

Pautas para la Utilización de la Digitalización

Guías
AGN



Comité Asesor de Reprografía y Automatización del Archivo General de la Nación

Coordinador Ingeniero Carlos Rojas Núñez

COLOMBIA



ARCHIVO
GENERAL
DE LA NACIÓN

Contenido

COLOMBIA



ARCHIVO
GENERAL
DE LA NACIÓN

SERIE: Guías y Manuales

Pautas para la Utilización
de la Digitalización

Comité Asesor de Reprografía y Automatización del Archivo General de la Nación

Table with 2 columns: Page number and Title. The page numbers are listed on the right side of the page. The titles are listed on the left side of the page. The table is partially obscured by a watermark and a box at the bottom.

Coordinador
Ing. Carlos Rojas Núñez

BIBLIOGRAFIA

Las ediciones expresadas o utilizadas en la obra son:

Indicador de Serie

1985. Política reprográfica establecida por el Archivo General de la Nación.

Colombia. Archivo General de la Nación. 1985.

Las publicaciones del Archivo General de la Nación en Colombia.

Impreso en Colombia.

ISBN 978-95-9242-114-8

Archivo General de la Nación de Colombia.

1985. Política reprográfica establecida por el Archivo General de la Nación.

Indicador de Serie

1985. Política reprográfica establecida por el Archivo General de la Nación.

Colombia. Archivo General de la Nación. 1985.

Las publicaciones del Archivo General de la Nación en Colombia.

Impreso en Colombia.

ISBN 978-95-9242-114-8

Archivo General de la Nación de Colombia.

ARCHIVO GENERAL DE LA NACIÓN

República de Colombia

Establecimiento público adscrito al Ministerio de Cultura

Consejo Directivo

Ministerio de Cultura
Ministra: Paula Marcela Moreno Zapata
Viceministra: María Claudia López Sorzano
Presidenta del Consejo

Presidencia de la República

Victor Manuel Díaz
Representante del Sr. Presidente

Representante Archivos del País

José Luis Vargas
Archivo Histórico de Antioquia

Academia Colombiana de Historia

Luis Horacio López
Delegado del Presidente

Colciencias

Paola Natalia Nieto
Representante del Sr. Director

Archivo General de la Nación

Sara González Hernández
Director General (E)

Comité Editorial

Sara González Hernández
Antonio Díaz Montiel
Mauricio Tovar González
William Martínez Jiménez
Dania Asprilla Yurgaqui

Comité de Reprografía

Archivo de Bogotá
Leonardo Guzmán
Alberto Cáceres Pineda
Archivo General de la Nación
Carlos Rojas

Ejército Nacional
Germán Villate

Biblioteca Luis Ángel Arango
Francisco Mondragón

Biblioteca Nacional
Sandra M. Angulo

Ecofilm

Jairo Rodríguez
Independiente

Jorge Montoya
Nelson Sánchez Tole

Inversiones Ajoyeco S.A.
Alberto Palacios
Jeannette Esperanza Medina

Sidein

Antonio Ospina
Icontec

Claudia Carrero
Ameritec S.A.

Jorge Quiroga
Binal

Carlos Barrera

Coordinación Editorial

Grupo de Divulgación y Medios
AGN

Impresión, Diseño y Diagramación

Buenos y Creativos S.A.S.

ISBN 978-958-8242-14-9

Archivo General de la Nación de Colombia

Carrera 6 No. 6-91

Teléfono: 3282888 Fax: 337 2019

E-mail: contacto@archivogeneral.gov.co

Página web: www.archivogeneral.gov.co

Bogotá, D.C., Colombia

2009

Impreso en Colombia

Las publicaciones del Archivo General de la Nación de Colombia están protegidas por lo dispuesto en la Ley 23 de 1982. Podrán reproducirse extractos sin autorización previa, indicando la fuente.

Las opiniones expresadas o utilizadas en la obra son responsabilidad exclusiva de sus autores.

Contenido

Introducción	5
1. Requisitos Archivísticos de un Sistema de Digitalización	9
Cuadro de clasificación.....	10
Principios documentales.....	13
2. Captura.....	15
Captura documental	15
Digitalización documental.....	16
Desempeño del escáner.....	17
Escáner <i>Hardware</i> y <i>Software</i>	17
Especificaciones técnicas de un escáner	18
Características en la captura documental	21
3. La Calidad en los Medios Digitales.....	27
El término Calidad	27
Cómo llegar a la imagen digital	28
El documento	29
La tecnología	30
El control de la calidad	31
La fiabilidad	33
4. Almacenamiento de Imágenes.....	37
¡El archivo tiene que considerar sus imágenes como datos de misión crítica!	37
5. Vocabulario Técnico.....	41
ANEXOS	53
BIBLIOGRAFÍA	57

Introducción

Hoy en día, cada vez son más numerosas las aplicaciones que usan documentos *digitalizados* de todo tipo: artículos de revistas, prensa, imágenes fotográficas, documentos manuscritos o antiguos, libros, etc. Sin duda, la continua caída de los precios del *Hardware* necesario para digitalizar, así como el imparable crecimiento de Internet, han contribuido de forma significativa a que las *imágenes digitales* nos sean hoy tan familiares en cualquier sistema de información electrónica. Los Archivos, las bibliotecas y centros de documentación tienen en esta tecnología un recurso cada vez más barato y relativamente sencillo para ofrecer más y mejores servicios a sus usuarios.

La digitalización de documentos comenzó a realizarse de forma masiva en bancos y compañías de seguros, allí donde la gran cantidad de papel suponía disponer de un gran espacio físico. El elevado precio del suelo frente a la gran capacidad de almacenamiento de los medios digitales llevó a estas empresas a la digitalización masiva de mucha documentación. Sin embargo, no todas las ventajas de esta tecnología residen en el ahorro de espacio físico (ni siquiera se considera hoy en día la más importante), sino en la facilidad y rapidez para recuperar un documento entre miles, el uso de un *facsímil* para preservar el original, la disponibilidad de acceso a documentos -antes reservados sólo a investigadores y personal autorizado- a cualquier usuario y en cualquier sitio, gracias también a Internet, etc.

Conocer las técnicas y conceptos básicos para la digitalización de documentos impresos constituye, por tanto, un requisito imprescindible para todos los profesionales de la información.

Los archivos, como centros de información que son, deben despojarse de la imagen cultural y tradicional, formada a través de años, en los que se les vio como el lugar donde se guardaban una serie de papeles viejos, sin ninguna utilidad, custodiados por funcionarios o empleados que por su vejez o mal comportamiento recibían esta "*distinción*"; los archivos deben constituirse en la **memoria de cada organización**.

La unidad básica de todo archivo es el documento, considerando éste como una unidad conformada por un elemento físico; el soporte, y un elemento inmaterial; la información, y con algunas características adicionales, se puede constituir en prueba irrefutable para garantizar derechos, libertades, autoridad, etc. O como testimonio de actuación para la posteridad.

Desde otro punto de vista, el documento es un objeto físico cuya finalidad es conservar y transmitir información mediante mensajes. La función del documento es la de hacer disponible un mensaje. El documento está, por consiguiente, constituido por un material idóneo a conservar y evidenciar los resultados de una actividad de representación física mediante un mensaje comprendido a los destinatarios.

El documento como entidad objetiva es el soporte material de la información y ésta constituye el contenido de aquél. En otras palabras, la interacción entre documento e información consiste en que la forma documental tiene la función de propiciar el conocimiento mediante la fijación y la conservación de la información. Esta es la razón por la cual se afirma que “documentar es predisponer los medios de conocimiento”. La gran masa de la información circula sobre documentos, pero es claro que el documento es un vehículo respecto de la información.

De esta manera prevalece el valor de la información frente a la del soporte, razón que conlleva los programas de reproducción sistemática de documentos creados en papel hacia otras tecnologías que puedan garantizar la permanencia de la información como el microfilme o que faciliten la consulta o transmisión por medios informáticos, como es el caso de la digitalización.

El estar inmerso en la sociedad de la información implica que el ser humano se acostumbre cada día más a las abstracciones, a representaciones del mundo real: dinero plástico, moneda digital, transferencia electrónica de fondos, digitalización de obras literarias, musicales y artísticas, ediciones electrónicas, obras multimedia, bancos de datos con archivos digitales, centrales de información, *home banking*, comercio electrónico a distancia, etc.

Existen programas de computador especializados en aplicaciones de Gestión de Archivos, que permiten transferir la documentación con soporte en papel a medios informáticos valiéndose de dispositivos de entrada como el scanner o la cámara digital y que ofrecen una ágil consulta por parte de los usuarios ya que estos programas cuentan con bases de datos integradas que se encargan de las búsquedas. Para que esto sea posible y funcione de manera eficiente es necesario que se parta de documentación archivísticamente organizada.

Una institución, a partir de actividades administrativas y técnicas tendientes a la planificación, gestión y organización de la documentación producida y recibida por las entidades, desde su generación hasta su finalización con el objeto de facilitar su utilización y conservación, se puede definir como las actividades tendientes a cumplir con este propósito estas actividades son el núcleo del ciclo vital del documento.

El Archivo General de la Nación (AGN), ha venido promoviendo la implementación de programas de gestión documental en las entidades públicas y privadas colombianas, ya que estas en su conjunto regulan la creación, recepción, distribución, consulta y disposición final de los documentos, por esto se deben tener en cuenta los siguientes requisitos archivísticos mínimos:

Contar con una unidad administrativa de la más alta jerarquía encargada del manejo documental de la institución y dotada de los recursos de talento humano y técnico necesarios para el desarrollo de la función archivística.

Haber desarrollado un Programa de Gestión Documental (PGD) que incluya:

Tablas de Retención Documental (TRD), aprobadas por el AGN para el caso de las entidades públicas, y actualizadas y aplicadas en todos los casos.

Disponer de archivos de gestión y de archivo central o intermedio debidamente organizados con los recursos necesarios e inventariados y no tener fondo documental acumulado no organizado.

Requisitos archivísticos de un sistema de digitalización

Partiendo de la definición registrada en la Ley 594 de 2000, relacionada con la Gestión documental, “*Conjunto de actividades administrativas y técnicas tendientes a la planificación, manejo y organización de la documentación producida y recibida por las entidades, desde su origen hasta su destino final, con el objeto de facilitar su utilización y conservación.*”, se puede determinar que las actividades tendientes a cumplir con este propósito están enmarcadas en el manejo del ciclo vital del documento.

El Archivo General de la Nación (AGN), ha venido promoviendo la implementación de programas de gestión documental en las entidades públicas y privadas colombianas, ya que éstas en su conjunto, regulan la creación, recepción, distribución, consulta y disposición final de los documentos, por esto se deben tener en cuenta los siguientes requisitos archivísticos, como mínimo:

Contar con una unidad administrativa de la más alta jerarquía encargada del manejo documental de la institución y dotada de los recursos de talento humano y técnicos necesarios para el desarrollo de la función archivística.

Haber desarrollado un Programa de Gestión Documental (PGD), que incluye:

Tener Tablas de Retención Documental (TRD), aprobadas por el AGN para el caso de las entidades públicas, y actualizadas y aplicadas en todos los casos.

Disponer de archivos de gestión y de archivo central o intermedio debidamente organizados, con los recursos necesarios e inventariados y no tener fondo documental acumulado sin organizar.

Tener documentados los procesos relacionados con el manejo de la documentación de la entidad.

Haber implementado programas de manejo centralizado de la correspondencia.

La implantación de un sistema de gestión de documento electrónico debe ser el objetivo de todas las instituciones colombianas, pero este objetivo no se puede cumplir sin el lleno de los requisitos archivísticos que se tienen para los documentos en soportes tradicionales; igualmente es indispensable que una vez se inicie la aplicación de un sistema de digitalización, se mantenga el soporte papel.

Cuadro de clasificación

El cuadro de clasificación es producto del análisis orgánico funcional de la entidad, hecho por el área responsable del manejo archivístico, que considere todos los documentos que produce o recibe la institución en cumplimiento de sus funciones e independiente del soporte en que se encuentre, se elabore o se reciba.

En otros ámbitos es llamado también “plan de clasificación” y su principal importancia radica en que normaliza al interior de una entidad cómo y dónde se archivarán los documentos.

El sistema de digitalización debe:

- Soportar en el cuadro de clasificación de la organización y ser compatible con él.
- Garantizar que se utilice un cuadro de clasificación que corresponda a la organización física y en el que los expedientes y documentos se puedan representar dispuestos en una jerarquía con un mínimo de tres niveles.
- Permitir que en el momento de la configuración se definan mecanismos de descripción que permitan la indización de la documentación.

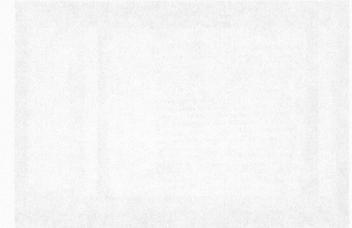
- Permitir a los administradores añadir nuevas series, subseries, expedientes y/o documentos, cuando no existan almacenados en ese punto.
- Permitir la definición y el uso simultáneo de varios cuadros de clasificación. Esto puede ser necesario, por ejemplo, tras la fusión de dos organizaciones; o cuando la entidad ha sufrido una reestructuración.
- Admitir un mecanismo de denominación de series, subseries, expedientes y/o documentos basado en términos incluidos en un vocabulario controlado y en relaciones extraídas de un tesoro que satisfaga la Norma ISO 2788 o la ISO 5964. Asimismo, es conveniente que permita la vinculación del listado de términos normalizado al cuadro de clasificación.
- Facilitar sin límite alguno al número de series, subseries, expedientes y/o documentos que se pueden definir.
- Permitir la opción de creación y mantenimiento automático de una lista o inventario de expedientes y documentos simples.
- Permitir a los administradores añadir (es decir, abrir) tipologías documentales electrónicas en cualquier expediente que no se haya cerrado.
- Grabar la fecha de apertura de cada nueva tipología e incluirla entre los metadatos de éste.
- Impedir que el usuario añada documentos electrónicos de archivo a un expediente cerrado.
- Permitir que el administrador reabra un expediente previamente cerrado de forma temporal y que, tras añadirle documentos de archivo, lo cierre de nuevo. Esta característica es útil con el objeto de hacer la rectificación de errores de los usuarios, pero debe quedar debidamente documentada y hacerse de acuerdo a unas políticas y procedimientos establecidos por el área de manejo archivístico de la entidad.

- Permitir la reubicación de series, subseries, expedientes y/o documentos completos dentro de la jerarquía, o en un lugar distinto del cuadro de clasificación. Asimismo, debe garantizar que todos los documentos electrónicos de archivo ya colocados sigan vinculados a las series, subseries, expedientes y/o documentos reubicados.

La utilización de esta herramienta se ha previsto sólo para casos excepcionales, como fusiones de organizaciones, otras reestructuraciones y corrección de errores.

- Dejar constancia clara en los metadatos y en la pista de auditoría del estado de cualquier serie, subserie, expediente y/o documento de archivo con anterioridad a su reclasificación, de modo que pueda conocerse fácilmente todo su historial.
- Exigir al administrador introducir las razones que le llevaron a la reclasificación de cualquier serie, subserie, expediente y/o documento archivo.
- Impedir en todo momento que se elimine un expediente o cualquier parte de su contenido, salvo en caso de eliminación conforme a la TRD, caso en el cual debe solicitar dos veces la conformidad del administrador para proceder al borrado de los datos.
- Permitir el cierre de un expediente por parte de un usuario o del administrador conforme al procedimiento específico para él.
- Cerrar expedientes de forma automática cuando se cumplan unos criterios determinados definidos en el momento de la configuración y que, cuando menos, incluyan expedientes delimitados por una fecha de corte anual, como el término del año natural, fiscal o cualquier otro ciclo anual definido; transcurso del tiempo desde una acción determinada, por ejemplo, desde la fecha más reciente en que se añadió un documento electrónico de archivo a ese expediente; grabar la fecha de cierre de un expediente e incluirla entre los metadatos del mismo.
- Impedir que un expediente abierto de forma temporal siga abierto después de que el administrador se desconecte del sistema, no obstante debe realizar las alertas previas a esta acción.

- Permitir a los usuarios crear referencias cruzadas entre los expedientes relacionados entre sí.
- Permitir crear entradas múltiples para un documento electrónico de archivo en varios expedientes electrónicos, sin duplicación física del documento electrónico de archivo.
- Disponer de herramientas que proporcionen al administrador estadísticas sobre distintos aspectos de la actividad en el cuadro de clasificación, incluido el número de series, sub-series, expedientes y/o documentos electrónicos creados, cerrados o eliminados en un período determinado, consultas, etc..
- Incluir de forma automática entre sus metadatos los atributos heredados de los metadatos del expediente al que pertenece, tales como su nombre y su código de clasificación, siempre que se abra una nueva tipología.



Principios documentales

A partir de la realización de las funciones, se crean estructuras documentales conocidas como series, las cuales están conformadas por documentos que se relacionan entre sí y tienen características similares en cuanto a estructura y contenido.

El documento tiene origen y se utiliza a partir de la realización de las funciones del negocio, con el objeto de apoyar los procesos de la entidad dentro de su entorno; éste, a través de su estructura y contenido debe satisfacer las necesidades por las cuales fue creado y a su vez, reflejar correctamente lo que se comunicó o se decidió.

Para que el documento cumpla con el objetivo por el cual fue creado, la entidad debe tener políticas y procesos claros en cuanto al manejo del ciclo vital, de tal forma que no se alteren los principios de autenticidad, confiabilidad, integridad y disposición.

El principio de autenticidad se encuentra presente cuando se puede probar que el documento es lo que pretende ser, es decir, cuando se produce un documento, éste puede pro-

bar que suplió una necesidad, cumplió con objetivo, está involucrado en una función y su estructura es acorde con lo establecido por la entidad y la legislación; también cuando se puede probar que el documento ha sido producido por la persona de la cual se afirma que lo produjo y que ha sido creado en el momento en el que se afirma que lo fue.

El principio de confiabilidad se encuentra presente cuando el documento es una representación completa y precisa de las actividades que conforman una función administrativa de la que da testimonio y garantiza el transcurso de las actividades posteriores.

El principio de integridad de un documento se refiere a que éste se encuentre completo e inalterado, es decir, un documento se considera íntegro cuando no contiene enmendaduras, supresiones, anotaciones o adiciones, en el caso de que se presenten, la entidad debe establecer políticas claras para su realización en cuanto a los responsables, circunstancias, indicaciones, entre otros.

El principio de disponibilidad de un documento se refiere que éste se pueda localizar, recuperar, presentar e interpretar; es decir, adicional a que el documento pueda ser prestado, se deben tener los mecanismos necesarios para identificar los documentos desde su momento de producción.

Los anteriores cuatro principios son características fundamentales inalterables que se deben contemplar desde la producción hasta la disposición final del documento independientemente de sus características internas o externas, el soporte, los formatos, entre otros.

Cuando se implementen procesos de digitalización se hace necesario que la entidad defina claramente cómo se pueden establecer los principios del documento en las imágenes digitales, de tal forma que la validez del documento tanto en papel como digital sea la misma.

Captura

Captura Documental

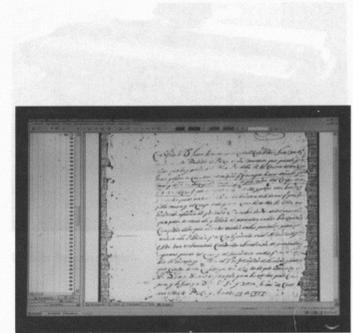
La captura documental es un proceso que involucra numerosos elementos. Los documentos deben inicialmente pasar por el proceso de preparación documental dadas las características físicas del documento las cuales son de importancia significativa en un proceso de captura.

La creación de la imagen no sólo involucra la digitalización sino también el mejoramiento de la misma, optimizando el uso de los equipos y minimizando las re-digitalizaciones de los documentos fuente, obteniendo imágenes aceptables.

Los documentos digitalizados requieren índices para la recuperación, los cuales pueden ser creados de forma manual o automática. El control de calidad se debe aplicar sobre las imágenes y sobre dichos índices para asegurar legibilidad y recuperabilidad. Finalmente una vez las imágenes y sus índices son aceptables, éstas deben ser liberadas del proceso de captura y pasar al proceso de almacenamiento.

Antes de obtener la imagen de un documento, se debe determinar las características físicas del papel; se debe conocer el tamaño del papel, el grosor, y la calidad entre otros, para evitar desastres y costos por reprocesos.

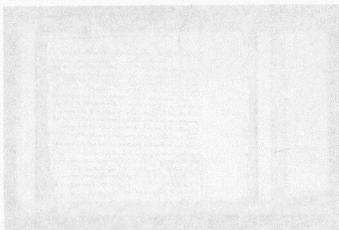
Una vez ha sido obtenida la información anteriormente descrita, se comparan las características del documento original con las especificaciones técnicas del escáner (si se cuenta con un equipo para ejecutar la digitalización) o si por el contrario es necesario obtener un nuevo equipo.



Consulta de documentación via web

Digitalización Documental

Digitalizar es el proceso de convertir una imagen en papel a una imagen que puede ser reconocida en su totalidad por un computador.



Escáner

El Escáner es el equipo o el *Hardware* usado para transformar la imagen que se encuentra en papel en una imagen digital y ser enviada a un computador. Los escáneres se presentan en una variedad de tamaños y formas, el objetivo es poder entender los tipos de escáner (incluyendo conexiones físicas) así como las interfaces de *Software*, *Drivers* y la funcionalidad que provee cada interfaz.

El proceso de digitalización se realiza en las siguientes etapas:

1. El documento es cargado en el dispositivo, sea de forma manual o automática.
2. La luz de la lámpara del escáner incide en el documento reflejando diferentes intensidades según sea gris, negro o color.
3. Las diferentes intensidades se convierten en señales eléctricas a través de un sensor. Actualmente estos dispositivos se denominan CCD.
4. Las señales eléctricas son codificadas en bits.
5. La imagen es comprimida y temporalmente almacenada.
6. La imagen es desplegada a través del *Software* en una pantalla de computador para poder hacer control de calidad.
7. Los índices de la imagen son creados de forma manual o automática.
8. La imagen y sus respectivos índices son enviados al almacenamiento permanente.

Desempeño del Escáner

Seleccionar el correcto escáner es una decisión clave y hay varias consideraciones que se deben tener en cuenta para ello. Entre otras, está el volumen diario, la velocidad, la confiabilidad del equipo, el tiempo total que puede tomar el equipo en completar la labor y la relación costo-mano de obra involucrada.

Es esencial seleccionar el escáner correcto para el trabajo y predecir el desempeño en tiempo real.

Escáner Hardware y Software

Existen tres grandes categorías para la clasificación de escáneres, la primera son los de escritorio que se refieren a los equipos de oficina con bajas capacidades de producción, la segunda son los de producción los cuales son usados para la digitalización masiva de archivos a muy alta velocidad y la tercera son escáneres para aplicaciones gráficas los cuales son equipos de baja velocidad con cama plana y resoluciones muy altas. Una aplicación puede llegar a requerir más de un tipo de escáner y es muy importante dimensionar el equipo adecuado.

Los tipos de escáner pueden ser: Con alimentador automático ó ADF, con cama plana o una mezcla de los dos anteriores.

Los escáneres dependiendo de su aplicación pueden digitalizar papel, rollo de microfilmación, microfichas, planos, fotos, radiografías, etc., y con las aplicaciones de *Software* que se ofrecen actualmente es posible hacer reconocimiento automático de caracteres (OCR), eliminación de los colores de fondo de un formulario, reconocimiento de códigos de barras e impresión para control en los documentos, entre otras.

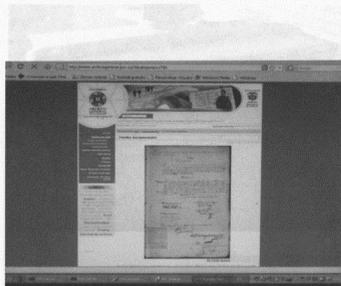
Interfaces

La interfaz es la forma a través de la cual se conecta un escáner a un computador. Hay dos tipos de interfaces a saber:



Escáner formato grande





Digitalización de fotografías

Interfaz de Hardware: Entre la cual se encuentran los puertos SCSI, USB, Paralelo y Serial.

Interfaz de Software: Se encuentran TWAIN (*Technology Without An Important Name*), que fue lanzada en 1992 para el proceso de traer una fotografía o cualquier otra imagen a un documento; en principio se diseñó para aplicaciones de bajo volumen, pero actualmente soporta la digitalización en producción. ISIS (*Image and Scanner Interface Specification*) Originalmente desarrollada por *Píxel Translations* para trabajo en alto volumen con opciones de rotación de imagen.

Actualmente las dos interfaces de *Software* o *Drivers* están diseñadas para soportar captura de imágenes bitonales, escala de grises y color.

Especificaciones Técnicas de un escáner

1. **Velocidad:** Se mide en PPMs (páginas por minuto). Es un factor que depende del tamaño del papel y de su orientación (horizontal o vertical).
2. **Ciclo útil:** Se mide en páginas diarias. Representa el total de número de páginas que un escáner puede producir. Adicionalmente es una medida de garantía de los equipos. Esta medida toma en consideración tiempos muertos por mantenimiento, configuraciones, etc.
3. **Resolución:** Se mide en DPIs (*Dot Per Inch* o puntos por pulgada). A mayor resolución mejor calidad de imagen y menor velocidad.
4. **Tipo de imagen:** Puede ser bitonal, escala de grises o color.
5. **Tamaño de la página:** Ofrece las medidas de los tamaños máximos y mínimos de papel que puede pasar por el escáner. Puede estar dado en centímetros o en pulgadas.
6. **Rendimiento promedio:** Permite determinar el rendimiento real del desempeño del escáner considerando problemas como imágenes pobres, con ángulos de inclinación, saltos, etc.

7. **Rendimiento real:** Corresponde a un porcentaje de la velocidad teórica del escáner. El valor de este porcentaje depende de la resolución de captura de la imagen y la cantidad de información que se deba procesar, ya que a mayor resolución, mayor cantidad de información, tamaños de archivos digitales mayores y por lo tanto menor velocidad.
8. **Formato de archivo:** Tipo de archivo en el cual quedará la imagen en el computador.
9. **Alimentador:** Especifica cómo es el sistema de alimentación del equipo. ADF ó cama plana.

Selección del Escáner

Los siguientes puntos deben ser tenidos en cuenta a la hora de seleccionar un escáner:

1. **Volumen diario:** Es el volumen actual de documentos que se deben digitalizar en un ciclo de trabajo, esta cantidad puede estar dada en volumen semanal, mensual o anual.
2. **Condición de los documentos:** Se debe considerar si los documentos están impresos por ambos lados, contienen grapas, están en carpetas, son formularios, tienen color, condición del papel y cómo se van ha indexar.
3. **Digitalización de archivo inactivo:** Se deben considerar la cantidad de documentos para convertir pertenecientes al archivo inactivo y sumarlo al volumen diario para poder determinar cuál debe ser el desempeño del escáner.
4. **Mano de obra:** Sin la mano de obra que mantenga el o los equipos operativos trabajando a su óptimo desempeño, el rendimiento promedio se verá afectado.

A continuación se mencionan algunas características que pueden impactar la labor requerida o el equipo necesario.

- Búsqueda y transporte de documentos.



Cámara planetaria digital



- Remoción de grapas.
- Tomar fuera los documentos de sus carpetas.
- Documentos arrugados.
- Documentos dañados por gotas de sustancias orgánicas o inorgánicas. Por ejemplo el café.
- Clasificación de documentos por tipo o por lotes.
- Adición o remoción de encabezados de lote.
- Seguimiento de los documentos escaneados.
- Regreso de los documentos al archivo.
- Documentos perdidos.
- Alimentación o múltiple o ángulos de inclinación en el proceso de digitalización de los documentos.
- Daño del equipo del escáner.
- Excesiva alimentación manual.
- Excesiva tasa de re-escaneo.
- Revisión de resultados de OCR (*Optical Character Recognition*).

Características en la captura documental

Resolución

En tecnología de imágenes la resolución se expresa en términos de DPI (*Dots per inch* o puntos por pulgada). La resolución actual está típicamente en el rango de los 100 DPI a los 1600 DPI, siendo el rango más común de digitalización de 200 DPI a 400 DPI. Es importante recordar que entre más alto sea el número de puntos por pulgada ó DPIs, mayor será la capacidad de almacenamiento requerida.

Puntos por pulgada cuadrada

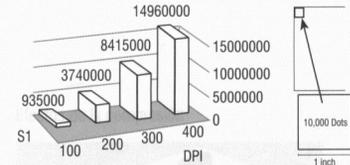
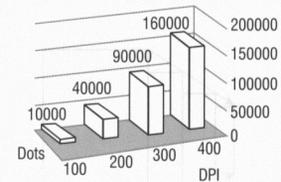
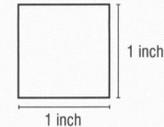
En imágenes bitonales, los puntos por pulgada cuadrada son iguales a los bits por pulgada cuadrada, pero en escala de grises y color las imágenes requieren una mayor cantidad de bits por punto o píxel y por lo tanto mayor memoria.

Píxeles en 8 ½" x 11"

En una resolución de 100 DPI, una pieza estándar de papel puede tener al menos un millón de puntos.

Algunas resoluciones típicas son:

Resolución	Resolución (DPI)
Fax – Estándar	100 DPI (98 x 204)
Fax – Fino	200 DPI (196 x 204)
Fax – Super fino	400 DPI (392 x 204)
Scan	200 – 400 DPI
Print	150 – 600 DPI



Selección de la resolución apropiada

Antes de seleccionar la resolución apropiada se debe considerar la aplicación en la cual se van a utilizar las imágenes, así por ejemplo, si la mayoría de las imágenes se van a utilizar en una estación de trabajo sin OCR (Reconocimiento Óptico de Caracteres), visibilidad del documento limitada y poca necesidad de impresión, es viable trabajar con una resolución de 200 DPI; por otra parte, si la mayoría de las imágenes van a ser utilizadas en procesos de OCR y de impresión, la resolución mínima recomendada es de 300 DPI.

Muchos de los motores de reconocimiento óptico de caracteres han sido calibrados para imágenes que se encuentren digitalizadas mínimos a 300 DPI.

La mejor aproximación para este tipo de procesos, es digitalizar el documento en diferentes resoluciones, de tal forma que imágenes con baja resolución serán de fácil acceso y consulta, mientras que imágenes con alta resolución serán mejores para impresión y procesos de OCR.

Resolución de Digitalización

El escáner determina la resolución de digitalización, los parámetros de calidad y el detalle de la imagen. Una vez que la resolución se ha determinado, es necesario calcular el número de bytes de información requeridos por pulgada cuadrada en el documento. Es importante recordar que la resolución es el número de puntos por pulgada tanto horizontal como vertical.

Número de puntos (píxeles o bits) por pulgada cuadrada

Dpi Horizontal	Dpi Vertical	Total de puntos en pulgada cuadrada
200	200	40000
300	300	90000
400	400	160000

Número de bytes por pulgada cuadrada

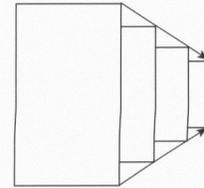
Dpi	Total de puntos en pulgada cuadrada	Bytes por pulgada cuadrada
200	40000	5000
300	90000	11250
400	160000	20000

Compresión

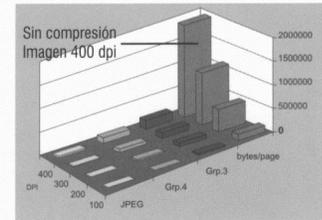
Después de ver la cantidad de memoria requerida para almacenar una página digitalizada, es fácil encontrar que poder tener suficiente memoria para el almacenamiento puede ser un inconveniente. La solución a este problema es la compresión.

La compresión es un algoritmo matemático o un esquema que remueve o combina bits innecesarios reduciendo así el tamaño de la imagen para almacenar. El tipo más común de compresión está basado en el estándar conocido como CCITT Grupo 3 (10:1) o Grupo 4 (20:1) estos algoritmos de compresión son considerados “sin pérdida de información” de tal manera que el mapa de bits original puede ser completamente reconstruido.

Otra forma de compresión típicamente utilizada en imágenes a color es denominada JPEG considerada “con pérdida de información” y puede entregar relaciones de compresión de hasta 100:1. Jpeg es el estándar de compresión de imágenes en color y en escala de grises.



Efectos de la compresión

**Resumen de relaciones de compresión.**

Estándar	Relación de compresión promedio
CCITT Grupo 3	10:1
CCITT Grupo 4	20:1
JPEG	100:1
DJVU	1500:1

Factores que intervienen en el grado de compresión

El grado de compresión de una imagen puede verse afectado por cualquiera de los siguientes factores:

- **Calidad del documento original:** Cuántas veces ha sido fotocopiado, si hay arrugas u otros componentes en el papel o si hay puntos en el papel.
- **Líneas Verticales:** Las formas utilizan líneas verticales las cuales no se comprimen bien. La compresión básica está basada en remover el espacio entre líneas de texto y las líneas verticales no tienen.
- **Grado de inclinación de la imagen.**
- **Contraste de la imagen.**



Es posible experimentar un fenómeno denominado compresión negativa. En este caso, el tamaño del documento después de su compresión resulta ser mayor al tamaño del documento antes de la compresión. Esto es cuando el documento es muy complejo para ser manejado por el algoritmo CCITT. Algunos sistemas son capaces de detectar este fenómeno y enviar una alerta para que el usuario decida si continúa con el proceso de compresión o decide trabajar con la imagen sin comprimir.

Cálculo del tamaño de la imagen

¿Cuál es el tamaño en KB de una imagen tomada de un documento de 8 1/2" x 11" digitalizado a 200 DPI?

Antes de proceder, es importante mencionar que el término usado más frecuentemente para referirse a la resolución de un escáner de imágenes es DPI (*Dots per Inch*), este término ha sido heredado de las tecnologías de impresión y estrictamente hablando hace referencia a la resolución de impresión, sin embargo suele usarse DPI para referirse a SPI (*Samples per Inch* o muestras por pulgada) que es el término más apropiado para referirse a resolución de digitalización.

Habiendo hecho esta aclaración, aquí está la explicación:

Bytes = vSPI X hSPI X área X Codificación de profundidad de color / 8

vSPI: SPI vertical

hSPI: SPI horizontal

Área: Es el área en pulgadas cuadradas del documento escaneado

Codificación de profundidad de color: Número de bits empleados para codificar la información de color de un documento. 1 bit para blanco y negro, 8 bits para 256 colores, 16 bits para 65.536 colores, 24 bits para 16 millones de colores.

Kilobyte = 1024 Bytes

Byte = 8 Bits

Para responder la pregunta aplicamos la fórmula:

Bytes = 200 X 200 X (11 X 8,5) X 1 / 8 = 467.500 Bytes

Kilobytes = 467.500 / 1024 = 456.543 Kilobytes

Por lo tanto una hoja de tamaño 11 pulgadas por 8,5 pulgadas, que ha sido escaneada a una resolución de 200 SPI en blanco y negro, ocupará un espacio de 456,5 Kb en memoria. Este tamaño se puede reducir utilizando los mecanismos de compresión mencionados anteriormente, y debe adicionarse con el espacio que ocupa la información de encabezado (Nombre del archivo, propiedades de la imagen, tipo de imagen, fecha de creación, etc.) dependiente del formato en que será almacenada.

Hay tres consideraciones para determinar los requerimientos de almacenamiento de una imagen.

- Resolución de digitalización.
- Tamaño del papel del documento original.
- Compresión.

La calidad en los medios digitales

El término calidad, no tiene para todas las personas el mismo significado y valor frente a un determinado producto (producto o servicio), y más cuando se habla de documentos (documentos en papel, microfilmes e imágenes digitales), convirtiéndose el término en algo subjetivo. Es por ello, que en este capítulo de esta guía, trataremos de relacionar los principales aspectos que se deben tener presentes en la determinación de la calidad.

Dentro de un orden mostraremos de forma sencilla, lo que para los expertos y tratadistas significa el término calidad, esto con el fin de que todos podamos hablar un mismo lenguaje. Luego, miraremos cuáles son las formas para llegar a la imagen digital, confrontándolos frente al significado de la calidad.

El término Calidad

A. V. Feigenbaum

Nos dice que dentro de la frase Control de Calidad, “la palabra calidad no tiene el significado de lo mejor en sentido absoluto – industrialmente quiere decir mejor dentro de ciertas condiciones del consumidor. Dentro de esas condiciones son importantes, el uso a que el producto se destina y su precio de venta”.

Douglas C. Montgomery

“Es esencial que los productos satisfagan los requisitos de los usuarios. Por lo tanto, la definición de calidad es aptitud para el uso”.

Jack Campanela (ASQC)

“La calidad se utiliza en un sentido más amplio y significa adecuación para el uso por parte del usuario. Cuando la calidad se define así, abarca tanto la calidad de diseño como la calidad de conformidad, y los programas de mejora de la calidad tienen que abarcar todas las fases de la vida del producto, desde el diseño hasta su uso por el cliente”.

Genichi Taguchi

“No está de acuerdo con enfocar la calidad como <conformidad con los límites de la especificación>. La diferencia entre un producto que está justo dentro de los límites de la especificación y un producto que está justo por fuera de los límites de la especificación es pequeña, sin embargo uno se considera <bueno> y el otro <malo>”.

“Taguchi, define la calidad como <--- la pérdida ocasionada a la sociedad desde el momento en que se expide el producto>. Esta pérdida incluye, entre otras cosas, la insatisfacción del cliente, más costes de garantía para el productor y las pérdidas debidas a que una compañía tiene mala reputación, lo que conduce a una subsiguiente pérdida de la participación en el mercado.” (Jack Campanela – ASQC)

J. M. Juran

“Entre las necesidades o exigencias del usuario, la extensión en que el producto sirve con éxito a los fines del usuario, durante el uso, se llama aptitud para el uso. Este de la aptitud o idoneidad para el uso, es popularmente llamado por nombres tales como calidad, es un concepto universal, aplicable a todos los bienes y servicios.”

“La idoneidad o aptitud para el uso se determina por aquellas características del producto que el usuario, puede reconocer como beneficiosas para él.”

Cómo llegar a la imagen digital

Los medios para llegar a la imagen digital son el documento, el microfilme y el escaneo directo de los documentos. De conformidad con el significado de calidad, “aptitud o idoneidad para el uso”, o lo que puede ser lo mismo, “cubrir las necesidades de la institución y de sus usuarios”, entraremos a mirar los elementos que entran en juego en la calidad de la imagen digital, no sin antes decir que todo esto requiere una verdadera planeación de la calidad del manejo documental, en donde se observen todos los elementos que en ello intervienen, incluidos los especialistas (no sólo los ingenieros de sistemas) :

El documento

Aunque anteriormente este tema se trató, es importante mirarlo desde otra óptica, en especial con lo que va a estar relacionado con la calidad; por lo tanto, lo primero que se debe considerar es el documento como medio de información y respaldo a la gestión empresarial, y que en definitiva es el protagonista para la aplicación de la tecnología de conservación e información.

El documento, desde un principio, siempre ha estado manejado en forma física, hasta que entró la tecnología de la fotocopiadora, del microfilme y posteriormente el manejo de la imagen digital. Cuando se documentan los procedimientos y se crean los formatos y los requisitos del funcionamiento operativo, nunca se piensa en los elementos de calidad que cada uno de los tipos documentales deben poseer, desde el punto de vista administrativo, informático, de conversión y traspaso a otras tecnologías (aún se obra de la misma manera); para el caso que nos atañe, estos problemas traen como consecuencia, otros, dentro de los que podemos mencionar: información inapropiada, falta de legibilidad, formatos con dimensiones que no son propias para los equipos de conversión, y que tampoco aprovechan las ventajas de la automatización, estos problemas, no los puede remediar la tecnología, salvo, muy parcialmente en la imagen digital (desarrugar un documento, dar un poco de contraste a la escritura), pero ello tiene un costo superior al trabajo base del sistema.

Para que los documentos tengan una respuesta con calidad por parte de la tecnología digital (incluso en el microfilme), se debe emprender una tarea, de diseño de los documentos, de vigilar el contraste de la información, crear una identificación de documentos por medio de códigos de barra (con calidad), y vigilar el trato que se le da al documento (rasaduras, arrugas, etc.), por último, en la medida de lo posible, los archivos empresariales no deben tener documentos propios generados en fotocopia y en papel de fax.

Dentro del diseño de los documentos, se deben tener presente los tipos y tamaños de las letras, los colores (tanto en los caracteres como en el papel, las fotografías, entre otros), ya que en el sistema digital estos factores cuentan en la capacidad de almacenaje de las imágenes.



Fig. 10.10. Trabajo documental. (1)

La tecnología

Para poder manejar y controlar la calidad en los procesos de conversión del documento a una nueva tecnología, es preciso entender que las fotocopiadoras, las microfilmadoras y los escáneres, trabajan bajo principios fotográficos, es decir, las imágenes se forman por medio de la iluminación del documento, en donde por reflexión, la imagen es captada y grabada en el medio correspondiente (papel, película, discos magnéticos); el color del papel y de las tintas tiene un gran efecto, pues estas características se pierden con postratamientos en blanco y negro, y su afectación de legibilidad depende del color; es por ello, que se da énfasis a la calidad del documento.

La microfilmación

Desde hace una década aproximadamente, la microfilmación ha estado a la par con la imagen digital, es así que existen microfilmadoras que realizan las dos operaciones; igualmente, se han creado equipos que se encargan de transformar la imagen micrográfica en imagen digital, bien indirectamente o en forma directa.

Para convertir indirectamente las imágenes micrográficas en imágenes digitales, se requiere de un lector de microfilmes digital (sus denominaciones dependen del fabricante) en donde se ha remplazado la parte de fotocopiado por un escáner y un software; el procedimiento es lento, ya que debe hacerse imagen por imagen, lo cual lo hace más costoso. Las deficiencias de calidad de las imágenes se reproducirán igualmente en el nuevo medio digital.

La forma directa de llevar las imágenes micrográficas a imágenes digitales, es el empleo del escáner, estos equipos trabajan a alta velocidad, por lo tanto, exigen que el microfilme sea de alta calidad técnica, es por ello que hoy las normas o estándares técnicos para microfilmación, han sido ajustados para responder a una buena calidad de las imágenes digitales.

El escaneo

Además de la calidad que deben tener los documentos (características de los formatos, un alto contraste entre el fondo del papel y la escritura, entre otros), se debe tener presente la

calidad de la organización de los documentos, la denominación o nombre (serie, subserie, nombre del contenido), estado de los documentos, este último factor es importante, ya que los escáneres que por lo general se emplean en la digitalización de los archivos, son del tipo rotativo, y pueden crear problemas de imagen y de daño en los documentos.

Otro de los factores que se debe tener presente al momento de digitalizar, es tener los documentos desencuadrados, para poder realizar la operación de escaneo, esta parte requiere de un proceso delicado, pues los documentos no pueden quedar con residuos de cosido y de pegantes, pues puede repercutir en el buen funcionamiento de los equipos. Sobre el aspecto de desencuadrar, se debe tener presente no cortar información.

La recuperación de la información

La recuperación de la información es básica en todos los sistemas (archivos, microfilmes e imágenes digitales), lo más importante es planear bien el contenido de las bases de datos, teniendo presente cuál ha de ser el uso y las aplicaciones que se requieren, teniendo presente no establecer características o parámetros que no se van a utilizar, y que en muchos casos no van a ser detectados por los usuarios del producto.

Se deben considerar dentro de las bases de datos: los archivos del Estado, estos deben seguir los lineamientos del Consejo Internacional de Archivos, con miras a poder difundir y recibir información a nivel mundial bajo parámetros estándar, por lo tanto, las normas ISAD(G) e ISAAR(CPF), son un requisito imprescindible. La empresa privada, por lo general tiene sólo necesidades internas, por lo tanto, y muy seguramente, se acoplarán más a los *Software* comerciales.

El control de la calidad

Como vimos en la definición de la calidad, la mayoría de los especialistas hablan de la aptitud o idoneidad para el uso; esta aptitud, no siempre es la misma en todos los archivos, como tampoco ésta, es conocida para un determinado producto por todas las personas dentro de una misma entidad, en especial, por quienes deben producirlos; de tal manera,



Sala de consulta documental - AGN

que la calidad se debe planear, estableciendo las características principales del producto y determinando los respectivos índices de calidad, en otras palabras, es establecer el estándar, sobre el que se evaluará dicha calidad.

Para el caso de los sistemas digitales, existen normas que fijan las características mínimas de calidad, una de ellas es la ANSI / AIIM MS44 -1998 (Rev. 1993), la cual se debe emplear con las tarjetas de escaneo (similar a la tarjeta de resolución en el microfilme) X 441 y X443.



M. M. - Imagen digital con resolución 300

En la imagen digital está conformada por puntos o caracteres denominados píxeles, y están demarcados por la cantidad de éstos en una pulgada, dando como parámetros: 100, 200, 300, 400 y 600 píxeles, éstos son equivalentes a los patrones que trae la tarjeta de resolución de los microfilmes, así: 2 para 100 píxeles, 3,9 para 200 píxeles, 5,9 para 300 píxeles, 7,9 para 400 píxeles y 11.8 para 600 píxeles. Como se puede observar, sólo en los 300 píxeles, se está en las mismas condiciones de resolución que nos proporciona el microfilme, en los 400 y 600, se estará en un nivel superior a la microfilmación.

Para la valoración o enjuiciamiento de la calidad, Joseph M. Juran nos dice:

“La aceptación del producto implica dos determinaciones diferentes: 1. ¿Es conforme el producto con la especificación?, 2. ¿Es adecuada para el uso?

Cada una de ellas tiene un objetivo distinto. La constatación de la conformidad con la especificación es necesaria para:

- Proteger al usuario cuando no se dispone aún de datos sobre el uso.
- Impartir criterios operativos a quienes desconocen la adecuación para el uso.
- Crear una atmósfera de orden.

- Proteger de inculpaciones injustificadas a los inocentes.

Por el contrario, la constatación de la adecuación para el uso es necesaria para:

- Proteger al usuario.
- Mantener la comerciabilidad del producto.
- Evitar el perfeccionismo”.

“Cuando el producto, en virtud de la inspección, se declarara conforme con la especificación, se supone también que es adecuado para el uso”. “Sólo cuando no es conforme con la especificación surge la segunda pregunta: ¿Es el producto (no conforme) adecuado para el uso?”

Lo dicho anteriormente, plantea la necesidad de que las instituciones, conjuntamente con la creación de las especificaciones, creen los parámetros específicos de valoración, es decir, determinar qué es un error y categorizarlos por grados de deméritos, esto con el fin de no desechar productos, que si bien no cumplen las especificaciones, sí son aptos para el uso.

La fiabilidad

“El diseño de un producto (bienes o servicios) debe reflejar la aptitud para el uso, y el producto debe estar de acuerdo con el diseño. La amplitud con que el producto cumple con el diseño se llama << calidad de conformidad >>,”la cual es la resultante de muchas variables que intervienen en su construcción y desarrollo, los que indudablemente tendrán que ver con su durabilidad y uso.

Para los productos (bienes o servicios) de larga duración, algunos factores relacionados con el tiempo entran en juego, éstos los denominaremos habilidades (disponibilidad, fiabilidad, mantenibilidad), y están íntimamente relacionadas y son vitales en la aptitud para el uso.

La disponibilidad: está relacionada con la continuidad de vida, por tal motivo los expertos en calidad han dedicado esfuerzos para minimizar la proporción de fallos de los productos y cómo restaurar el servicio rápidamente en caso de fallar. “Se dice que un producto está disponible cuando está en estado operativo”.

Fiabilidad: “si los productos no fallaran nunca, la disponibilidad sería del 100%. Sin embargo, los productos fallan”. Por lo tanto, “fiabilidad es la probabilidad de que el producto realice las funciones para las que fue construido de manera satisfactoria, durante un lapso de tiempo determinado dentro de ciertas condiciones ambientales, definidas durante el diseño”.

En los sistemas digitales, las expectativas de vida de los sistemas y de la conservación de las imágenes digitales, no supera los 15 años, debido a la obsolescencia tecnológica, que nos presentan día a día los desarrollos de *Hardware* y *Software*; por tal motivo, se recomienda que en el diseño del sistema se prevea, la posibilidad de transferir las imágenes a otro medio más seguro y duradero, como por ejemplo la microfilmación.

Los sistemas híbridos (digital y microfilmación), no obstante aún sus altos costos, proveen la alternativa de emplear la primera en aquellos documentos que se requieren permanentemente para mantener un sistema de información ágil y rápido para la toma de decisiones, y emplear la microfilmación en los documentos de carácter vital e histórico de las instituciones.

Dado que todas las tecnologías están sujetas a fallas, y que las nuevas tecnologías se están introduciendo a intervalos cada vez más reducidos, es posible quedar demasiado absortos en preocupaciones acerca de la vida útil de los medios de almacenamiento digital. La supervivencia a largo plazo requiere de un plan exhaustivo que incluye atención respecto a la vida de los medios, el entorno de almacenamiento, los procedimientos de manipulación, detección de errores, copias de seguridad (*backup*), respuesta ante desastres y control de la obsolescencia de *Hardware* y formatos.

Mantenibilidad: “la necesidad de continuidad del servicio ha estimulado, también, la realización de muchos esfuerzos para mejorar el mantenimiento de productos de larga vida”. Este mantenimiento se realiza de 2 maneras (en equipos y sistemas):

Mantenimiento preventivo o programado.

Mantenimiento no programado, que consiste en el restablecimiento del servicio en caso de fallo. La efectividad del mantenimiento está fuertemente influida por la tecnología de base, y por la disponibilidad de las piezas de repuesto (apoyo logístico).

Para terminar es indispensable reflexionar sobre “el daño económico derivado de las averías en piezas y en funcionamiento de estos sistemas, tanto por el creciente valor de los aparatos como por las cada vez mayores responsabilidades que se les confía”. La fiabilidad cuesta y, en consecuencia, es necesario valorar cada vez el nivel más económico, teniendo en cuenta las ventajas que de él se pueden derivar.

Bibliografía:

Mauricio Harbin y Gregorio Invrea, EL CONTROL DE CALIDAD, Ediciones Deustos , S.A. España 1979,

Dale H. Besterfield, CONTROL DE CALIDAD, cuarta edición, Prentice Hall Hispanoamericana, S.A. México, 1994.

Joseph M. Juran, MANUAL DE CONTROL DE LA CALIDAD, cuarta edición, Editorial Reverté, S.A. España 1990

Almacenamiento de imágenes

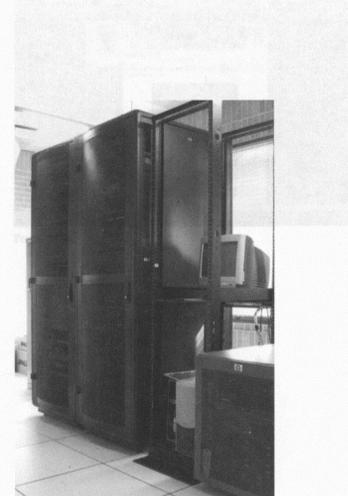
¡El archivo tiene que considerar sus imágenes como datos de misión crítica!

Tal y como se menciona en el entorno comercial, “*Los datos y la información son el activo más valioso del negocio*”. El archivo y su consecuente masa documental digitalizada, es por tanto, un elemento de valor crítico del cual depende en gran medida la disponibilidad y servicio de acceso, es decir, el cumplimiento de una importante parte de sus objetivos misionales, su razón de ser.

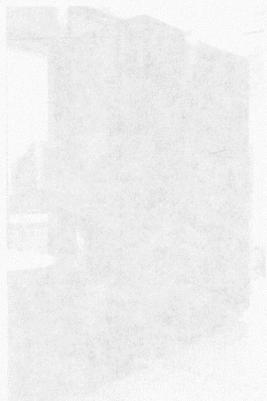
Un almacenamiento masivo, que considere las imágenes como materiales de valor crítico en un centro de información, debe aportar la tranquilidad necesaria del archivo, en relación con la disponibilidad, escalabilidad, actualización y seguridad de las imágenes. Las soluciones de almacenamiento deberán aportar un completo soporte que permita servir, compartir y proteger la información.

Por sus características innatas, por la tradición del tratamiento archivístico en nuestro País y especialmente porque no hay archivo pequeño; la digitalización de los documentos tiene inmerso un gran volumen de datos. Bajo cualquier perspectiva tecnológica de digitalización de archivos, es evidente que la institución requerirá de un almacenamiento considerable en donde repose el producto del escaneo. Este alojamiento de cientos y miles de páginas se tendrá que orientar no sólo con fines de servicio, sino muy seguramente con opciones de respaldo y seguridad que, como mínimo, duplicará la necesidad del almacenamiento requerido.

Hace un par de años, las jerarquías de almacenamiento asignaron diferentes tipos de tratamiento para la información, dependiendo del nivel del uso. Se almacenaron en disco magnético – *on line*- los datos de mayor frecuencia, mientras que los de menor concurrencia en acceso fueron almacenados en medios ópticos.



Servidores de almacenamiento de imágenes



Simultáneamente, aquella información que carecía de actividad regular fue dispuesta fuera de línea en cintas magnéticas entendidas también con propósitos de respaldo. Algunos sistemas, inclusive, llegaron a implementar tendencias híbridas de éxito, en donde el almacenamiento óptico servía por demanda imágenes de alta consulta – *near line* –, pero imposibles económicamente de colocar masivamente en línea, por lo que aparecieron entonces los administradores de discos.

Los importantes avances tecnológicos hacen prever que los archivos no tendrán impactos negativos frente a la necesidad de espacio para el almacenamiento de sus imágenes. Sin embargo, resulta necesario que el proyecto digital contemple su modalidad y recursos de almacenamiento masivo para lograr la disponibilidad y seguridad que persiguen estos documentos.

La tendencia a considerar los medios ópticos de transporte, como sistemas de almacenamiento masivo está desapareciendo. Hoy en día, los archivos en particular y todas las instituciones en general, disponen de medios y sistemas más eficaces, baratos y más funcionales para guardar su información. Más útiles a los propósitos de servicio y respaldo de documentos y por tanto, mejor adaptados a las necesidades de los archivos. Si bien se trata de un estudio técnico de almacenamiento masivo, que muy seguramente tendrá liderazgo en competencia informática, es indudable que los CD- ROM y DVD, entre otros, no son los más apropiados para los efectos e intereses de los repositorios.

Las diversas opciones que brinda la tecnología: redes de área de almacenamiento (SAN), almacenamiento conectado a redes (NAS) y tecnologías de almacenamiento de conexión directa, así como las nuevas tendencias técnicas de los dispositivos de almacenamiento (discos SATA, por ejemplo) soportaran soluciones escalables y económicas, basadas en las necesidades institucionales, como una parte integral de la infraestructura general de tecnologías de la información del Archivo.

El tema de almacenamiento masivo es un componente que requerirá de aportaciones profesionales y soporte técnico e institucional del Archivo, para que la planeación del almacenamiento garantice su administración, desempeño, aumento de capacidad, consolidación

y puesta en marcha; todo esto enmarcado en una etapa previa de evaluación del almacenamiento requerido, que identifique los riesgos potenciales del éxito de las operaciones.

Como valor agregado de la tecnología, los sistemas de almacenamiento masivo de datos (imágenes) vienen caracterizados con técnicas de auto monitoreo, análisis de fiabilidad y recuperación, así como de disponibilidad de redundancia y duplicación de datos. Esto, sin duda, es importante para una institución en donde la información soportada en imágenes marca la diferencia entre un archivo y facilita el acceso y la difusión, a la vez que fortalece la preservación de los materiales originales.



Portal AGN

agregaciones jurídicas, producida y conservada electrónicamente.

El documento digital es un producto de información y de comunicación, se crea, se utiliza y tiene su ciclo de vida en un entorno digital. Se trata de un documento dinámico, en tanto puede variar espacial, temporal e informativamente, e interactivo, en cuanto el usuario lector posee capacidad de decisión en su lectura/exploración/navegación. En breve tiempo será adaptativo, siendo capaz de variar en consonancia con las necesidades de los usuarios individuales.

Las características del documento digital hacen que este sea una unidad informativo-documental de significado completo, compuesta de uno o de varios elementos informativos relacionados entre sí, y que a su vez puede conectarse con otros documentos o con otros elementos, con los cuales pueda establecerse una relación lógica.

Respecto al documento "tradicional", el documento digital puede ofrecer al usuario lector herramientas para desarrollar su propio proceso cognitivo de lectura, comprensión y selección, así como de localización y recuperación de nuevos documentos asociados temáticamente. Mediante estos recursos se integra la interactividad comunicativa como uno de los componentes fundamentales del documento digital. El ejemplo más evidente de documentos digitales lo tenemos en las páginas del *World Wide Web*, aunque no ha sido el primero, ni es el único sistema de documentos digitales. Las aplicaciones hipertextuales, por ejemplo de ellas: HTML, XML, PHP, son técnicas de producción de componentes de documentos digitales; por encima de ellas están el diseño y la arquitectura de la información.

Vocabulario técnico

Definiciones de **Archivo electrónico**

En computación se refiere a cualquier documento electrónico que contiene información que puede ser leída, vista u oída de un programa informático.

Definiciones de **Documento electrónico**

Toda representación de hechos, actos o transacciones jurídicas, producida y conservada electrónicamente.

El documento digital es un producto de información y de comunicación, se crea, se utiliza y tiene su ciclo de vida en un entorno digital. Se trata de un documento dinámico, en cuanto puede variar espacial, temporal e informativamente, e interactivo, en cuanto el usuario/lector posee capacidad de decisión en su lectura/exploración/navegación. En breve tiempo será adaptativo, siendo capaz de variar en consonancia con las necesidades de los usuarios individuales.

Las características del documento digital hacen que éste sea una unidad informativo-documental de significado completo, compuesta de uno o de varios elementos informativos enlazados entre sí, y que a su vez puede conectarse con otros documentos o con otros elementos, con los cuales pueda establecerse una relación lógica.

Frente al documento “tradicional”, el documento digital puede ofrecer al usuario/lector herramientas para desarrollar su propio proceso cognitivo de lectura, comprensión y selección, así como de localización y recuperación de nuevos documentos asociados temáticamente. Mediante estos recursos se integra la interactividad comunicativa como uno de los componentes fundamentales del documento digital. El ejemplo más evidente de documentos digitales lo tenemos en las páginas del *World Wide Web*, aunque no ha sido ni el primero, ni es el único sistema de documentos digitales. Las aplicaciones hipertextuales son ejemplos de ello: HTML, XML, PHP.. son técnicas de producción de componentes de documentos digitales; por encima de ellas están el diseño y la arquitectura de información.

#Digital: Señal que utiliza valores discretos en lugar de un espectro continuo de valores. Proceso de convertir a formato digital los elementos (polígonos, líneas, puntos) que aparecen impresos o dibujados en un mapa o una fotografía aérea.

En general, consiste en trabajar con números en lugar de señales de tipo sinusoidal. Es descomponer las señales en un conjunto de números (dígitos) y recomponerlas en el punto de recepción. La información de los ordenadores está ya digitalizada, salvo en contadas excepciones. La digitalización de todas las señales permite utilizar las mismas técnicas tanto en transmisión como conmutación, con independencia de que las señales correspondan a voz, datos o imágenes.

Proceso mediante el que una imagen original se convierte en información digital que puede ser manipulada e impresa a través de un ordenador.

Definiciones de **Escanear**

v. acción de pasar un documento por un aditamento electrónico que transfiere un texto o una imagen a la computadora.

Acción de utilizar un escáner. // Acción generalmente automática realizada por un programa o un dispositivo que se encarga de “mirar” y/o verificar el funcionamiento de algo. Por ejemplo los *War Dialers*. Otro ejemplo son los escaneos de puertos, para verificar cuáles están activos, etc.

Definiciones de **Indexar**

Organizar una base de datos utilizando un criterio determinado. Por ejemplo, los buscadores de Internet indexan y organizan las páginas web por su contenido e importancia.

Definiciones de **Catalogación**

Apuntar, registrar ordenadamente libros, documentos, etc., formando catálogo de ellos.

Proceso técnico mediante el cual se representa en forma sintética o condensada todos los detalles técnicos, gráficos, tipográficos y bibliográficos de una obra determinada, con el propósito de ubicarla dentro de una colección determinada y de identificarla y recuperarla rápidamente.

Definiciones de **Escáner**

Un escáner es un dispositivo para obtener o “leer” imágenes (escáner de ordenador o de barras) o encontrar un objeto o señal (escáner de un aeropuerto, o de radio).---Hay mucha variedad de escáneres. Entre los que obtienen o leen imágenes hay: *Escáner de ordenador: se utiliza para introducir imágenes de papel, libros, negativos o diapositivas. El Escáner 3-D es una variación de este para modelos tridimensionales.

Un periférico de digitalización de documentos (fotografías, diapositivas, dibujos, textos), aunque también puede digitalizar objetos pequeños. Es una herramienta fundamental para poner imágenes en un sitio web personal.

Aparato capaz de introducir información óptica (documentos, fotos, etc.) en el ordenador. Periférico computacional que traduce un documento o imagen impresa a un formato digital.

Un escáner es un periférico que lee una imagen punto a punto, la digitaliza (la convierte en ceros y unos) y la transmite al ordenador.

Mecanismo que convierte una página impresa o una imagen en una representación digital que puede ser vista y manipulada en el ordenador.

Es un explorador óptico que reconoce y digitaliza las fotografías, los dibujos y las imágenes. Con un *Software* adecuado identifica los textos para incorporarlos después sin necesidad de teclearlos.

Dispositivo electro óptico que sirve para convertir documentos físicos (dibujos, fotografías, textos...) en documentos digitales susceptibles de tratamiento informático.

Dispositivo electroóptico que permite incorporar imágenes planas a un archivo de computadora, para luego ser grabado.

El plural es escáneres. Periférico que permite transferir una imagen desde un papel o superficie y transformarlos en gráficos a computadora (proceso llamado digitalización). En la actualidad existen escáneres que capturan objetos en tres dimensiones. // Todo programa dedicado a escanear, por ejemplo los *War Dialers*.

Es un aparato que lee fotos, textos o dibujos y los transforma en ficheros informáticos para que el ordenador los almacene. Hay varios tipos: los de mano, los más baratos; los de do-

cumentos, que escanean hojas o folios; y los planos, que son capaces de escanear hasta un libro.

Dispositivo de entrada que utiliza la luz para captar información impresa como textos, fotografías, gráficos, etc. y transferirla al ordenador en formato digital.

Unidad de intercambio de información de entrada que digitaliza una imagen para su introducción en el computador.

Definiciones de **Software**:

Software -también conocido como programática y aplicación informática- es la parte lógica del ordenador, esto es, el conjunto de programas que puede ejecutar el **Hardware** para la realización de las tareas de computación a las que se destina. Es el conjunto de instrucciones que permite la utilización del equipo.

En informática, se llama así a los programas y aplicaciones no físicas.

Programas, procedimientos y reglas para la ejecución de tareas específicas en un sistema de cómputo.

La parte “que no se puede tocar” de un ordenador: los programas y los datos.

Equipo lógico de ordenador, formado por aplicaciones y programas.

También conocido como soporte lógico, compendia todo tipo de programas, utilidades, aplicaciones, sistemas operativos, *drivers* que hacen posible que el usuario pueda trabajar con la máquina. El término está totalmente integrado en nuestro idioma ya que, al igual que sucede con **Hardware**, no ha habido nadie capaz de encontrar una traducción capaz de englobar el concepto en una sola palabra.

Un programa de computadora grabado en algún medio de almacenamiento y cargado en la memoria de la computadora (RAM) para su ejecución.

Son los programas, incluyendo procedimientos, utilidades, sistemas operativos, programas de aplicación y paquetes informáticos, implementados para un sistema informático.

Es el conjunto de programas que dirige el funcionamiento de un ordenador.

Software: En computación, procedimientos y reglas lógicas escritas en la forma de programas y aplicaciones, que definen el modo de operación de la computadora. Tienen carácter virtual (en contraposición con el *Hardware*) y están almacenadas en los diferentes tipos de memoria de lectura/escritura.

A diferencia del *Hardware*, es lo que no se ve, es decir los programas y aplicaciones que están guardadas en un disco duro, CD-ROM ó disquetes.

Es el conjunto de instrucciones o programas usados por una computadora para hacer una determinada tarea. Lo contrario al *Hardware*.

Referente a la parte lógica (flexible y programable) de un dispositivo electrónico, incluyendo la microprogramación, el sistema operativo y todas las aplicaciones que puedan albergar sus memorias, ya sean volátiles o no.

Los programas del ordenador o computadora.

Conjunto de programas, documentos, procesamientos y rutinas asociadas con la operación de un sistema de computadoras, es decir, la parte intangible o lógica de una computadora. (Ver Programa)

Programas, aplicaciones utilizadas en los ordenadores (sistema operativo, procesador de texto, hoja de cálculo...).

Todos los componentes lógicos de un sistema informático, que incluyen tanto a las aplicaciones como a los datos sobre los cuales operan. Desde una óptica técnica, se subdivide en *Software* de sistema y *Software* de aplicación.

Término general que designa los diversos tipos de programas usados en computación.

Procedimientos completos, programas, reglas y rutinas que permiten a un ordenador completar una serie de instrucciones o comandos. Esas instrucciones son normalmente definidas como programas. Cuando está activo, reside en estado eléctrico en la memoria de un ordenador.

Conjunto de instrucciones escritas en un determinado lenguaje, que dirigen a un ordenador para la ejecución de una serie de operaciones, con el objetivo de resolver un problema que se ha definido previamente.

Soporte lógico: Conjunto de componentes lógicos (instrucciones o datos) que hacen funcionar una computadora o posibilitan la operación de una red. Se considera *Software* a todo aquello que se pueda almacenar electrónicamente en un sistema computacional. Se suele clasificar el *Software* en dos grupos: *Software* de sistemas (sistema operativo y utilidades) y *Software* de aplicaciones (procesadores de texto, hojas de cálculo, y sistemas de administración de bases de datos).

Se refiere a todas las aplicaciones o programas que se encuentran funcionando en cualquier equipo computacional o de comunicación. Las aplicaciones o *Software* más común son desarrollados con fines administrativos (Procesadores de palabras, hojas de cálculo, contabilidad, entre otros).

Programa de cómputo, sistema de cómputo.

Los sistemas lógicos o programas con los que el hardware realiza aplicaciones (el sistema operativo, los procesadores de texto, los editores de imágenes, los graficadores o programas de diseño, los presentadores, los manejadores de bases de datos, los sistemas de correo electrónico, los navegadores de web, etc.).

Programas escritos en un lenguaje que la computadora entiende y puede ejecutar para realizar una tarea.

Definiciones de **computador**

Una computadora (Hispanoamérica) u ordenador (España) es un dispositivo electrónico compuesto básicamente de un procesador, una memoria y los dispositivos de entrada/salida (E/S).

Sinónimo de Ordenador.

Definiciones de **imagen digital**

Un registro codificado digitalmente de la intensidad de la reflectancia o la radiación de un objeto o área. Cada elemento de la imagen digital tiene un valor de intensidad único para cada una de las bandas del espectro electromagnético empleadas.

Definiciones de **CD**

Un Disco Compacto.

Normalmente se refiere a Compact Disc. Otras abreviaturas relacionadas son: CD-ROM, CD-R y CD-RW. Un Compact Disc para ordenador será capaz de almacenar cerca de 650 Mb de información.

Compact Disc. Disco Compacto. Disco Óptico de 12 cm de diámetro para almacenamiento binario. Su capacidad "formateado" es de 660 Mb. Usado en principio para almacenar audio. Cuando se usa para almacenamiento de datos genéricos es llamado CD-ROM.

Se enciende cuando se recibe una señal de datos (portadora) válida de un módem.

Abreviatura del término inglés Compact Disc, que significa Disco Compacto de Sonido, hecho de una placa de metal recubierta de plástico y cuyo campo físico es de 3 pulgadas de diámetro que se coloca bajo un aparato lector que emite un rayo láser.

Compact Disc.

Compact Disk o Disco Compacto. Disco óptico circular de 12 cm de diámetro y que pesa unos pocos gramos para el almacenamiento de información de forma binaria. La información se almacena de manera digital en éstos. Almacenan hasta 640 MB. Existen gran variedad de CD: CD-ROM, CD-RW ó CD-R... Pueden almacenar música en pistas o *tracks*, las cuales con un programa *ripper* pueden ser guardadas en el disco duro en forma de MP3.

Compact Disc. Sistema de almacenamiento de sólo lectura capaz de almacenar hasta 650MB de datos.

Definiciones de **Memoria RAM**

RAM es el acrónimo inglés de *Random-Access Memory* (memoria de acceso aleatorio).

Siglas de *Random Access Memory*, que lo podríamos traducir como Memoria de Acceso Aleatorio. Es un conjunto de Chips que junto con el microprocesador es parte fundamental del ordenador, es una memoria volátil ya que los datos almacenados en ella se pierden al apagar el ordenador. El tamaño de la RAM determina la rapidez y comodidad de trabajo ante el ordenador, así como el número de programas que podemos utilizar de forma simultánea.

La MEMORIA RAM (Random Acces Memory). Está integrada por uno o más chips y se utiliza como memoria de trabajo donde podemos guardar o borrar nuestros programas y datos.

Random Access Memory. Es la memoria de acceso aleatorio cuyo contenido permanecerá presente mientras el computador permanezca encendido.

Definiciones de **disco duro**

Se llama disco duro (en inglés hard disk, abreviado con frecuencia HD) al dispositivo encargado de almacenar información de forma persistente en un ordenador.

Un disco rígido contiene un medio magnético rotante (en forma de discos) que gira rápidamente. Hay pequeñas cabezas que flotan sobre la superficie de cada disco, y sirven para leer y escribir en el disco a medida que rota. En inglés: «Hard Disk».

Es un dispositivo de almacenamiento, que nació como evolución del *diskette*. Tiene una capacidad mucho mayor (hoy en día es habitual que pasen de los 2.5Gb) y es mucho más rápido, pero no está diseñado para ser llevado de un sitio a otro, sino para permanecer dentro del ordenador (salvo algunas pocas excepciones, que sí son portables).

Sinónimo de la “Unidad C:”. Es un dispositivo interno de grabación y almacenamiento de datos, donde se guardan los programas y los ficheros que queremos conservar. Su característica más importante es el tamaño o la capacidad de almacenamiento de los datos, que se mide en Gigabytes (Gb). 1Gb = 1.024 Mb.

Dispositivo en el que se almacenan los programas, información y aplicaciones que el usuario utiliza cotidianamente en su computadora. Las más recientes versiones cuentan con capacidades mínimas de dos a tres mil millones de bytes (2GB a 3GB).

Hard Disk. Dispositivo de almacenamiento permanente de la información computadorizada. Los datos o programas almacenados en este dispositivo no se pierden cuando la computadora se apaga. Pueden presentarse en diferentes capacidades y formas.

Unidad de almacenamiento permanente de información. Éste es el que guarda la información cuando apagamos el computador. Aquí se guardan la mayoría de los programas y el

sistema operativo. Su capacidad de almacenamiento se mide en Megabytes (Mb) o Gigabytes (Gb). $1024 \text{ Mb} = 1 \text{ Gb}$.

Dispositivo magnético integrado (en algunos casos) en las computadoras personales. La información es grabada en estos medios de almacenamiento magnéticos en el código de representación que use el computador, ya sea EBDIC ó ASCII, tienen capacidades que van desde 10 megabytes hasta 800 megabytes y pueden ser usados tanto en microcomputadoras, como en minicomputadoras y *mainframes*; siendo sus medidas físicas distintas según el equipo al que estén destinados.

Definiciones de **Bytes**

Se describe como la unidad básica de almacenamiento de información, generalmente equivalente a ocho bits, pero el tamaño del byte depende del código de información en el que se defina.

Es una medida que determina el tamaño de los archivos que puedes utilizar para publicitar la imagen de tus productos y/o servicios dentro de tu tienda virtual o en las páginas de eje comercial. Puedes asesorarte con el distribuidor que más te convenga y, si tienes alguna duda al respecto, no dudes en escribirnos a: clientes@ejecomercial.com.mx.

Definiciones de **KB**

Kb es la abreviatura de: #Kilobases = 1.000 pares de bases de ADN. #Kilobits = 1.024 (210) bites. A veces se confunde con KB, que es la abreviatura correcta de Kilobytes.

Un kilobyte que se abrevia como K, KB ó K-byte son 1024 (210) bytes. Se trata de una unidad de medida común para la capacidad de memoria o almacenamiento de las microcomputadoras. Las PC de IBM más antiguas, por ejemplo, tenían una capacidad máxima de 640 K, o alrededor de 640 000 caracteres de datos.

Kilobyte, múltiplo del byte equivalente a 1024 bytes. Más correcta, aunque menos utilizada, es la forma “kb”; también se emplea “Kb”.

Unidad empleada para medir la longitud de los fragmentos de ADN constituidos por una serie de bases. $1 \text{ Kb} = 1.000 \text{ bases}$.

Abreviatura de KiloByte.

Definiciones de *Backup*

Copia de seguridad. Acción de copiar documentos, archivos o ficheros de tal forma que puedan recuperarse en caso de fallo en el sistema.

Copia de seguridad.

Permite copiar el contenido de toda la ruta especificada fuera del disco actual para efecto de recabar respaldos de tipo permanente.

Referente al equipo y a los procedimientos a los sistemas de computación o las bases de datos donde se produce una copia y se almacena en el caso de que se necesite recuperar información.

Acción de copiar archivos o datos, de forma que estén disponibles en caso de que un fallo produzca la pérdida de los originales. Se recomienda hacer esta sencilla acción continuamente para evitar problemas irremediables.

Hacer un *backup*. “Bacap”, en argot. Son copias de seguridad. Por ejemplo, si tiene archivos muy importantes en el disco rígido y los copia en un disquete por una cuestión de seguridad, lo que hizo fue un *backup*. Es la copia de datos que se realiza sobre medio de almacenamiento externo, tal como un disquete, cinta, CD-ROM, o cualquier otro soporte posible. Como la información almacenada en el disco rígido es susceptible a accidentes o pérdidas, es necesario tener una copia de respaldo.

Copia de seguridad. Es la copia de información (total o parcial) del disco duro, de un CD, de un disquete u otro medio de almacén informativo. La copia es almacenada en otro disquete, otro disco duro, otro CD, discos virtuales o cualquier otro medio de almacenamiento informativo. Es utilizada para respaldar información valiosa para el usuario, en caso de que la original se pierda.

Salvaguardia. Se entiende por hacer *backup* el obtener una copia de seguridad de un disco, una *diskette* o de un conjunto de ficheros.

Copia de seguridad. Se hace para prevenir una posible pérdida de información.

Copia de Seguridad. Los programas de *backup* son los que permiten meter en disquetes o en cintas (de *backup*) la información del disco duro, para recuperarla si hubiese algún problema.

Respaldo de archivos, equipos o procedimientos alternativos disponibles para usarlos en caso de emergencias producidas por fallas totales o parciales de un sistema computacional.

Definición de KB

Kilobytes. Medida de memoria o almacenaje de mil bytes aproximadamente.

Lector de pantalla. Es un programa de *Software* que lee en voz alta al usuario el contenido de la pantalla. Lo usan principalmente los ciegos. Habitualmente los lectores de pantalla pueden leer textos que estén impresos, no pintados.

Unidad de información equivalente a 1.024 bytes.

Abreviatura para kilobyte.

Abreviatura de Kilobyte. Otra unidad de medida de información.

Puede ser usado como abreviación del teclado del ordenador o, más comúnmente “Kb” significa un kilobyte de datos.

Kilobytes. Unidad de tamaño de archivos. Equivale a 1024 bytes.

centros.edu.xunta.es/iesxunqueira1/tecnoweb/lecciones/dicc.htm

Medida de información. Contiene 1024 Bytes.

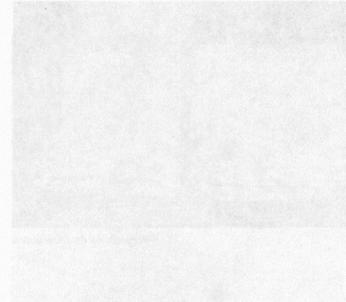
Definiciones de Megabytes

(Megabytes) 1.024 Kilobytes.

Definiciones de Memoria

En informática la Memoria es el lugar donde el ordenador almacena los programas y los datos que está utilizando.

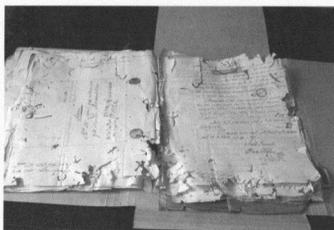
Los ordenadores sólo pueden procesar inmediatamente la información que tienen en su



memoria. Por eso, cuando deseamos ejecutar un programa, debemos cargarlo (almacenarlo) previamente en la memoria.

Zona de almacenamiento para los datos binarios o de programas. En una computadora, la memoria se divide en dos partes: una memoria electrónica rápida e integrada a la computadora y una memoria externa, más lenta, compuesta, por ejemplo, de unidades de disco o de cintas que utilizan medios de almacenaje magnético. La memoria interna puede ser ROM, RAM etc. y la externa disquetes, etc. La memoria se caracteriza por su tiempo de acceso, su capacidad, su naturaleza, etc.

Módulos electrónicos en los cuales el computador almacena la información cuando está encendido. Existen de tres clases de memoria que utiliza el procesador: RAM, video, y caché. El caché viene en cantidades más pequeñas y se utiliza para acceder rápidamente a la información. La RAM se utiliza para almacenar la información de los programas y el sistema operativo. La memoria de video se utiliza para almacenar la información que se muestra en la pantalla.



Preparación de documentos

ANEXOS

ANEXO A

Preparación de los Documentos para Proceso

Es sabido que un buen trabajo archivístico previo a la aplicación de cualquier técnica reprográfica, potencia la calidad final obtenida a un menor costo, por la mejor disposición de los documentos y consecuente disminución de registros para direccionamiento.

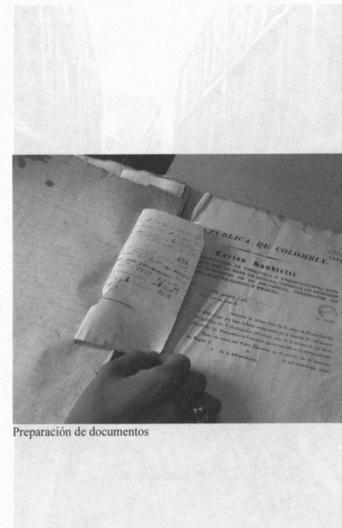
Preparación documental

Ésta es un área que normalmente es olvidada. La condición del archivo y la cantidad de mano de obra necesaria para preparar la digitalización es usualmente subestimada. En muchos casos, las organizaciones o compañías piensan que sus documentos están en mejores condiciones de lo que realmente están y este proceso puede definir el éxito y la ventana de tiempo en un proyecto de digitalización.

La preparación documental incluye hacer separación de los documentos en color de los documentos en blanco y negro, separación por tamaño, grosor, y documentos que incluyan códigos de barra y se utilicen como separadores. El anterior proceso es lo que permitirá el escaneo por lotes; lo cual hace referencia a la digitalización de grandes cantidades de documentos mejorando tiempos de captura y optimizando el tiempo.

Las tasas de digitalización están basadas en un mundo ideal en donde nunca se presentan errores; pero estos procesos están basados en la condición del material que se piensa capturar y un inadecuado proceso de clasificación podría ocasionar fallas en el *Hardware* disminuyendo la productividad.

A continuación, consideraciones que se deben tener en cuenta antes de determinar qué se debe convertir:



Preparación de documentos



Depósito documental AGN

Condición de los documentos: ¿Los documentos contienen grapas?, ¿contienen notas en *Post-it* que deban ser integradas a la imagen digital?, ¿hay notas manuscritas que se deban integrar a la imagen digital?

Preparación de carpetas: ¿Los documentos están en carpetas?, ¿las carpetas serán parte del proceso de indexación?

Pre-Indexación: ¿Los documentos tienen campos predeterminados que serán usados en el proceso de indexación?, ¿quién o qué proveerá la información de indexación?

Post-Conversión: Cuando la imagen esté lista ¿qué pasará con los documentos?

La preparación documental que simplemente consiste en disponer la pieza física de papel para ser digitalizada, puede ser la clave del éxito en un proyecto de digitalización. Una pobre preparación documental puede ocasionar sobrecostos y demoras en el proceso.

En Normas Técnicas Colombianas específicas de Reprografía, como la NTC 3723 Anexo A, y otros documentos publicados por el Archivo General de la Nación, se insiste en la necesidad de preparar correctamente los documentos para el proceso de transcripción; por esta razón, retomamos las recomendaciones y las contextualizamos a la técnica reprográfica de que se trata:

Debe examinarse el archivo documental con el fin de cerciorarse que todos los documentos estén en el orden requerido según el sistema de archivo vigente. Las hojas, páginas, fólderes u otras unidades archivísticas que estén mal colocadas, deberán intercalarse en el orden correcto. Se deben retirar los documentos extraños o no-registrados. Debe hacerse todo lo posible por obtener los documentos faltantes, de modo que se logre un copiado o transcripción de registros tan completa como sea posible.

Si resulta imposible la consecución de páginas o documentos faltantes, se deben insertar símbolos o tarjetas de “documentos faltantes” como los descritos en la NTC 4080; otros

símbolos gráficos apropiados o tarjetas especiales de la referida norma que sean aplicables para la descripción de la condición de los documentos y del archivo deben ser colocados entre las páginas del archivo antes de comenzar el copiado.

Si el tamaño o la forma física no permiten el copiado de ciertos documentos, se deben registrar debidamente las tarjetas que identifiquen cada ítem y den explicación de sus omisiones. Frecuentemente, es necesaria una paginación (foliación), para mantener el orden y asegurar que la colección documental esté completa. Tenga en cuenta que, antes de proceder a copiar o transcribir las imágenes, todos los ganchos, clips u otros sujetadores deben ser retirados de los documentos. Este procedimiento es obligatorio si se usa escáner de alimentación automática (Proceso similar al de los equipos rotativos de microfilmación).

Se pueden ignorar las imperfecciones que no afectan el texto o las ilustraciones. Los documentos con rasgaduras menores usualmente se pueden copiar correctamente sin que haya pérdida de texto. La manipulación de los originales no es tan crítica si los documentos se van a descartar después de la aprobación de las imágenes obtenidas en la transcripción; sin embargo, recomendamos prudencia.

Los documentos que **no** tengan carácter histórico pueden repararse, si es necesario, con cinta adhesiva transparente (preferiblemente por el vuelto del documento donde no hay escritura). Se debe tener especial cuidado con los documentos delicados o en mal estado, ya que el paso de estos por los escáneres rotativos puede deteriorar aún más los papeles.

El operario de un escáner rotativo debe estar conciente de que el adhesivo de los bordes de la cinta pegante puesta sobre los papeles reparados, puede interferir con la alimentación de los documentos, causar atascamiento del escáner, dañar o ensuciar las guías de vidrio, obstruyendo así el paso óptico.

Los documentos de carácter histórico no deben pasarse por equipos rotativos; debe usarse solamente escáner de cama plana; Los defectos mayores de los documentos históricos no deben repararse antes de copiarse puesto que se puede reducir la legibilidad del texto. En este

caso, **no** deberá usarse la cinta adhesiva. Los fragmentos deben colocarse preferiblemente en un fólter plástico para su copiado y observar las recomendaciones de los restauradores.

Si se tiene que desencuadernar para hacer legibles las márgenes del documento, deberá hacerse bajo supervisión de un restaurador con experiencia.

Cada entidad debe definir previamente la política aplicable al proceso, en especial respecto a las características que espera capturar y retener en las imágenes a fin de conservar los aspectos relativos a su autenticidad; en tal caso, cuando por ejemplo la densidad del documento original varía considerablemente, se pueden tomar dos imágenes con diferentes resoluciones. Debe usarse entonces el símbolo de “repetición de imagen” (ver NTC 4080). Asimismo, no se deben retirar las manchas; en lugar de ello, se puede usar un *Software* especializado para hacer más legible la imagen copiada, (siendo conservada tanto en su archivo original como en su forma mejorada).

Aunque los medios permiten contener una gran cantidad de imágenes, no está demás recomendar que, se debe determinar o prever el número de imágenes que cabrán adecuadamente en una unidad de grabación; Dicha información es obtenible generalmente a partir del cálculo de resolución versus memoria requerida, por el número de documentos. Esto tiene por objeto diseñar los cortes de modo que sean lógicos en el archivo. Por ejemplo, finalizando temas, fechas o volúmenes.

BIBLIOGRAFÍA

- BRITISH STANDARD. Storage, transportation and maintenance of media for use in data processing and information storage. London, (BS 4783)
- Bundles for the Perpetual Preservation of electronic documents and associated objects. London, (BS 7978)
- CASILIMAS Rojas, Clara Inés y Ramírez Moreno, Juan Carlos. Fondos Acumulados Manual de Organización. Bogotá: Archivo General de la Nación, 2004, 80 p.
- COLOMBIA. ARCHIVO GENERAL DE LA NACIÓN. Documentos Sistema Nacional de Archivos. Bogotá: 1994, 63 p.
- Plan estratégico del Archivo General de la Nación 2003 – 2006. Bogotá: 2003, 40 p.
- Reglamento General de Archivos. Bogotá: 1997, 94 p.
- COLOMBIA. PRESIDENCIA DE LA REPÚBLICA. Guías para el fortalecimiento de las unidades de control interno de las entidades gubernamentales en los aspectos de informática y telemática. Bogotá: 1994, 120 p.
- CORTES Alonso, Vicenta. Manual de Archivos Municipales. 2° edición. Madrid: Anabad, 1989, 158 p.
- DAVENPORT, Thomas H. Ecología de la información. México: Oxford University Press, 1999, 296 p.
- DOYLE, Murielle y FRENIERE André. La preparación de manuales de gestión de documentos para las administraciones públicas: estudio del RAMP. Bogotá: Archivo General de la Nación, 1991, 47 p.

EUROPEAN COMMISSION BY SERCO CONSULTING WITH FUNDING FROM THE IDABC PROGRAMME. Model Requirements for the Management of Electronic Records. 2008. <http://www.moreq2.eu/index.htm>

GALLEGO Domínguez, Olga y López Gómez, Pedro. Introducción a la Archivística. Bilbao: Departamento de cultura y turismo del Gobierno Vasco, 1989, 121 p.

GODOY, Julia et al. Tablas de retención y transferencias documentales. Bogotá: Archivo General de la Nación, 2001, 92 p.

GRUPO DE ARCHIVEROS MUNICIPALES DE MADRID. Tipología Documental Municipal. Madrid: Grafes, 1992, 130 p.

GUZMÁN Mora, Leonardo. Pautas para el establecimiento de políticas y planes de automatización de archivos. Bogotá: Archivo General de la Nación, 1998, 54 p.

HEREDIA Herrera, Antonia. Archivística General. Teoría y Práctica. 4ª edición. Sevilla: Servicio de publicaciones de la Diputación de Sevilla, 1992, 389 p.

INSTITUTO COLOMBIANO DE NORMAS TÉCNICAS. Compendio Tesis y otros trabajos de grupo. Bogotá: 2002, 128 p.

INTERNATIONAL ORGANIZATION FOR STANDARDIZATION. Codes for the representation of names of languages. (ISO 639).

----- Codes for the representation of names of countries (ISO 316616)

----- Data elements and interchange formats – Information interchange – Representation of dates and times (ISO 8601)

----- Information technology – 8-bit single-byte coded graphic character sets (ISO 8859).

- Information technology – database languages – SQL (ISO 9075)
- Information technology – Universal Multiple-Octet Coded Character Set (ISO 10646)
- Information retrieval – application service definition and protocol specification (ISO 23950)
- Information and documentation — Records management. (ISO/TR 15489) Parts 1 and 2.

LA CONSERVACIÓN DEL PATRIMONIO DOCUMENTAL. (2° : 1996: Bogotá) Memorias del V seminario del Sistema Nacional de Archivos. Bogotá, Archivo General de la Nación, 1997, 212 p.

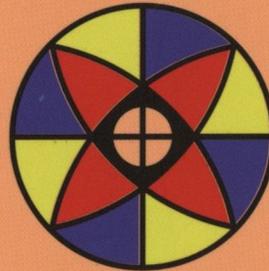
LOS ARCHIVOS DE CARA AL SIGLO XXI. (5° : 1994: Bogotá) Memorias del III seminario del Sistema Nacional de Archivos. Bogotá, Archivo General de la Nación, 1995, 236 p.

LOS ARCHIVOS Y LAS NUEVAS TECNOLOGÍAS. (3° : 2001: Bogotá) Memorias del X seminario del Sistema Nacional de Archivos. Bogotá, Archivo General de la Nación, 2001, 250 p.

RHOADS, James B. La función de la gestión de documentos y archivos en los sistemas nacionales de información: estudio del RAMP. Bogotá: Archivo General de la Nación, 1995, 82 p.

SISTEMA NACIONAL DE ARCHIVOS DE COLOMBIA: BALANCE Y PERSPECTIVAS. (7° : 1999: Bogotá) Memorias del VIII seminario del Sistema Nacional de Archivos. Bogotá, Archivo General de la Nación, 2000, 253 p.

COLOMBIA



**ARCHIVO
GENERAL
DE LA NACIÓN**

Archivo General de la Nación - Colombia

Establecimiento Público adscrito al Ministerio de Cultura

Carrera 6a N° 6-61 - Tel.: 3282888 Email: contacto@archivogeneral.gov.co

Página Web: www.archivogeneral.gov.co

Bogotá, D.C. - Colombia