacionados con los costos. Los presupuestos de mantenimiento de los sistemas de digitalización de imágenes deberían anticipar estas necesidades.

Otro inconveniente es la confusa proliferación de nuevas tecnologías. Esto es particularmente cierto en dos áreas. Una es la de las interfases de hardware para los discos magnéticos. Para aprovechar la densidad de almacenamiento reciente (y los consecuentes aumentos en la velocidad de recuperación de datos) se deben desarrollar nuevas interfases de hardware que puedan estar a la altura de las unidades. De lo contrario, las unidades más rápidas no tendrían ventajas.

El resultado ha sido una intensa competencia por aumentar la velocidad de flujo de datos que pueden manejar las interfases, en la que las partes interesadas en las interfases intentan superar a los demás y ganar una porción más grande del mercado de las aplicaciones de alto rendimiento. Entre los ejemplos de esto podemos nombrar el cambio de USB 1.1 a 2.0, la introducción regular de nuevos estándares SCSI y los cambios inminentes de IEEE 1394a a 1394b y del ATA paralelo al ATA serial. Las nuevas versiones ofrecen un rendimiento superior pero pueden causar problemas tales como incompatibilidades (con dispositivos de versiones anteriores y con el sistema informático mismo), falta de soporte del sistema operativo y un retraso en la disponibilidad de los drivers del dispositivo.

La otra área en donde la proliferación de la tecnología ha provocado confusión y dolores de cabeza a los usuarios es el área de los formatos de los medios de disco compacto. Esto es especialmente cierto en lo que respecta a los formatos de DVD de alta densidad, en donde compiten por lo menos cinco formatos distintos, incluyendo tres formatos regrabables (DVDRAM, DVD+RW y DVD-RW). La falta de estandarización da lugar a incompatibilidades entre las unidades y los medios y dificulta que los usuarios se decidan por una de las tecnologías

Los requisitos de capacidad pueden ser estimados por medio de un simple cálculo:

Almacenamiento total necesario = cantidad de archivos de imágenes x tamaño de archivo promedio x 1,25

Ejemplo: Una colección de 3000 imágenes de texto de aproximadamente 75KB cada una, requerirá cerca de 225MB de almacenamiento. Sin embargo, muchos otros factores pueden aumentar las necesidades de almacenamiento. Un texto OCR (reconocimiento óptico de caracteres) para las mismas páginas, podría necesitar 3KB por página, a alrededor de 1/25 del espacio requerido para el archivo de imagen correspondiente. La cantidad y tamaño de los archivos derivados, así como también si se los almacena en forma permanente o se los crea a las carreras también podría agregar más a los requisitos de almacenamiento. Además, todas las tecnologías de almacenamiento comprenden cierta cantidad de espacio desperdiciado. La cantidad precisa depende de factores como por ejemplo la tecnología de almacenamiento utilizada, la capacidad total, el tamaño de la partición y el tamaño de archivo promedio. Se pueden necesitar algunos experimentos para determinar el porcentaje aproximado de espacio desperdiciado, pero se lo debe tener en cuenta al estimar las necesidades de almacenamiento. La fórmula anterior es un factor, en una gran mayoría, para cubrir dichas preocupaciones.

La entrega comprende los procesos de hacer llegar las imágenes digitales y los archivos auxiliares a sus usuarios. Los componentes más importantes son redes y dispositivos de visualización (principalmente monitores e impresoras). Esta es la etapa de la cadena en la

cual conocer a los usuarios se vuelve al menos tan importante como conocer los documentos.

A menos que las imágenes digitales sean estrictamente para uso dentro de la empresa, algunos componentes de la entrega están fuera de control. Por ejemplo, si la mayoría de los usuarios están conectados a Internet con módems de 56Kbps, una colección de preciosas imágenes a color de 24 bits, con un tamaño promedio de 500KB y en la cual cada una de ellas tarda más de dos minutos en bajar, frustrará a los usuarios.

La entrega exitosa a una audiencia combinada de usuarios dentro de la empresa y fuera de la misma requerirá una cuidadosa planeación realizada con anterioridad. Si los recursos lo permiten, el mejor acercamiento es ofrecer múltiples versiones de imágenes, aprovechando la capacidad superior cuando exista, pero también soportando las conexiones de ancho de banda bajo con imágenes de calidad inferior. Hay que tener cuidado con el enfoque del "común denominador más bajo", que puede parecer igualitario, pero en última instancia priva a los usuarios con mejores equipos del valor potencial de sus imágenes.

Las decisiones acerca de los formatos de archivo, las relaciones de compresión, y la aplicación de escalas tendrán un impacto sobre la entrega.

Las redes son probablemente la parte menos visible de la infraestructura técnica. Las tarjetas de red están escondidas dentro de las computadoras; el hardware de redes se guarda en cuartos de máquinas o "gabinetes" de comunicaciones; y los cables se ocultan bajo tierra, en las paredes y / o van en forma aérea. Pero nada puede detener una iniciativa de digitalización de imágenes de manera más rápida que una red de un tamaño más pequeño de lo necesario, demasiado lenta o no confiable. Se tiene necesidad de redes veloces y confiables para poder transportar los archivos durante la creación y gestión de archivos. Una colección de imágenes digitales muy utilizada ejercerá una mayor exigencia sobre su red.

Las decisiones respecto de la infraestructura de redes por lo general se toman en el nivel institucional. Las grandes instituciones anticipan el crecimiento de las necesidades generales de redes y están preparadas para manipular volúmenes significativos de tráfico en la red. Las pequeñas instituciones podrían descubrir que una iniciativa de digitalización de imágenes impone exigencias a una red existente, las cuales tienen implicaciones para toda la organización. Incluso la limitación del uso de ciertas redes de gran intensidad a horarios tradicionales de poco tráfico puede interferir con otras actividades. Al comienzo de la etapa de planeación se debería discutir con los administradores de redes acerca de las demandas de las redes que se proyectan.

Una organización que ha utilizado su red principalmente para el e-mail y en parte para la navegación en la Web puede descubrir que su conexión a Internet es completamente inadecuada para ofrecer grandes volúmenes de imágenes digitales. La mayoría de las conexiones a Internet son asimétricas, permitiendo que más datos pasen hacia el usuario (desde Internet) que hacia la red (hacia Internet). Una conexión de Internet que permite que grandes volúmenes de datos pasen hacia la red puede ser muy costosa.

La Compatibilidad todavía puede ser un problema en algunas instituciones. A pesar de que el protocolo de comunicaciones de Internet TCP/IP (Protocolo de Control de Transmisión / Protocolo de Internet) se ha vuelto indispensable, en algunas instituciones

todavía se utilizan protocolos heredados. Hay que verificar con los administradores de red para asegurarse de que sus planes son compatibles con la red existente.

La Fiabilidad es otra inquietud. Los cortes de la red reducen la productividad y frustran a los usuarios. A pesar de que la fiabilidad del hardware de red ha mejorado en los últimos años, todavía ocurren fallas. Algunas redes más antiguas han "crecido como locas" y son una

mezcla de diferentes tecnologías, cableados y hardware. La responsabilidad dividida de la administración de la red también debilita la fiabilidad.

La Seguridad de Internet es una preocupación creciente. Los servidores de imágenes están sujetos a violaciones a la seguridad, haciendo peligrar potencialmente el acceso a los usuarios autorizados, o dejando a los datos vulnerables ante una posible eliminación o modificación realizada en forma intencional. Los administradores de sistemas y de redes pueden proponer remedios tales como firewalls (cortafuegos), software de control especial, o el requisito de autenticación de todos los usuarios. Algunas medidas de seguridad pueden ser onerosas, ya sea porque requieren personal especializado para mantenerlas o porque restringen el acceso a sus materiales más de lo que se quisiera.

El Gasto puede o no ser un asunto significativo. En algunas instituciones, simplemente se conectará el equipo a las redes existentes y se estará en camino. Pero si la tarea requiere una actualización fundamental de la red o un nuevo tipo de conexión a Internet, el gasto puede ser muy importante. La clave es probar, y puede ser necesario que se tenga que retroceder y ajustar los planes a algo que pueda realizarse dentro de la infraestructura existente.

Los asuntos de velocidad y de capacidad están determinados por una cantidad de factores. Algunos están dentro de control, otros no. Al igual que respecto de tantos otros asuntos de rendimiento, el evitar los cuellos de botella es un objetivo importante. La transmisión por red está regulada por el enlace más lento. Los factores que afectan la entrega por red incluyen:

- Capacidad de transporte (ancho de banda) de la red de área local;
- Ancho de banda de la conexión a Internet de la institución;
- Velocidad y capacidad del servidor de red;
- Tasa de velocidad de lectura y transferencia de datos de los dispositivos de almacenamiento;
- Tamaño de archivo de imagen;
- Demanda de usuarios en un momento dado;
- Cantidad de tráfico que compita en la red (en todos los niveles de red);
- Velocidad de cualquier paso del procesamiento "l a las carreras"; • Tiempo requerido para autenticación y otros chequeos de seguridad; • Capacidades de la computadora del usuario final, incluyendo:
 - Velocidad de la CPU; o Cacheo (caching) de Ram / disco; o Rendimiento del subsistema de video; o Velocidad de la conexión a Internet.

Los esfuerzos continúan aumentando el rendimiento de las redes. Una estrategia es aumentar la velocidad en las redes existentes a través de nuevas formas de compresión

o presentación. Sin embargo, la necesidad de reducir el tamaño de archivo para aumentar la velocidad de la entrega puede ser un asunto de tiempo limitado dado que los conductos de información de banda ancha y las capacidades de transferencia de datos a altas velocidades en forma inalámbrica se están desarrollando actualmente para respaldar la investigación, comercio electrónico y el entretenimiento. La creciente utilización de servicios de cable módem y DSL para uso residencial facilitará las inquietudes acerca del ancho de banda en el extremo del usuario. El potencial de la televisión digital, en especial la Televisión de Alta Definición (HDTV) para proporcionar nuevos y diferentes tipos de información a una gran gama de usuarios -incluyendo el acceso a recursos culturales digitalizados- está en pleno desarrollo. Las normas Comisión Federal de Comunicaciones de los Estados Unidos (FCC) exigieron que todas las emisiones analógicas fueran retiradas hacia fines de 2007. Comenzando con Internet 2, el gobierno de los Estados Unidos está activamente financiando esfuerzos para construir la Internet de la Próxima Generación (NGI) para unir los laboratorios de investigación y las universidades a redes de alta velocidad que son entre 100 a 1000 veces más rápidas que la Internet actual. Diseñada para manejar grandes volúmenes de información, la NGI facilita el acceso a las imágenes digitales y vuelve práctico el audio de alta calidad y la transferencia de imágenes en movimiento.

Si los dispositivos de almacenamiento se encuentran entre las tecnologías de más rápida evolución, los monitores se encuentran entre las más lentas.

Un monitor será como la ventana del usuario hacia su colección de imágenes digitales. Como en el caso de las redes, algunas veces el monitor está bajo su control, otras no. Cuando lo está, entonces la oportunidad es minimizar los aspectos en los que hay que ceder inherentes a la tecnología actual de los monitores. Además de elegir un producto de calidad, las características del tipo de configuración de la resolución, calibración, luz externa, e incluso con qué frecuencia se limpia la pantalla, pueden afectar la calidad percibida de la imagen.

Los determinantes en la calidad de la imagen en los monitores son:

- Resolución de la pantalla
- Tamaño de la pantalla
- Espacio entre puntos
- Velocidad de actualización
- Profundidad de bits
- Rendimiento del monitor
- Rendimiento de la tarjeta de video

Obviamente, el dispositivo de mejor visualización no puede corregir problemas de imagen que resulten del uso de equipos inadecuados o de malas decisiones tomadas en los pasos de captura o de procesamiento. Las limitaciones en el manejo del color de los sistemas operativos y del software de visualización de imágenes (en particular los navegadores Web) también pueden afectar la imagen final.

Sin embargo, suponiendo que se ha hecho un esfuerzo considerable en la captura de la imagen, es sensato elegir un dispositivo de visualización que realce las imágenes para lograr el mejor efecto. No todos los monitores CRT se crean de igual modo. Entre los

monitores CRT, por ejemplo, el diseño de máscara de sombras se luce para el caso de texto, mientras que los diseños con rejilla de apertura producen mejores imágenes a pesar de que las delgadas líneas horizontales cerca de la parte superior y la parte inferior de la pantalla pueden ser molestas.

Cuando se adquiere un equipo, las pantallas de 17 pulgadas deberían ser consideradas el mínimo para la mayoría de los objetivos de visualización, los monitores de 19 pulgadas se han vuelto muy populares y se empiezan a utilizar ampliamente los monitores de 21 pulgadas que aceptan 1920 x 1440 píxeles, aunque muchas veces el tamaño del monitor está limitado por el espacio disponible en los escritorios. Por el lado de la producción, los monitores más grandes reducen la fatiga visual del personal que realiza el trabajo de control de calidad.

Mientras que las computadoras sean grandes, los dispositivos de visualización tengan baja resolución y sean incómodos para la vista, la tecnología de batería esté en su etapa inicial, y la infraestructura de comunicaciones esté unida por cables, el deseo de crear impresiones de las imágenes digitales perdurará. Sin embargo, no deberían subestimarse los costos de hacer que imágenes de alta resolución realmente estén disponibles online, en formatos que se puedan imprimir por medio de una cantidad de plataformas y una variedad de impresoras. Antes de realizar promesas de entregar imágenes de calidad en forma impresa en un entorno de red, hay que verificar que la infraestructura técnica esté al nivel de la tarea, y hay que considerar los costos de almacenamiento adicional asociados con el acceso online.

Hoy en día, la impresión en blanco y negro está dominada por dos tecnologías: las impresoras a inyección de tinta (chorro de tinta), que echan chorros de tinta líquida sobre el papel a través de pequeños inyectores; y las impresoras láser, que utilizan una fuente de luz para crear cargas en un tambor fotoconductor, permitiéndole atraer partículas de tinta seca (toner) que se funden en el papel. Las impresoras a inyección de tinta se han vuelto muy económicas, pero son más lentas que las impresoras láser y por lo general no están diseñadas para la impresión de grandes volúmenes. Las impresoras láser de producción de gama alta pueden producir bastante más de 100 páginas por minuto a 600 dpi.

Ambas tecnologías han sido adoptadas para el color. Las impresoras a inyección de tinta a color vienen en modelos de 3 ó 4 colores. Las impresoras láser a color son mucho más costosas, tanto por el precio inicial como por los insumos. Las impresoras a color, ya sean a inyección de tinta o láser, son sustancialmente más lentas que sus equivalentes en blanco y negro. Las impresoras a color a inyección de tinta imprimen en promedio 5 páginas de texto e por minuto y 1 página de gráficos de página completa por minuto. Las impresoras láser a color son más rápidas, con un promedio de 12 páginas de texto por minuto y 2 páginas de gráficos de página completa por minuto.

Se encuentran disponibles muchas otras tecnologías para la impresión a color. Las mismas incluyen sublimación de tintura, tinta sólida y cera térmica.

La sublimación de tintura es en especial significativa dado que puede producir impresiones de color verdadero de tono continuo, a pesar de ser extremadamente lentas y requerir un papel recubierto especial.

Los principales criterios de evaluación para seleccionar las impresoras adecuadas son los siguientes:

- Resolución y espaciado de los puntos;
- Reproducción del color;
- Representación tonal;
- Capacidades de mejoras de la imagen;
- Tamaño de documento soportado;
- Comparación entre la impresión a simple faz o a doble faz (simple /doble);
- Medios soportados (papel liso, papel recubierto, transparencias, sobres);
- Velocidad y capacidad;
- Idiomas de descripción de página y formatos de imagen sin tratamiento soportados;
- Capacidades de red;
- Costo.

Proceso de Administración Electrónica de Documentos

- a. Clasificación y etiquetado. Los documentos en papel, son recibidos y etiquetados con una etiqueta de código de barra que identifica de manera única a dicho documento (RUT).
- b. Digitalización. Los documentos ya etiquetados son digitalizados y creados como archivos electrónicos con el número de identificador (RUT).
- c. Control de Calidad. Los documentos digitalizados, corresponden a una "foto" del documento en cuestión y como tal debe ser evaluada su calidad de imagen, resolución, etc., no siendo posible, en esta fase acceder al "contenido" o lo "escrito dentro del documento". Los documentos que son rechazados por calidad insuficiente, son re-solicitados para digitalización, en caso contrario pasa a la fase siguiente.
- d. Bodega. Los documentos digitalizados que pasan la prueba de calidad son notificados para su almacenamiento, o para que continúen con el flujo normal de operación que posean.
- e. Indexación Complementaria. Los documentos digitalizados, que han pasado la fase de calidad, son procesados, convirtiendo las "fotografías" o imágenes escaneadas de los papeles originales, en texto útil para efectos de búsqueda e indexación. En este punto se logra EXTRAER desde las imágenes de los papeles, el texto que efectivamente se encuentra escrito en ellos. Metadatos.

Biblioteca Institucional. Los documentos convertidos, esto es, Imagen + Texto, son cargados a la Biblioteca Institucional y al Gestor Documental, haciendo posible la búsqueda de información no solo por sus referencias, sino también por el contenido o texto que se encuentre escrito en dichos papeles.

4.4. SOLUCIÓN DE LA PROBLEMÁTICA PLANTEADA

Objetivo General/

Conservar los expedientes judiciales y la documentación administrativa que constituyan el verdadero patrimonio nacional en materia de administración de justicia, en razón de su relevancia jurídica y documental, mediante su óptima organización, clasificación, catalogación y digitalización, en su caso, a fin de contribuir a difundir su consulta en forma

ágil y eficaz, transparentando así, el ejercicio del servicio público que brindan el Alto Tribunal, los Tribunales y Juzgados Federales.

Alcance.

Juzgados de Distrito

Durante el año 2009 y 2010, llevar a cabo la transparencia, valoración, catalogación y depuración de los expedientes generados por los Juzgados de Distrito de los años 1951 a 2003.

En el año 2010, llevar a cabo la catalogación y digitalización de los expedientes históricos de los años 1950 y anteriores.

Pleno y Salas de la Suprema Corte

Respecto de los expedientes generados en el Pleno y las Salas de la Suprema Corte, en el año 2010 se concluirá la catalogación y digitalización de los expedientes.

La publicación de los expedientes digitalizados del siglo XX se llevará a cabo en el año 2009; la digitalización de los expedientes del siglo XIX y XXI se efectuará entre los años 2010 y 2011.

En cuanto a la digitalización de dichos expedientes, se estima que en el año 2012 se podrá concluir el proceso.

En cuanto al acervo histórico generado por los Tribunales de Circuito de los años 1950 y anteriores, en el año 2010, se llevará a cabo la catalogación. La digitalización se tiene prevista efectuarse en los años 2010 y 2011.

El acervo a procesar está integrado por expedientes jurídicos, con las siguientes características:

- Tamaño: carta 20%, oficio 75%, otros 5% (Entre los otros se consideran tarjetas de tamaños varios, planos arquitectónicos, papeles media carta, doble carta, doble oficio, fotografías, revistas, etc)
- Tipo de papel usado: Bond 75gr; 58%; cebolla 19%, cartulinas y cartoncillo para portadas 20%, varios (periódicos y papeles diversos) 3%.
- Los colores más usados en los materiales mencionados son: Blanco en un 80%, amarillo 7%, cremas 5% y varios en un 8% (Rojo, azul o morado). • Impresión :
- En máquina de escribir 20%, impresoras laser (toner) e inyección de tinta 45%, copias papel carbón 20%, varios (offset, manuscritos e impresión de matriz de puntos) 15%.
 - Los tipos de tinta usados, son: De resinas plásticas 60%, de carbón 30%, varias (ferrosas, de aceite y vegetales) el 10%. ○ Modo impresión: a dos caras 90%, a una cara 10%.
 - Claridad escritura: Buena 65%, Regular 25% y Mala 10%. ○ Tintas: Negra 81%, azul 11% y roja y verde 8%.

Descripción	Cantidades	Estado de los Expedientes

La documentación generada por los Juzgados de Distrito.	Aproximadamente 23,850 metros lineales de los expedientes depurados y catalogados, generados por los Juzgados de Distrito de 1951 a 2003.	Integro en un 95% y el 5% restante presenta deterioros con o rasgaduras, enmendaduras con cinta adhesiva u hojas rotas.
Acervo histórico generado por los órganos jurisdiccionales.	Aproximadamente 4,991 metros lineales de los expedientes catalogados, generados en 1950 y años anteriores.	Integro en un 90% y el restante presenta deterioros como rasgaduras, enmendaduras con cinta adhesiva u hojas rotas.
Documentación generada por el Pleno y Salas. (SCJN)	Aproximadamente 1,524 metros lineales de los expedientes generados por el Alto Tribunal de los siglos XIX y XXI.	Integro en un 95% y el 5% restante presenta deterioros como rasgaduras, enmendaduras con cinta adhesiva u hojas rotas.

- Composición de los expedientes:
 - El 68% de los expedientes contienen: Sección Principal, Sentencias y Actuaciones posteriores.
 - El 32% solo Sección principal.
- Contenido de hojas: de 5 a 30,000, promedio 200.
- Folios: no tienen y/o no son confiables.
- Hojas para identificación y separación de secciones: Se manejan identificadores impresos para el reconocimiento físico y por software de los documentos.
- Los identificadores se deben imprimir en papel Bond de 75gr.
 - 0 10% de expedientes requieren dos identificadores.
 - 0 45% de expedientes requieren tres identificadores.
 - 0 45% de expedientes requieren cuatro identificadores.
- Estado físico de documentos: Normal 40%, Reseco 10%, Húmedo 5%, Rotos 15%, con polvo 20%, Contaminación de Hongos 5% e incompletos 5%.

Plan de Digitalización

De acuerdo a los estándares, el estimado de volúmenes de documentación considerados para procesar.

	Concepto	Año 2009		
		Octubre	Noviembre	Diciembre
Juzgados de Distrito 1951 a 2003	Expedientes/ Mensual	54,560	54,670	54,560
	Fo•as Mensual	5 208 000	5 218 500	5 208 000
SCJN siglo	Ex edientes/mensual		22 848	22 848
	Fo•as/ mensual		1 632 204	1 632 204
Totales	Ex edientes/mensual	54 560	77 518	77 408
	Fo•as mensual	5 208 000	6 850 704	6 840 204

Para el 2010 se manejarán todos los expedientes y su contenido de la siguiente manera:

- Juzgados de Distrito de 1951 a 2003: de enero a diciembre en un promedio mensual de 77,990 expedientes con un volumen de hojas cercano a las 7.5 millones cada mes.
- Suprema Corte de Justicia de la Nación. Siglos XIX y XXI.: de enero a mayo se manejarán cada mes, 22,864 expedientes con 1,634, 347 hojas promedio.
- Históricos de las Casas de la Cultura: de junio a diciembre se manejarán un promedio mensual de 33 mil expedientes con 1,648,463 hojas.

En suma durante 2010, cada mes en promedio se revisarán poco mas de cien mil expedientes de enero a junio y ciento once mil de julio a diciembre, lo que significará que cada mes se tengan más de nueve millones de hojas útiles.

Requerimientos técnicos de las imágenes y metadatos

Se requiere obtener las imágenes electrónicas de los Expedientes Judiciales con una resolución de 300 DPI, las cuales deberán ser indexadas por expediente y número de O imágenes, identificando para cada uno:

- a) Primera sección: Desde la carátula hasta antes de la sentencia o resolución que ponga fin al procedimiento.
- b) Segunda sección: Sentencia o resolución que ponga fin al procedimiento.
- c) Tercera sección: El resto de la información.

Se deben escanear todas las hojas, inclusive aquellas en blanco, lo que se pretende es conservar en forma íntegra los Expedientes Judiciales.

ARCHIVO MAESTRO-CONSERVACION

IMAGEN	Blanco y negro
RESOLUCION DE ESCANEADO	300 cipi para todos los documentos
TIPO DE ARCHIVO	TIFF no comprimido
ARCHIVO DE CONSULTA	
IMAGEN	Blanco y negro
RESOLUCION	72 dpi para todos los documentos
TIPO DE ARCHIVO	PDF

Se requiere obtener las imágenes electrónicas de los Expedientes Judiciales, los cuales deberán ser indexados por expediente y número de imágenes, identificando para cada uno:

- a) Primera sección: Desde la carátula hasta antes de la sentencia o resolución que ponga fin al procedimiento.
- b) Segunda sección: sentencia o resolución que ponga fin al procedimiento}
- c) Tercera sección: el Resto de la información

Normativa

- MoReq 2 (Model Requirements for the Management of Electronic Record), es un modelo de requisitos funcionales para la gestión de documentos electrónicos de archivo elaborado a través del programa IDA de la Comisión Europea.
- ISO 15489

Planeación y Control de Producción

ee

La administración de la producción se realizará con base en la filosofía MRP, es decir deberá existir inventario suficiente para iniciar la operación, los puntos de enlace o de paro arranque de cada línea dependerá de la cantidad de expedientes existentes para asegurar la continuidad de la operación dentro de las políticas establecidas.

Para perfeccionar los mecanismos de control y disponibilidad del acervo de este Alto Tribunal, es necesario contar con un software para administración y control de archivos que sea integral, seguro, confiable, eficiente y que permita la identificación individualizada y el control de cada expediente mediante elementos tecnológicos de identificación de radiofrecuencia.

Volumen de documentos a procesar, en rincio.	2/100,000 documentos 105 000 000 imágenes
---	--

¿se cuenta con una base de datos para la administración		No
	X	Si

En cuanto tiempo requiere que se realice el servicio		Inmediato
		1 mes
		3 meses
	X	Otro del 2009 al 2013
Soporte de la documentación		Libros con escritura anti ua
		Libros / o revistas a color
	X	Documentos legales, certificados, pagos, etc.
		Documentos técnicos científicos
		Planos
		Otros. Es ecifi ue
Características	X	Documentación blanco ne ro
		Documentación a color
		Texto manuscrito
		Archivos de ima en
Tamaño de los originales	X	Carta
	X	Oficio
		Doble carta
		Otros
Estado de conservación de los documentos	X	Deteriorados
		Faxes térmicos o en mal estado
		Fotoco ias
	X	Documentos originales, claros y en buen estado
Cómo se encuentran almacenados actualmente los documentos		Ho•as sueltas
	X	Car eta de 3 erforaciones
		Cuántos documentos tiene cada ex ediente o car eta
		Los documentos tienen información por anverso reverso
		En fóliders sueltos
	X	Libros em astados
	X	Ex edientes documentos ordenados

La documentación puede traer	X	Gras
	X	Cli s
Retoque de la imagen	X	Cosidos
	X	Sin reto que
	X	En ra ados
	X	De ar como el ori inal
	X	Atados
Una vez digitalizados los documentos serán		Amastados
		Destruídos
		Eliminación de ecas o untos sueltos
	X	Reintegrado a sus expedientes o car etas ara su uarda
		Eliminación de bordes ne ros
Tipo de entrega		Aún or de liberar
	X	Otro
Resolución de escáner		200 DPI
		300 DPI
	X	400 DPI
		600 DPI
		Otro
Tipo de salida de imagen		Otro
	X	TIF
Indexación: promedio de Imágenes por documento con la misma información a indexar		1 Toneladas de imá enes
		JPG
		2 Imá enes
		BMP
		3 Imá enes
	X	PDF
Indexación: cada documento de imágenes será indexada con		No. de imá enes 50
		1 Cam o nombre del archivo
		2 Cam os
		3 Cam os
Promedio de número de caracteres por campo	X	No. de cam os 35
		10 caracteres or cam o
		20 caracteres or cam o
		Otro
Formato de salida a base de datos	X	40 caracteres or cam o
		Access
		Excel

		Word
	X	Otro: Se requiere elaborar propuesta
Finalidad de la digitalización	X	Preservación
	X	Difusión interna
	X	Publicación en Internet

5. CONCLUSIONES

Llevar a cabo un proyecto de digitalización significa establecer un equilibrio entre las necesidades de los usuarios reales y potenciales, la infraestructura tecnológica utilizada por el proyecto y los recursos financieros y humanos disponibles, Además, las posibilidades

C tecnológicas cámaras digitales, de las estaciones sistemas de operativos, trabajo y dispositivos memoria interna, de captura almacenamiento de imágenes (escáneres calidad de

visualización, capacidad de las redes y velocidad) deben estar claras antes de iniciar el proyecto, así como los medios para presentar los datos de las imágenes a los usuarios.

El análisis de las necesidades técnicas de un proyectos de digitalización se lleva a cabo por lo general en el marco de un proyecto piloto o estudio, Las necesidades técnicas se centran fundamentalmente en el hardware y en el software, ambos factores dinámicos en el desarrollo del futuro previsible.

La adquisición de un equipo también deberá incorporar un contrato de mantenimiento por un mínimo de tres años y, si es posible, por cinco. Por lo general, se considera que después de tres años, el equipo requerirá una actualización y después de cinco, será obsoleto y necesitará ser sustituido.

La selección del software deberá basarse en una seria consideración de soluciones de código abierto (es decir, disponibles públicamente).

Las instituciones que inician un programa de digitalización de documentos, deben establecer unos objetivos del proyecto que deben traducirse en resultados concretos (base de datos accesible desde cualquier punto, local o remoto, de la organización, debidamente autorizado, etc.)

La obtención de estos objetivos y resultados por parte de la institución, dependerá de los recursos y procesos que se dispongan para conseguirlos. Los recursos incluirán: personal, finanzas, espacio, tiempo y capacidades técnicas. Los procesos, tras la elección del Sistema de Gestión Electrónica de Documentos (SGED) escogido, incluirán: digitalización, procesamiento de imágenes, indexación/ control de calidad, preservación y gestión.

En el momento de acometer el proyecto de digitalización de documentos, la institución puede ee optar por realizar la totalidad de los procesos internamente o externalizarlo, es decir contratarlo a una empresa especializada, externa.

Existen ventajas y desventajas respecto a la elección de externalización o la creación de capacidades internas para realizar ia digitalización de documentos.

La externalización sólo será viable si la institución tiene una buena comprensión de los objetivos a corto y a largo plazo, y puede especificar de forma completa sus requisitos.

Ventajas:

- Contención del costo y riesgo limitado: se paga por resultados, con frecuencia se pacta un precio fijo por documento,

- _ Los proveedores normalmente están preparados para absorber grandes volúmenes de documentos en los tiempos que se requieran,
 - Los costos de capacitación del personal y obsolescencia tecnológica, son absorbidos por el proveedor.
 - Se dispone de una amplia gama de servicios contratables: digitalización de imágenes, indexación manual o automática, desarrollo de base de datos, procesamiento de imagen, impresión, almacenamiento y copias de seguridad.

Desventajas:

- _ La institución se aleja de las funciones de digitalización (frecuentemente los servicios se llevan a cabo fuera de las instalaciones) y temporalmente de sus documentos.
- _ Vulnerabilidad, debido a la posible inestabilidad del proveedor.
- _ Inexperiencia del proveedor respecto a la cultura de la institución.
- _ Falta de control en la contratación, producción y calidad de la digitalización y procesos derivados.
- _ Casuística relacionada con el transporte, manipulación y seguridad de los documentos.

En cuanto a la producción propia, la institución deberá disponer del personal, las instalaciones y equipamiento de hardware y software necesarios para producir internamente la digitalización de documentos.

Ventajas:

- _ Los requisitos se definen en forma progresiva, en lugar de hacerlos al inicio.
- _ Se mantiene el control directo sobre la digitalización de documentos y toda la gama de funciones derivadas.
- Se prevé una seguridad consistente y de calidad.
- _ Se asegura una manipulación adecuada y la accesibilidad respecto al fondo documental,

Desventajas:

- Gran inversión y tiempo prolongado en la etapa inicial del proyecto.
- _ No hay precio fijo por documento digitalizado.

- La institución paga los gastos (incluyendo los costos de tiempo de inactividad, capacitación/experiencia del personal y obsolescencia tecnológica), en vez de pagar por los productos.

- Capacidad de producción (personal y equipamiento) limitada.

La decisión de externalizar o producir internamente la digitalización de documentos vendrá condicionada normalmente por el fondo documental disponible (con sus problemas de confidencialidad, manipulación, etc.), el tipo de institución (con agilidad o no para la contratación de personal/adquisición de equipamiento y su capacidad financiera) y los requisitos de tiempo previstos para poner en pleno rendimiento un SGED.

Frecuentemente se acude a fórmulas mixtas que incluyen la externalización de digitalización de documentos históricos, con la producción interna del día a día, a partir del momento en que se decide la implantación del sistema. Esta externalización puede ser también parcial (sólo personal o sólo equipamientos) y contratarse para ser ejecutada en las mismas instalaciones de la institución contratante, donde se encuentra físicamente el fondo documental.

Finalmente se señalarán los beneficios que puede acarrear la digitalización de archivos históricos judiciales, del Poder Judicial de la Federación.

Para el país:

- Conservar la documentación con relevancia documentab jurídica o histórica.
- Hacer más eficiente la consulta de la información.
- Impulsar la transparencia y el acceso a la información.
- Brindar el acceso a la información en forma ágil y eficaz.

Para el Poder Judicial de la Federación:

- Coadyuvar en la función jurisdiccional, al reducir tiempos de respuesta de las solicitudes de información,
- Agilizar la localización de información, mediante búsquedas catalográficas.
- Visualizar documentos en línea, a fin de ahorrar tiempos de consulta.
- Conservar el patrimonio histórico, jurídico y documental, por su valor como referente de la historia jurídica del país.
- Consolidar la conservación y consulta electrónicas de los archivos históricos judiciales.
- Hacer más eficiente la organización, clasificación, conservación y consulta de los archivos históricos judiciales.

- _ Contribuir a la solución de los problemas de espacio físico para el resguardo de expedientes judiciales, al conservar aquéllos con relevancia documental/ jurídica o histórica.
- _ Generar estándares que orienten los trabajos y procesos técnicos en materia de archivística judicial.

Para el público, en general:

- _ Acceder a información acorde a sus requerimientos, en forma sencilla, con base en el que se conserve, clasifique y catalogue.
- _ Acotar tiempos en la consulta de información.
- _ Consultar en forma inmediata, a través de internet, la documentación digitalizada, de conformidad con la normativa sobre transparencia y acceso a la información.

6. REFERENCIAS

SAMBARINO, Julia. "Digitalización: Concepto Preliminares" Escuela Superior de Cómputo del Instituto Politécnico Nacional (IPN), México, 2009. P.3.

Idem

Dbidem p.5

7. BIBLIOGRAFIA

- AVERON D.M. "Electronic Imaging: What It's All About" Avedon associates, 1991. • BACA, M. , "Introduction to metadata: pathways to digital information". Version 2.0. Malibu, CA, Getty Standards Program, 2001. <http://www.getty.edu/research/institute/standards/intrometada/index.html>
- BESSER, T. & TRANT, J. "Introduction to imaging: issues in constructing an image database. Malibu", CA, Getty Information Institute 1995. <http://www.getty.edu/research/institute/standards/introimages/>
- BROWN D.S. "Image Capture Beyond 24-Bit RGB". RLG DigiNews. Vol. 3 Num. 5, 15 Octubre 1999. <http://www.rlq.org/preserv/diqinews/diainews3-5.html>
- CHAPMAN, S. "Guidelines for image capture in NATIONAL PRESERVATION OFFICE/RESEARCH LIBRARIES GROUP Guidelines for digital imaging: papers given at the joint" NPO/RLG Preservation Conference, 1998. London 1998. <http://www.rlq.org/preserv/joint/chapman.html>
- D'ALLEYRAND M.R. "Image Storage and Retrieval Systems: A new approach to records management". McGraw-Hill Book, 1989.
- ESTER, M: "Digital image collections: issues and practice. Washington, DC, The Commission on Preservation and Access". (Publication 67), 1996. <http://www.clir.org/pubs/abstract/pub67.html>
- FREY, F. & REILLY, J. "Digital imaging for photographic collections: foundations for technical standards". Rochester; image Permanence Institute, 1999. http://www.rit.edu/~661www1/sub_pages/frameset2.html
- V • FULTON W. "A few scanning tips". 2004. <http://www.scantips.com/>
- GLOVER G. "Image Scanning for Desktop Publishers", Windscrest, 1990.
- GREEN W.B. "Instruction to Electronic Document Management Systems", Academic Press, 1993
- GREESTEIN D., George G, "Benchmark for digital reproductions of monographs and serials". Digital Library Federation. 15 Enero 2002. <http://diglib.org/standards/bmarkfin.htm#what>
- HOWAR B., Trant J. "Introduction to Imaging: Issues in Constructing an Image Database. Getty Art History Information Program, Santa Monica, CA, 1995. <http://www.getty.edu/research/institute/stands/introimages/>
- INSTITUTE OF MUSEUM & LIBRARY SERVICES " A framework of guidance for building good digital collections", November 6, 2000, 2001. <http://www.ims.gov/pubs/forumframework.htm> (Formally endorsed by Digital Library Federation, 1 de Marzo de 2002. <http://www.diglib.org/standards/imsframe.htm>) • JORDAH G. "Color everywhere, but not a byte to spare", Imaging Magazine, Vol. 6, Num. 5, Mayo 1997.
- KENNEY, A.R. & CHAPMAN, S. "Tutorial: digital resolution requirements for replacing text-based material: methods for benchmarking image quality". Washington, DC, Council on Library & Information Resources. (Publication 53), 1996.
- KENNEY, A:R: "Digital benchmarking for conversion and access in KENNEY, A:R. & RIEGER, O. Moving theory into practice: digital imaging for libraries and Archives". Mountain View, VA, Research Libraries Group (RLG), 2000.

- KOULOPOULOS T.M., Frappaolo C. "Electronic Document Management Systems", Computing McGraw-Hill, 1995.
- LAGOZE, C. & PAYETTE, S. "Metadata: principles, practice and challenges in KENNEY, A: R. & RIEGER, O. Moving theory into practice: digital imaging for libraries and archives". Mountain View, VA, Research Libraries Group (RLG), 2000.
- LAWRENCE G.W. et al. "Risk Management of Digital Information: A File Format Investigation", Council on Library and Information Resources. Junio 2000. <http://www.clir.org/pubs/abstract/pub93abst.html>
- LIBRARY OF CONGRESS. Preservation Directorate "Digitizing library collections for preservation and archiving: a handbook for curators". Washington, DC. , 1997.
- LIIKANEN E. "Capture, Indexing & Auto-Categorization" Intelligent methods for acquisition and retrieval of information stored in digital archives. ALLM Industry White Paper. ALLM International Europe, 2002.
- I-UNT P. "High-Speed Scanners Put Documents in the Fast Lane". Imaging & Document Solutions. Vol. 7, Num. 11, Noviembre 1998.
- LUNT P. "New Directions in Data Capture". Imaging Document Solutions, Vol. 9, Num. 3, Mar-zo 2000.
- LUPOVICI, C. & MASANÈS, J. "Metadata for the long term preservation of electronic publications". The Hague, Koninklijke Bibliotheek (NEDLIB report series 2), 2000. • OSTROW, S. "Digitizing historical pictorial collections for the Internet". Washington, DC, Council on Library and Information Resources. (Publication 71), 1998. <http://www.clir.org/pubs/reports/pub71.html> e OTTEN K. "integrated Document & Image Management". Association for Information & Image Management, 1987.
- SITTS, M.K. "Handbook for digital projects: a management too/ for preservation and access". Andover, MA, Northeast Document Conservation Center, 2000. <http://www.nedcc.org/dighand.htm>
- VANSTON L.K. et al. " Computer-Based Imaging and Telecommunications" Technology Futures, 1991.
- WIGGINS B. "Document Imaging: A Management Guide, Meckler Publishing, 1994. • WITTEN, I.H. et al. "Greenstone: a comprehensive open source digital library software system", 2001 <http://www.nzdl.org/>

Bibliografía en línea

- Ocr con redes neuronales:
http://www.cc.uah.es/hilera/docs/1993/c_jiacsel/c_jiacsel.html
- Icr (reconocimiento de manuscrito): <http://www.eslova.com/firmasindex.html>
- Modelos ocultos de markov: [http:// en.wikipedia.org/wiki/Hidden Markov model](http://en.wikipedia.org/wiki/Hidden_Markov_model)
 - Algoritmos de ocr:
<http://201.218.41.233:8080/dspace/bitstream/123456789/18/2/ALGORITMOS.pdf>
 - Procesamiento de imágenes (Libro en español):
<http://books.google.com/books?id=h4Gj8GuwPVkC&pg=PA43&lpa=PA43&dq=binarizaci%C3%B3n&source=bl&ots=Cp-w3cXD6y&sig=gW-qv4krl-2xBIN4sIP448->

[540E&hl=en&ei=4tOySsi3N4fOsQ024YiRDA&sa=X&oi=book
result&ct=result&resnum=6•von e+page&q=binarizaci%C3%B3n&f=false](https://www.google.com/search?q=540E&hl=en&ei=4tOySsi3N4fOsQ024YiRDA&sa=X&oi=book_result&ct=result&resnum=6•von+e+page&q=binarizaci%C3%B3n&f=false)

8. AGRADECIMIENTOS

A mi esposo, Dr. Gabriel Ricardo Manuell Lee, por su apoyo, amor y respeto de más de 20 años.

A mis hijas, Ofelia y Gabriela y a mis siete hermanos, además de cuñados, cuñadas, concuños y concuñas, así como sobrinos y sobrinas y demás familiares, por su cariño y oraciones.

A mis padres y a mi suegro, a los que les hubiera dado una gran alegría. A mi suegra que siempre me ha animado a salir adelante.

Al Ingeniero Javier Jiménez Espriú, por su apoyo y guía.

Al Ingeniero José Luis Antón Macín y a Maricela, su esposa, ya que con sus consejos e impulso permitieron que concretara este proyecto.

A mi querido amigo Doctor Salvador Landeros Ayala, con el que varios sueños los hemos podido convertir en realidad.

Al Maestro Ingeniero Raúl González Apaolasa, por su disposición, consejo, apoyo y compañero de muchas lides.

Al Doctor Diódoro Guerra Rodríguez, del que he aprendido lo importante que es sumar.

Al Doctor Octavio Rascón Chávez y al Ingeniero José Antonio Ceballos, quienes me impulsaron para lograr la aceptación de la Academia.

Al Doctor Jorge Sosa, al Doctor Hugo César Coyote Estrada y al Doctor Rolando Menchaca, por su apoyo en la Comisión de Comunicaciones y Electrónica.

Al Ingeniero Sigberto Alarcón, por su ayuda en todos los trabajos importantes que he desarrollado en mi vida profesional.

A la Maestra Julia Sambarino, quien me introdujo en este tema.

Al Doctor Adolfo Guzmán Arenas por sus recomendaciones para abordar esta temática.

A la Lic. Eva María Belaunzarán, a Adriana Salas, a Brenda Morales, a Armando Sánchez Lugo, y demás colaboradores del despacho Manuell Barrera y Asociados.

OA otros mis proyectos.amigos, compañeros y colaboradores que me han acompañado para cristalizar éste y

SEMBLANZA DE LA DRA. MONICA MA. DEL ROSARIO BARRERA R.

Ingeniera en Electrónica y de Comunicaciones, Universidad Iberoamericana (1976-1980). Maestrías en: Administración (1988-1990) y en Administración Pública (1990-1992), del Instituto Tecnológico y de Estudios Superiores de Monterrey (ITESM). Doctora en Administración Pública de la UNAM (1994-1998). Corresponsable en Instalaciones.

Secretaria Ejecutiva de Obras, Recursos Materiales y Servicios Generales del Consejo de la Judicatura Federal. Fue Secretaria Ejecutiva de Servicios en la Suprema Corte de Justicia de la Nación (2009); Coordinadora General de Planeación y Centros SCT y Consejera Suplente en el Consejo de Administración de PEMEX de 2003-2006. Titular del órgano Interno de Control en la Secretaría de Comunicaciones y Transportes 2001-2003; Contralor Interno: en Pemex-Refinación 1997-1999; de la Secretaría de Energía 1995; de la antes Secretaría de la Contraloría General de la Federación de 1988 a 1994; Directora General de Gas de la Secretaría de Energía 1995-1996; o entre de la Auditoría otros cargos, Superior en los de sectores la Federación, público, en privado 2001 y y de social. la Contaduría Finalista para Mayor ocupar de Hacienda la titularidadde la

Asamblea Legislativa del D.F., 2007. Registro PF020, como Testigo social.

C

Presidenta de 1994 a 1998 de la Asociación de Ingenieros y Arquitectos de México. Presidenta de la Asociación de Ingenieros Universitarios, Mecánicos Electricistas (1990-1992). Es Vicepresidenta del Instituto Mexicano de Auditoría Técnica. Es miembro del Colegio de Ingenieros Mecánicos Electricistas, del Colegio de Ingenieros en Comunicaciones Eléctricas y Electrónica y O del Colegio de Contadores Públicos. Es presidente fundador del Instituto Mexicano para la Integridad Social, A.C. Testimonio a la Excelencia Profesional 1992 de la Asociación Mexicana de O Ingenieros Mecánicos Electricistas. En 1996 fue Excelencia Profesional por el ITESM.

O Ha sido profesora en diversas instituciones y publicado artículos en revistas en materias de administración, auditoría pública e ingeniería.