

PONTIFICIA UNIVERSIDAD CATÓLICA DEL ECUADOR

FACULTAD DE INGENIERÍA

ESCUELA DE SISTEMAS



TEMA:

***DESARROLLO DE PROTOTIPO DE UN
SISTEMA DE MANEJO DE DOCUMENTOS***

AUTORA:

MARÍA AUGUSTA MATEUS HERRERA

DIRECTOR DE DISERTACIÓN:

MSC. GUIDO OCHOA

QUITO, NOVIEMBRE DEL 2003

Gracias a Dios por brindarme
la vida, la salud, y el amor
en cada uno de los instantes de mi vida.

Gracias a mis Padres,
esos seres comprensivos y cariñosos
que me dieron la oportunidad de
educarme y prepararme
en esta carrera que culmino
con el presente trabajo.

Gracias a mis hermanos
Rosita y David
por ser mi alegría y compañía
sincera todos los días.

Gracias a mis amigos en especial
a Vivi y a mi familia
por ser el apoyo en todo aspecto.
Y a los profesores que marcaron
mi vida universitaria con sus
enseñanzas y consejos
que recordaré y pondré en práctica
todo el resto de mi vida.

A todos Ustedes dedico este logro con mucho cariño.

PROLOGO

En la actualidad las empresas generan gran cantidad de documentos como fuentes de información, en distintos formatos texto, imagen, audio y video y diversos medios físicos como papel, cd, disquette, etc. La problemática que surge a partir de ello es cómo organizarlos y almacenarlos de tal forma que puedan ser recuperados eficaz y rápidamente.

Para conseguirlo se presentan los sistemas de gestión documental que facilitan la administración de documentos en una empresa, enfocándose en tres tareas principales: digitalización, indexación y recuperación de documentos.

El presente trabajo es un conjunto de conocimientos y experiencias en el manejo documental. Está dirigido a las personas y empresas que buscan una alternativa para administrar documentos recibidos y generados, y que requieran conservarlos digitalmente por las ventajas que conlleva en el ahorro de espacio, tiempo y papel.

Los primeros capítulos detallan la teoría del manejo documental, los tipos de documentos que se pueden encontrar en una empresa y el ciclo vital de los documentos, además se dedica un capítulo a los sistemas automatizados para el manejo de documentos, su arquitectura, clasificación y aplicaciones en diversas áreas; y también los tipos de presentaciones de un documento para texto, imagen, audio y video.

En los capítulos finales se verá el análisis, diseño e implementación del prototipo de sistema de manejo de documentos planteado, así como su funcionalidad y pruebas realizadas.

TABLA DE CONTENIDOS

1 INTRODUCCION A LA ADMINISTRACION DE DOCUMENTOS 1

1.1 MANEJO DE DOCUMENTOS.....	1
1.1.1 Historia.....	1
1.1.2 Objetivos.....	3
1.1.3 Ventajas.....	4
1.2 DISCIPLINAS QUE INVOLUCRAN GESTIÓN DOCUMENTAL.....	5
1.2.1 Archivística.....	5
1.2.2 Biblioteconomía.....	6
1.2.3 Documentación.....	7
1.3 TIPOS DE DOCUMENTOS.....	8
1.4 CICLO VITAL DE LOS DOCUMENTOS.....	10
1.5 VALOR DE LOS DOCUMENTOS.....	13

2 SISTEMAS DE GESTIÓN DOCUMENTAL..... 15

2.1 ARQUITECTURA DE LOS SISTEMAS DE GESTION DOCUMENTAL.....	15
2.2 TIPOS DE SISTEMAS.....	17
2.3 CARACTERÍSTICAS DE LOS SISTEMAS ESTUDIADOS.....	19
2.4 FUNCIONES BÁSICAS.....	23
2.5 VENTAJAS Y PROBLEMAS QUE SE SOLUCIONAN.....	23
2.6 SISTEMAS DE BASES DE DATOS RELACIONALES VS DOCUMENTALES.....	25
2.7 APLICACIONES DE LOS SISTEMAS DE MANEJO DE DOCUMENTOS.....	26
2.8 NUEVAS TENDENCIA TECNOLÓGICAS – TRATAMIENTO AUDIOVISUAL DE LA INFORMACIÓN.....	27

3 PRESENTACIÓN Y CONSERVACIÓN DE LA INFORMACIÓN..... 29

3.1 TEXTO.....	29
3.1.1 Proceso de digitalización.....	30
3.1.2 Formatos de almacenamiento.....	31

3.1.3	<i>Ventajas y desventajas</i>	32
3.2	IMÁGENES.....	33
3.2.1	<i>Proceso de digitalización</i>	33
3.2.1.1	<i>Cómo funciona un escáner</i>	33
3.2.1.2	<i>Resolución de un escáner</i>	34
3.2.1.3	<i>Tipos de escáner</i>	34
3.2.2	<i>Parámetros para la digitalización de imágenes</i>	36
3.2.3	<i>Formatos de almacenamiento</i>	37
3.2.4	<i>Ventajas y desventajas</i>	38
3.3	AUDIO.....	39
3.3.1	<i>Proceso de digitalización</i>	39
3.3.1.1	<i>Parámetros para la digitalización de audio</i>	40
3.3.1.2	<i>Soportes digitales de audio</i>	40
3.3.2	<i>Formatos de almacenamiento</i>	41
3.3.3	<i>Ventajas y desventajas</i>	42
3.4	VIDEO.....	43
3.4.1	<i>Proceso de digitalización</i>	43
3.4.2	<i>Formatos de almacenamiento</i>	44
3.4.2.1	<i>Métodos de compresión</i>	44
3.4.3	<i>Ventajas y desventajas</i>	45
4	ANÁLISIS Y DISEÑO DEL SISTEMA	46
4.1	ANÁLISIS.....	46
4.1.1	<i>Requerimientos</i>	47
4.1.2	<i>Alcance del sistema</i>	48
4.2	DISEÑO.....	54
5	DESARROLLO, IMPLEMENTACIÓN Y PRUEBAS	57
5.1	DESARROLLO DEL SISTEMA.....	57
5.2	IMPLEMENTACIÓN DEL SISTEMA.....	59
5.2.1	<i>Recursos de Software</i>	59
5.2.2	<i>Recursos de Hardware</i>	59

5.2.3 Funcionamiento de la aplicación.....	60
5.3 PRUEBAS.....	73
6 CONCLUSIONES Y RECOMENDACIONES	76
6.1 CONCLUSIONES.....	76
6.2 RECOMENDACIONES	78
ANEXOS	80
ANEXO 1 DICCIONARIO DE DATOS	80
ANEXO 2 CÓDIGO DE XELUPLOAD.ASP PARA SUBIR ARCHIVOS AL SERVIDOR.....	82
ANEXO 3 ALGORITMO DE BÚSQUEDA DE DOCUMENTOS	87
<i>Función Sacarterminos.....</i>	87
<i>Procedimiento Almacenado SP_CALC_C</i>	87
<i>Procedimiento Almacenado SP_UPD_SUMA</i>	89
<i>Procedimiento Almacenado SP_CALC_C_AND.....</i>	91
<i>Procedimiento Almacenado SP_UPD_CERO.....</i>	93
<i>Vista V_DOC_TIPO.....</i>	93
<i>Llamada desde la interfaz de búsqueda.....</i>	93
BIBLIOGRAFÍA	95
LIBROS	95
SITIOS WEB	95

INDICE DE TABLAS

TABLA 2-1: CUADRO COMPARATIVO BASE DE DATOS RELACIONES VS DOCUMENTALES	25
TABLA 3-1: TIPOS DE ESCÁNER	35
TABLA 3-2: BITS POR PIXEL	36
TABLA 3-3: VENTAJAS Y DESVENTAJAS EN LA DIGITALIZACIÓN DE IMÁGENES	39
TABLA 3-4: SOPORTES DIGITALES DE AUDIO	41
TABLA 3-5: TIPOS DE RESOLUCIÓN DE AUDIO	42
TABLA 5-1: TIEMPOS DE RESPUESTA EN SUBIR DOCUMENTOS	74
TABLA 5-2: TIEMPOS DE RESPUESTA EN BUSCAR DOCUMENTOS	74

INDICE DE GRAFICOS

GRÁFICO 1-1: CICLO VITAL DE LOS DOCUMENTOS	12
GRÁFICO 2-1: ARQUITECTURA DE SISTEMAS INFORMÁTICOS DE RECUPERACION DE DOCUMENTOS	15
GRÁFICO 4-1: DIAGRAMA DE PROCESOS PARA MANEJADOR DE DOCUMENTOS	50
GRÁFICO 4-2: DIAGRAMA DE PROCESOS PARA BUSCADOR DE DOCUMENTOS	53
GRÁFICO 4-3: DIAGRAMA DE PROCESOS PARA MÓDULO ADMINISTRATIVO	54
GRÁFICO 4-4: MODELO E-R	55

INDICE DE CUADROS

CUADRO 3-1: FRECUENCIAS DE GRABACIÓN MÁS TÍPICAS	40
CUADRO 3-2: FORMATOS DE PRESENTACIÓN DE LOS PRINCIPALES MATERIALES PARA DOCUMENTOS AUDIOVISUALES	44

INTRODUCCIÓN A LA ADMINISTRACIÓN DE DOCUMENTOS

1.1 Manejo de documentos

El manejo de documentos lo constituyen un conjunto de procedimientos diseñados para organizar de manera ideal, localizar eficaz y rápidamente; y proteger al máximo los documentos que deben conservarse y, posteriormente, destruir en forma sistemática los que ya no son útiles.

Es así, como el manejo de documentos incluye todas aquellas operaciones que permiten administrar y controlar la masa documental de una empresa, mediante técnicas de clasificación y ordenamiento que organicen de forma eficiente la información. Dichas técnicas se establecen basándose en el conocimiento de las actividades y los tipos de documentos que tiene la empresa.

1.1.1 Historia

En tiempos antiguos la mayor parte de transacciones financieras se realizaban verbalmente; se respetaba la palabra de las partes y aún de los antepasados en casos de ventas de terrenos, derecho de tránsito y otros tratos. Muchas personas podían vivir con solamente tres pruebas escritas de su

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

existencia: el registro de su nacimiento, el acta de matrimonio y el acta de defunción. Por ende, en esos tiempos no era necesario el manejo de documentos.

Con el paso de los años y el aumento de la población mundial, la producción de documentos es cada vez mayor. La palabra ya no tiene la misma validez y por ello se acompaña de documentos escritos. La imprenta, la dactilografía y el procesamiento de textos contribuyeron a este fenómeno.

A finales del siglo XIX en Estados Unidos, surge una teoría para la gestión de la administración mediante la agilización de la gestión documental. Se trataba de una administración que producía muchos documentos y su mala gestión repercutía en la administración. Para 1870, se plantea en el Congreso la eliminación de documentos innecesarios, aunque no se sigue adelante con este proyecto, se logra crear conciencia y así se crean solo documentos imprescindibles, y se comienza además una política de expurgo. Pero solo hasta la 2da. Guerra Mundial, se plantean verdaderas soluciones al problema de la administración documental, donde frente a la necesidad de consultar documentos para tomar decisiones rápidas se revela el mal funcionamiento de la gestión de documentos. Gracias a ello, en 1950 aparece la primera ley de archivos, cuya teoría plantea reducir gastos y aumentar la eficacia. Se crea una política de control de documentos desde que se crean, e incluso antes, y así se plantea en 1947 el ciclo vital de documentos.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

En 1934, se crea el Archivo Nacional en Estados Unidos. Posteriormente, en 1952, los países europeos van adoptando poco a poco este sistema, destacándose Reino Unido y España.

Las empresas de hoy se enfrentan al reto de manejar grandes cantidades de documentos que aumentan de forma desenfrenada. Un estudio realizado por la firma Deloitte & Touche en 1995, estimó que casi un tercio de las horas laborales son dedicadas a la búsqueda, lectura, modificación, escritura o crítica de un documento. Este mismo estudio indicó que la masa de papel consume un cuarto de los rubros laborales y cerca del 15% de los presupuestos futuros en las empresas.¹ Claramente, se puede ver que es inminente encontrar una solución a este problema.

La cada vez mejor y robusta infraestructura tecnológica de computadoras y redes, y el bajo costo de dispositivos de almacenamiento; ha hecho que las compañías piensen en introducir a sus ambientes soluciones para el manejo de documentos en los que el papel sea almacenado en baratos sistemas ópticos, en lugar de costosos archivadores.

1.1.2 Objetivos

El principal objetivo del manejo de documentos es:

Organizar los documentos de tal forma que puedan ser encontrados inmediata y eficazmente por cualquier persona.

¹ Documtreev – A Modern View of Information Management, www.treev.com

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Otros objetivos que se buscan son:

- Establecer un control de la documentación emitida o recibida por la empresa.
- Eliminar documentos que no sean necesarios
- Determinar el tiempo de conservación de los documentos
- Eliminar copias inútiles de documentos
- Proteger documentos confidenciales y/o vitales para la empresa.

Además de los objetivos que un adecuado manejo de documentos puede lograr existen otros objetivos que, específicamente un sistema automatizado de manejo de documentos debe cumplir, y son:

- Proveer los mecanismos para convertir documentos en papel a formatos electrónicos.
- Ahorrar tiempo, espacio y papel.
- Permitir que varios usuarios accedan a un documento al mismo tiempo.
- Proteger a los documentos de pérdidas y deterioros por el paso del tiempo.
- Facilitar la recuperación segura, rápida y selectiva de un documento
- Eliminar el problema de pérdida o traspapelado de documentos
- Eliminar operaciones mecánicas

1.1.3 Ventajas

Mediante un adecuado manejo documental se perciben las siguientes ventajas:

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

- Arreglo lógico y organizado de documentos
- La búsqueda de un documento es fácil y rápida
- Se utilizan procedimientos para la gestión documental.
- Se reducen costos relacionados con el manejo de documentos tales como tiempo, mano de obra, espacio físico.
- Se protegen documentos importantes para la empresa y/o confidenciales.
- Se destruyen duplicados innecesarios de documentos o información que ya no se requiere.
- Se facilita el entrenamiento del personal nuevo.
- Se eliminan riesgos de perder o traspapelar un documento.
- Se aplican políticas para determinar el tiempo de conservación de los documentos.

1.2 Disciplinas que involucran gestión documental

Las disciplinas que a continuación se mencionan brevemente, enfocan su función en la gestión documental, aunque sólo recientemente hayan aplicado software para ello.

1.2.1 Archivística

La archivística es la ciencia que estudia principios y técnicas a aplicarse en los archivos.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Esta disciplina ha evolucionado desde tiempos históricos, así tenemos a la Archivística Empírica que empezó con las civilizaciones más antiguas (400 A.C), que manejaba archivos de palacios y templos. Luego, en el siglo XVI pasó a ser tratada como Disciplina Jurídica cuando los archivos fueron elaborados por juristas. Más tarde, en el siglo XIX, con la Archivística como Disciplina Histográfica nace el principio fundamental de esta ciencia, el principio de la procedencia. Según la cual se ordenan los documentos de acuerdo a la institución que los haya creado.

Ya en los tiempos modernos, se crean sistemas nacionales de archivo y se integra a la Archivística en la gestión de las organizaciones, con ello también surge el principio del ciclo vital del documento.

Actualmente el reto más importante para la Archivística son los documentos electrónicos. Puesto que, esta ciencia debe conseguir que se puedan consultar los archivos y además, que a lo largo de su vida tengan garantía de autenticidad.

1.2.2 *Biblioteconomía*

La Biblioteconomía tiene por objeto estudiar los principios teóricos y prácticos de organización de las bibliotecas como unidades de información documental.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

La biblioteca es un sistema de información que ha dado lugar a distintos tipos de bibliotecas: nacionales, públicas, especializadas; y manifestándose en diversos campos (adquisición, préstamo, etc.) a través de diferentes formas de organización (redes y sistemas) y en distintos ámbitos (local, nacional e internacional). A su vez, está sometida a importantes cambios externos como son: la diversificación de los soportes en los que aparece la información, el desarrollo de la informática y las telecomunicaciones.

1.2.3 Documentación

Como Otlet define en su libro *Traité de Documentation* (1934), la Documentación, es la disciplina relacionada al estudio del documento como fuente de información para basándose en ella, obtener nueva información o una toma de decisiones que generará nuevamente información.²

Además se la define como el estudio de los procedimientos requeridos para el tratamiento técnico de los documentos para la eficaz difusión de los mensajes que ellos contienen.

La Documentación como actividad abarca cuatro modalidades:

1. Recopilación, organización, almacenamiento, recuperación y difusión sistemática de información especializada, científica o técnica.

² Artículo El concepto de la Documentación en la doctrina española, Juan Ros García, Universidad de Murcia.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

2. Adquisición, organización, almacenamiento, recuperación y difusión de documentos.
3. Colección de documentos sobre una materia determinada.
4. Información descriptiva necesaria para iniciar, desarrollar, operar y mantener sistemas informáticos.

Para la actualidad, surge la *Documentación Automática*³, la misma que engloba no sólo el proceso documental completo, desde la producción de un documento hasta su difusión, sino también la intervención de la tecnología para ofrecer al usuario información “optimizada”, es decir actualizada, universal, rápida y veraz.

1.3 Tipos de documentos

Los documentos pueden tener diversas formas físicas (papel, microficha, disquete, disco óptico, cinta magnética, planos, fotografías, banda sonora, película, video, etc) y estar bajo muchas formas intelectuales (estados contables, presupuestos, contratos, etc).

Entre los principales tipos de documentos *administrativos* que se puede encontrar en una empresa están los siguientes:

³ Hacia un concepto básico de Documentación Automática/Documática, www.tramullas.com/documatic/1-4.html.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

1. De la administración de la empresa
2. Financieros
3. De la administración del personal.
4. De la administración de bienes
5. Legales

De la administración de la empresa:

Dentro de esta clasificación se encuentran los documentos constitutivos de la organización, las actas de instancias de decisión, las políticas, normas y directivas, los reportes y estudios.

Documentos financieros:

Son los más numerosos y complejos, tienen presentaciones muy diversas y suelen cambiar de formato e incluso de título muy frecuentemente. Por lo general estos documentos van acompañados de documentos de soporte (justificantes, notas manuscritas, etc). Entre ellos están el presupuesto, el estado financiero, los documentos relacionados con impuestos, documentos que registran préstamos, pagos y contabilidad.

De administración de personal:

Se utilizan para administrar y contratar al personal. Principalmente forman parte de este grupo, los organigramas de la empresa, convenios colectivos, descripciones de tareas y expedientes del personal.

De administración de bienes:

Son aquellos que tienen que ver con muebles e inmuebles, equipos y terrenos de una empresa; así como con las transacciones que sobre los bienes se realizan, tales como: compra, venta, mantenimiento, reparación o transformación. Además cubren el área de seguridad, vigilancia y prevención. Los más representativos son las licitaciones, ofertas e inventarios.

Documentos legales:

Estos documentos son exclusivamente de carácter jurídico y por ende implican movimiento jurídico de parte de la organización, los más importantes son: los contratos, los acuerdos y las reclamaciones.

1.4 Ciclo vital de los documentos

Para propósitos de la gestión documental se atribuye una edad al documento, basándose en tres períodos: activo, semiactivo e inactivo; con lo cual se conoce documentos activos, semiactivos e inactivos.

Documento activo

Es aquel de uso cotidiano, razón por la cual es recomendable conservarlo cerca de los usuarios.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Documento semiactivo

Se consulta ocasionalmente y por ende, su presencia física cerca de los usuarios no es necesaria.

Documento inactivo

Muy pocas veces es consultado. A menos que posea un valor histórico, legal o administrativo podrá ser destruido dentro de poco tiempo.

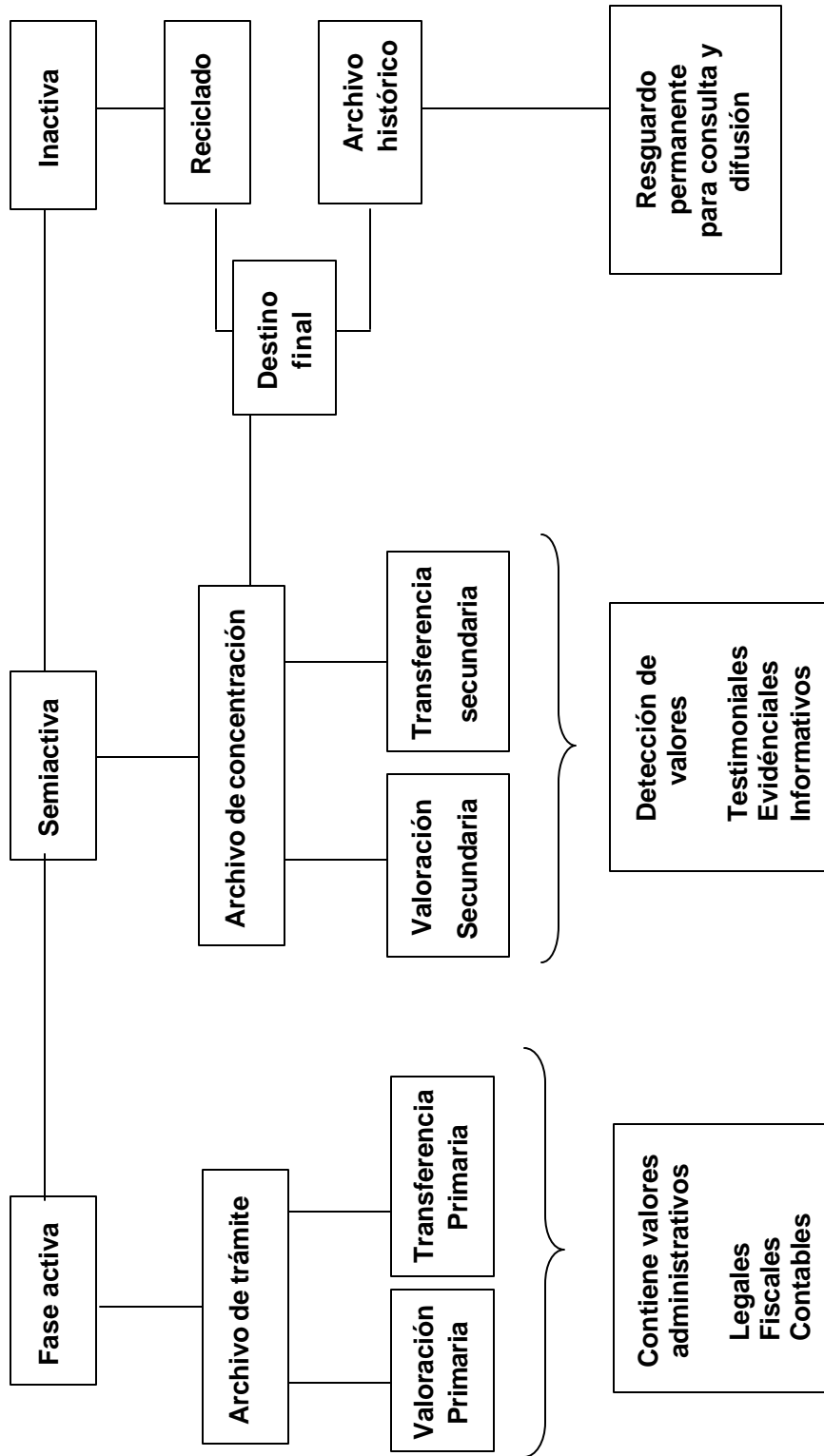


Gráfico 1-1 Ciclo Vital de los documentos

Fuente: Introducción al manejo de Documentos, Guylene Ethier, México 2000

1.5 Valor de los documentos

Con el fin de establecer su permanencia como archivo, todo documento sigue un proceso de valoración documental, por medio del cual puede adquirir un valor primario o secundario.

Valor Primario

Representa el valor informativo concreto e inmediato de todo documento que se genera como producto de un proceso, actividad o circuito de trabajo. A la vez que responde a un aspecto operacional y práctico, sirve como elemento de prueba y verificación.

Valor Secundario

Es el valor enfocado en la conservación y preservación del documento en el transcurso de su vida y que está encaminado a convertirse en antecedente e información de investigación. Es el que interesa a los investigadores de información retrospectiva. Surge una vez agotado el valor inmediato o primario. Los documentos que tienen este valor se conservan permanentemente.

Además es posible categorizar a los documentos según su origen en los siguientes valores: administrativo, legal o histórico. Todo documento que no encaje en ninguno de estos valores, eventualmente podrá ser destruido.

Valor administrativo

Un documento posee este valor si se utiliza para relatar, enmarcar y justificar una o más decisiones administrativas. Es importante recalcar, que sólo documentos oficiales adquieren este valor, aquellos documentos que sirvieron para la preparación del documento oficial no se les adjudica un valor administrativo.

Valor legal

Cuando un documento da un carácter oficial a un acto legal se le atribuye este valor. Tal es el caso de un juicio emitido en el caso de un reclamo que involucre a la empresa.

Valor histórico

Pocos son los documentos que adquieren este valor, básicamente son aquellos documentos que tienen que ver con la creación, organización, modificación o momentos importantes de una empresa.

SISTEMAS DE GESTIÓN DOCUMENTAL

Reciben esta denominación aquellas aplicaciones que en mayor o menor medida, incorporan las funciones típicas de las Ciencias de la Documentación a un entorno informático. Este tipo de sistemas se ha popularizado mucho en los entornos empresariales y su finalidad es la de establecer un control de la documentación recibida y generada por la organización, independientemente de recibir o no un tratamiento documental completo. Todo este proceso de mejora en el manejo de documentos dio lugar a la rama de la informática documental.

2.1 Arquitectura de los sistemas de gestión documental

La estructura básica de los sistemas de manejo de documentos se presenta en el siguiente esquema:

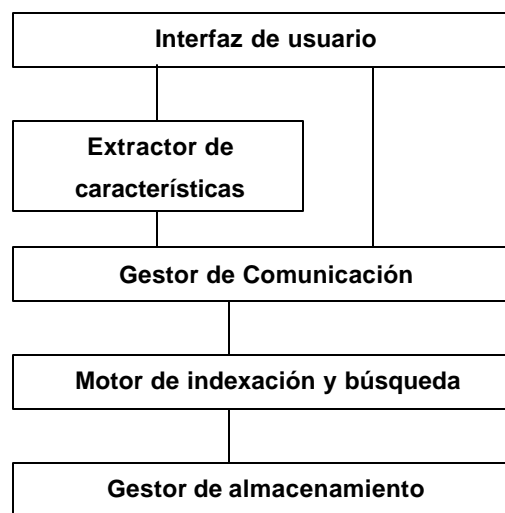


Gráfico 2-1 Arquitectura de sistemas informáticos de recuperación de documentos
Fuente: Multimedia Database Management, Guajun Lu, Artech House Boston -London

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Para explicarlo cabe mencionar que las operaciones (fases) más importantes en un sistema de gestión documental son:

- Inserción de un documento (diversos formatos: texto, imagen, audio, video)
- Búsqueda y recuperación de un documento.

Previamente a la inserción, el usuario obtiene el archivo digital del documento en mención. Luego, utiliza la *INTERFAZ DE USUARIO* para indicar al sistema el formato del archivo. El contenido y las características (propiedades del documento y su contenido) son ingresados de forma manual o automática con el *EXTRACTOR DE CARACTERISTICAS*. Estas propiedades y el archivo son enviados al servidor mediante los *GESTORES DE COMUNICACION*.

En el servidor, las propiedades del documento son organizadas (insertadas) de acuerdo a mecanismos de indexación que aseguran una eficiente recuperación del documento a través del *MOTOR DE BUSQUEDA E INDEXACION*. La información de indexación y el archivo original se almacenan gracias al *GESTOR DE ALMACENAMIENTO*.

Durante la búsqueda y recuperación de un documento, el usuario indica mediante parámetros de búsqueda el documento que desea consultar en la *INTERFAZ DE USUARIO*. También la *INTERFAZ DE USUARIO* le permite navegar entre los documentos almacenados. Con los parámetros de la consulta se obtiene las demás características del documento y se las envía al servidor mediante los *GESTORES DE COMUNICACION*. El *MOTOR DE BUSQUEDA*

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

E INDEXACION encuentra en la base de datos los ítems que mejor satisfagan las condiciones de búsqueda. Estos ítems son recuperados a través del GESTOR DE ALMACENAMIENTO y enviados al usuario mediante los GESTORES DE COMUNICACION. Finalmente, la INTERFAZ DE USUARIO despliega la lista de ítems obtenida de la búsqueda.

Otros modelos más complejos de software de manejo documental incluyen el GESTOR TESAURO, que se encarga de guardar sinónimos y relaciones entre la información que se ingresa al sistema, para así hacer la búsqueda más rápida.

2.2 Tipos de sistemas

Sistemas de gestión de bases de datos documentales: incorporan todas las características de los sistemas de gestión de base de datos tradicionales incluyendo creación y mantenimiento de bases de datos documentales, usuarios, controles de seguridad e incluso lenguajes propios de programación. Un software que utiliza este tipo de sistema es Inmagic.

Sistemas de indexación: crean ficheros de índice que permiten una limitada recuperación a través de ellos, utilizan operadores booleanos. Tal es el caso del software ZyIndex.

Sistemas de exploración o escáneres: además de las características de los sistemas de indexación, las aplicaciones que pertenecen a esta clasificación,

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

permiten buscar en el contenido de los documentos, de diferentes formatos, la cadena de caracteres que se expresa en la ecuación de búsqueda. Las *ecuaciones de búsqueda* constituyen una serie de términos de interés que utilizan un conjunto de operadores que expresan la relación que existe entre dichos términos.⁴

DtSearch es una aplicación de este tipo.

Sistemas de gestión bibliográfica: sistemas exclusivos para el tratamiento de bibliografías que permite almacenamiento, recuperación de referencias bibliográficas y exportación a diferentes procesadores de textos. Un software que pertenece a este tipo de sistemas es ProCite.

Sistemas de recuperación de información: son exclusivamente para recuperar información de bases de datos documentales no modificables. Aunque posee poderosos mecanismos de búsqueda su funcionalidad queda reducida a la consulta y exportación de documentos. Pertenecen a esta clasificación las siguientes aplicaciones: SPIRS (Silver Platter Information Retrieval System) y DIALOG.

Sistemas hipertextuales: mediante hipertextos e hipermedias es posible organizar, acceder y explorar documentos de diversos tipos. Los motores, tutoriales y presentaciones pertenecen a este tipo de sistemas.

⁴ Consultas al STRID, Las Ecuaciones de Búsqueda, www.tramullas.com/documatic/4-6.html

Sistemas de Gestión Documental o de Gestión Electrónica de Documentos:

aplicaciones que incluyen características de gestión de bases de datos y esquemas de obtención de una imagen (copia del documento original – escáner) más una descripción textual del documento y de su contenido. Tal es el caso del software Panagon.

Sistemas o Gestores de información personal: integran los documentos de diferentes tipos que son de interés para un usuario en su trabajo diario.

Sistemas compuestos: son los más completos ya que cubren la cadena documental como la gestión administrativa y por ende soportan actividades que se presentan en un archivo, biblioteca o centro de documentación. Por lo general está formado por un motor documental, encargado de manejar las bases de datos documentales que cubren los catálogos; y un motor relacional, que cubre las tareas administrativas. Las aplicaciones Absys y Libertas son sistemas documentales de este tipo.

2.3 Características de los sistemas estudiados

Los sistemas de gestión documental estudiados fueron los siguientes:

- Panagon de Filenet
- Weblink de LaserFiche
- AlchemyWeb de IMR

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Panagon

La empresa FileNET fundada en 1982, al discontinuar su producto Watermark, para reemplazarlo, ingresa al mercado con el software Panagon Web Content Manager para manejo de documentos. El mismo que está disponible en entorno web para Internet, Intranet o Extranet y ofrece flexibilidad y escalabilidad a empresas en constante crecimiento. Este sistema cuenta con un manejador de contenido donde a través de una base de datos relacional (Oracle o SQL Server) se almacenan los documentos. La interfaz de usuario está disponible gracias al servicio IIS (Internet Information Services) que reside en el servidor web.

Como tareas administrativas están las siguientes:

- Manejo de usuarios, dependiendo de la infraestructura de red en la que se encuentra puede ser configurado para utilizar el mismo usuario de red, evitando así una autenticación adicional del usuario.
- Manejo de “clases documentales” y “propiedades propias”, que consisten en verdaderos ficheros de los tipos de documentos que se pueden ingresar, como es el caso de una clase documental Contrato que podrá ser definida con las propiedades de: Número, Descripción, Fecha de Expiración, etc.
- Definición de plantillas de búsqueda que puedan ser usadas vía web, según la “clase documental”.

Para el proceso de ingreso de documentos permite asociar al documento a una “clase documental” General u otra definida por el usuario, y una vez digitados

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

los índices, el archivo sube al servidor en la carpeta elegida. Además si es un documento Word, Excel, Power Point, permite indexar su contenido.

Para el proceso de búsqueda de un documento es posible usar las plantillas de búsqueda que permiten buscar por las “propiedades” de las “clases documentales”, o también por contenido. Es posible el uso de operadores booleanos.

El documento puede ser abierto en un visor propio del sistema, o en la aplicación nativa. Además las operaciones de mover o copiar documentos entre directorios son permitidas.

WebLink

Desde 1987, la empresa LaserFiche se ha especializado en gestión de documentos y lanzó al mercado el software WebLink para el manejo de documentos, con acceso para entornos web vía Intranet o Internet. Este sistema al igual que el anterior permite definir “plantillas” para documentos con campos descriptores fijos como el caso de Facturas con los campos de: Número, Comprador, Ciudad, Fecha, Código de producto, etc.

Una interfaz bastante simple y que cuenta con las siguientes operaciones básicas:

- Navegar: para buscar documentos de directorio en directorio.
- Buscar: para búsquedas simples o con campos descriptores.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

- Anotaciones: es posible colocar notas electrónicas a un documento dentro de WebLink.

Al desplegar el documento permite obtener una copia en formato PDF del documento, bajar a disco local una copia del mismo, o imprimir. Además de las opciones de Zoom In-Out.

Los tipos de búsqueda permitidos son búsquedas simples, avanzadas, por aproximación (fuzzy logic), booleana y por los campos descriptores; o por una combinación de todos. También permite buscar por nombre de archivo, nombre de carpeta o cualquier otra palabra en la página.

Es usado en empresas de diversos tipos como CIA, FBI, La Armada de los Estados Unidos, Fuerza Aérea, etc.

Alchemy Web

Este sistema de la empresa IMR (Information Management Research) disponible para Intranets y Extranets, cuenta con tecnología XML. A diferencia de los anteriores está disponible en el Internet. Mediante una subscripción es posible ingresar documentos a él, cada archivo ingresado tiene una tarifa fija.

Para el ingreso de un archivo solo es necesario proveer el título del documento y un resumen, que servirán posteriormente para su búsqueda y recuperación. Permite búsquedas simples y avanzadas usando operadores booleanos.

2.4 Funciones básicas

En general un sistema de gestión documental debe englobar tareas de digitalización, indexación y recuperación de documentos en diferentes formatos electrónicos para facilitar la administración del documento desde que se recibe o se genera.

Además de ello sus funciones son:

- Permite consultar por diversos criterios documentos de varios tipos: contratos, documentos financieros, administrativos, etc.
- Acceso al documento íntegro en formato digital.
- Simplificar la organización electrónica de documentos.

2.5 Ventajas y problemas que se solucionan

Las ventajas de tener un sistema de gestión documental se resumen en el ahorro de tres recursos importantes en una empresa:

- Espacio
- Tiempo
- Papel (dinero)

Ahorro de espacio.- está determinado que la información contenida en 50 cajas de tamaño normal con 400 páginas dentro de ellas, cabe en un disco de 1Gb. Esto

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

hace que el espacio de todo un sistema de estanterías sea reemplazado por un disco que ocupa no más de 0.5m² de superficie.

Problemas solucionados:

- ❖ Superficie útil desaprovechada, suciedad, deterioro, etc.
- ❖ Seguridad física, incendio, humedad.

Ahorro de tiempo.- los procesos de almacenamiento y recuperación de un documento pueden ser sorprendentemente disminuidos con un sistema electrónico de manejo de documentos, ya que permite buscar por nombres, o cualquier clave indexada, visualizar en pantalla, enviarlo por fax o e-mail en cuestión de segundos.

Problemas solucionados:

- ❖ Tiempo de archivado
- ❖ Tiempo de recuperación
- ❖ Imposibilidad de clasificación por múltiples campos.

Ahorro de papel.- con este tipo de sistemas se evita enviar copias de documentos.

Problemas solucionados:

- ❖ Duplicidad de documentos.
- ❖ Falta de confidencialidad.
- ❖ Dificultad de acceso a la información desde sucursales de la empresa o sitios remotos.

2.6 Sistemas de bases de datos relacionales vs documentales

	<i>RDBMS (Sistemas gestores de Bases de Datos Relacionales)</i>	<i>DDBMS (Sistemas gestores de Bases de Datos Documentales)</i>
<i>Entidad</i>	Tabla	Documento
<i>Estructura</i>	Registro y campos de longitud definida	Registros y campos de longitud indefinida.
	Campos repetitivos. Diferentes tablas de datos enlazadas entre sí.	Campos clave únicos. Por lo general una tabla que en cada registro contiene una unidad diferente y en cada campo de este registro algún rasgo que define la información contenida en él.
	Búsqueda restringida a uno o más campos.	Es posible buscar por todas y cada una de las palabras de los campos.
<i>Utilidad</i>	Procesos que trabajan con datos sueltos	Cuando la información a recoger es en cierto sentido demasiado indefinida.
	Sistemas contables	Sistema de fichas de libros.

Tabla 2-1 Cuadro comparativo Base de Datos Relacionales vs Documentales

Elaboración: María Augusta Mateus

Fuente: Base de SGBDD-Mini-COMO Versión Preliminar, Ignacio de Miguel Diaz

<http://imigueldiaz.eresmas.com/SGBDD-Mini-COMO.htm>

Una base de datos documental es una colección de documentos relacionados reunidos en una sola unidad. Los documentos individuales pueden ser enormes o diminutos, pero estos pueden soportar relación entre ellos.

Una base de datos relacional está formada de unidades más pequeñas llamadas registros. Por otro lado, en una base de datos documental, un registro puede ser un documento entero, una sección dentro de un documento, una sola página, o un fragmento de texto dentro de una página.

Cuando se realiza una búsqueda en una base de datos relacional, se puede obtener uno o más registros que contengan la información buscada.

Un registro puede contener regiones más pequeñas de datos llamada Campos. Un campo normalmente define un tipo particular de dato común a varios o a todos los registros dentro de una base de datos. Por ejemplo, en una base de datos de memos en una empresa, un memo conforma un registro, y se tienen campos como: Para, De, Fecha, Asunto, y el Texto. El alcance de una búsqueda puede ser restringido a uno o más campos.

2.7 Aplicaciones de los sistemas de manejo de documentos

Este tipo de software tiene una gran variedad de campos de aplicación, por ejemplo citaremos:

- **Medicina:** el doctor que toma una imagen ultrasonido y necesita obtener todas las imágenes con un grado similar de hipertrofia ventricular izquierda al patrón que posee.
- **Seguridad:** el policía que posee una fotografía de un individuo y desea obtener todas las imágenes similares a él que exista en una base de datos que guarda información criminalística.
- **Educación:** un estudiante que escanea una imagen de un animal y desea obtener toda la información referente a él como (sonidos, imágenes, texto) de una base de datos académica. Otra alternativa de búsqueda es que el estudiante con el sonido del animal, recupere de la base de datos imágenes e información del animal.

- Periodismo: un reportero que escribe el artículo de una persona necesita obtener todas las imágenes y la información de periódicos de los últimos 20 años.
- Entretenimiento: un espectador desea recuperar los video clips que sea similares al que está viendo.

2.8 Nuevas tendencias tecnológicas – Tratamiento Audiovisual de la Información

Por definición, los documentos audiovisuales son aquellos que registran información acústica y óptica. Los centros de documentación audiovisuales son recientes, y están enfocados principalmente a medios de comunicación o empresas afines. Su objetivo principal es lograr conservación y acrecentamiento del material audiovisual con valor histórico, documental, informativo, cultural y/o recreativo; ya que forma parte del patrimonio de la empresa o bien, para reutilizarlos con fines comerciales o de reproducción.

Además el procesamiento digital de videos obtiene un interés especial debido a que es la base para muchas aplicaciones importantes como la televisión digital, Internet, la multimedia, el sonido digital y en general el entretenimiento.

Los materiales básicos para fijar información audiovisual son: película film, material magnético y material plástico (discos ópticos). La calidad del material audiovisual no solo depende del proceso de digitalización que se utilice,

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

sino también de las óptimas condiciones en que se realizó la grabación que se trata de digitalizar.

Entre los centros especializados en el tratamiento de este tipo de información están el Institut Nationale de l'Audiovisuel (INA) de Francia y el Canal Sur Radio y Televisión de España. Para estas empresas su meta es contar con grandes salas multimedia que reproduzcan documentos como dibujos originales en color, material cartográfico, grabados, minutos de sonido; todos ellos almacenados en un solo lugar y fáciles de acceder. Por los esfuerzos que se están realizando, se pronostica que el futuro está en los audiovisuales pues a través de ellos es posible transmitir de forma exacta y completa cualquier información.

Un ejemplo es la empresa Coca Cola que utiliza un sistema de gestión de archivos de texto, imágenes fijas y videos. Con el propósito de poner a disposición de sus empleados en todo el mundo, un siglo entero de campañas de marketing y publicidad de entre decenas de miles de video clips, fotografías y melodías de las más famosas campañas publicitarias de la compañía para ser utilizados para futuros proyectos.

PRESENTACIÓN Y CONSERVACIÓN DE LA INFORMACIÓN

La información puede llegar a nosotros en varias presentaciones, en forma de texto, imágenes, videos, animaciones, sonidos, etc. Para poder manipularlos con una computadora es necesario digitalizarlos.

El concepto de digitalización se aplica para todo aquello que puede ser representado por ceros y unos en una computadora. Textos electrónicos, imágenes, audio y video son susceptibles de digitalización, en formatos, codificaciones y representaciones diferentes. Se toma cada fuente y se la convierte a formato digital, almacenándolos en archivos que se distinguen unos de otros mediante el empleo de extensiones (doc, txt, jpg, gif, tif, wav, etc.) de acuerdo a su naturaleza.

3.1 Texto

Los textos digitales pueden tener diferentes orígenes: cuando parten de la creación y evolución totalmente digital en la cual no existe ningún medio impreso, se genera directamente en algún dispositivo que permite almacenarlo en un medio electrónico. Por otro lado, cuando el texto ha surgido del procesamiento de

documentos impresos en papel, entonces es escaneado y luego convertidos a formato digital mediante la técnica de OCR.

3.1.1 Proceso de digitalización

El proceso de OCR (Optical Carácter Recognition) permite transformar textos en papel a imágenes (con el uso de un escáner), y luego ser interpretados y convertidos a texto digital; para almacenarlos en los formatos de texto comunes (.doc, .txt, .RTF o plano).

El primer programa de OCR fue presentado en 1959 por Intelligent Machine Corporation y era capaz de leer apenas un tipo de letra de tamaño fijo. Luego, se diseñaron programas capaces de leer diez o más tipos de fuentes mediante técnicas de coincidencia de patrones, en las que se comparaba la imagen obtenida, con una “biblioteca” de imágenes almacenadas.

Para finales de los 80 se desarrollan sistemas de redes neuronales, capaces de aprender a partir de ejemplos; para lo cual se alimentó al sistema con más de 10.000 ejemplos para cada caracter en inglés, extraídos de todo tipo de material real, incluso de páginas dañadas. Esto dio como resultado un sistema de OCR bastante inteligente y eficaz.

Hoy en día, se usan algoritmos expertos, diseñados específicamente para atender diferentes campos del reconocimiento de caracteres, así por ejemplo, unos

se especializan en estilos de fuentes, otros en diversidad de vocabulario, hay otros que identifican daños en las hojas, etc.

En el mercado existen diferentes tipos de software dedicados a este proceso, uno de ellos y sobre el cual realicé pruebas es el OmniPage Pro 11, una aplicación bastante precisa y completa ya que permite dividir al documento en secciones para digitalizar por separado texto, columnas, tablas y gráficos. Además incluye un método de corrección contextual donde términos dudosos se resaltan para facilitar la corrección.

3.1.2 Formatos de almacenamiento

Una vez realizado el proceso necesario para obtener el texto digital, a éste se lo convierte en SGML (Standard General Markup Language), el lenguaje estándar para la representación del texto por sus bondades y características para la estructuración de texto. Otros formatos usados son: ODA (Office Document Architecture), LaTeX y PDF (Portable Document Format).

- **SGML:** permite la descripción de información estructurada que es independiente de como la información es procesada.
- **ODA:** especifica cómo los componentes lógicos y de diseño de un documento se componen e interrelacionan.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

- **PDF:** basado en lenguaje Postscript, PDF es designado para representar la apariencia de impresión, incluye estructura lógica y el texto puede incluir imágenes y gráficos.
- **Hypermedia (HyTime):** es un estándar ISO para la representación de documentos hypermedia usando SGML. Este estándar especifica la estructura del documento y la asociación entre los elementos del documento.
- **MHEG:** un estándar ISO en modelo orientado a objetos para el intercambio entre sistemas multimedia. Especifica el formato del documento usado para intercambiar información entre sistemas independientes.

3.1.3 *Ventajas y desventajas*

La principal ventaja es que gracias a los programas de reconocimiento óptico de caracteres (OCR) es posible convertir imágenes escaneadas a texto y luego editarlas en cualquier procesador de texto sin tener que mecanografiarlo. Esto se traduce en un ahorro de tiempo y esfuerzo.

Aunque para tipografías sencillas este resulta ser un proceso útil, para documentos manuscritos antiguos o de baja calidad, el OCR no es una buena opción ya que el proceso tiende a confundir letras y el tiempo que puede llevar en la corrección del mismo puede ser mayor que la captura manual del documento.

3.2 Imágenes

3.2.1 Proceso de digitalización

Cuando el texto no es posible convertirlo a formato electrónico, deberá permanecer como imagen. Digitalizar imágenes es simplemente convertir los formatos análogos como fotografías o dibujos a un formato que el computador pueda manipular. Para lograrlo se utiliza un dispositivo llamado escáner o una cámara digital.

3.2.1.1 Cómo funciona un escáner

Un escáner es un aparato para digitalizar imágenes, es decir que toma una imagen (fotografía, dibujo o texto) y la convierte a un formato que se pueda modificar y almacenar en un ordenador.

El escaneado de una imagen se realiza con el barrido del documento por una fuente luminosa, las zonas claras reflejan más luz que las partes oscuras. Luego mediante espejos se conduce la luz reflejada hacia un dispositivo fotosensor denominado CCD (Charge Coupled Device) que recibe la luz y la transforma en señales eléctricas controladas por la intensidad y el color de la imagen, del mismo modo que lo hace un ojo.

Las señales eléctricas son recibidas por un convertidor analógico-digital llamado DAC (Digital Analog Converter) que las convierte en bits, que luego son transmitidos al ordenador.

3.2.1.2 Resolución de un escáner

Un escáner tiene tres diferentes tipos de resolución:

- 1. Resolución óptica o real**, que es el número de puntos individuales de una imagen que es capaz de captar el CCD. Esta es la resolución más importante porque define los límites físicos del escáner. Se expresa dando los dpi (puntos por pulgada) horizontales x dpi verticales.
- 2. Resolución interpolada** está dada por puntos creados por la computadora, que acomoda entre los puntos captados en la resolución real. Estos puntos deben sus características a los puntos reales que tengan al lado. Esto aumenta mucho la resolución, pero siempre depende de la resolución óptica o real.
- 3. Resolución de escaneado**, que es definida por el usuario. Puede ir desde un mínimo de 75 dpi (puntos por pulgada) hasta el máximo que pueda alcanzar el escáner.

3.2.1.3 Tipos de escáner

A continuación, los tipos más conocidos de escáner:

Tipos de escáner	Funcionamiento	Características	Utilidad
<p>Escáner Plano</p> <p>Escáner con Alimentador de Hojas</p>	<p>La fuente de luz y el sensor CCD se encuentran acoplados en un brazo móvil que se desliza sobre el documento. El documento se encuentra inmóvil sobre una placa de vidrio.</p> <p>El sensor CCD y la fuente de luz permanecen fijos, mientras lo que se mueve es el documento por efecto de un mecanismo de rodillos, cinta, tambor o de vacío.</p>	<ul style="list-style-type: none"> - El más conocido y vendido. - Dimensiones de área de lectura: 22 x 28 cm. - Resolución real: 300 x 400 dpi y con interpolación de hasta 1600 dpi. - Diseñados para documentos de tamaño uniforme y con solidez suficiente para soportar manipulación brusca. - Existen en el mercado escáner blanco y negro, y a color. - Casi generalmente presentan resoluciones relativamente bajas. 	<ul style="list-style-type: none"> - Util para escanear: fotografías, hojas sueltas, periódicos, libros encuadernados. - Es posible digitalizar transparencias, diapositivas o negativos con los adaptadores adecuados. - Utilizados en lugares donde se requiere escanear grandes volúmenes de hojas.
<p>Escáner de tambor</p>	<p>En lugar de utilizar el sensor CCD, se utiliza un sistema de tubos fotomultiplicadores (PMT) en el bloque lector.</p> <p>Un sistema de transmisión fotomecánica recorre la imagen punto por punto, obteniendo una gran resolución y gama dinámica entre bajas y altas luces.</p>	<ul style="list-style-type: none"> - Consiguen la mayor resolución conocida (hasta 4000 dpi). - Produce imágenes en colores primarios, que se convierte a CMYK cuando el lector recorre la imagen. - Son escáner lentos. - Requiere una gran habilidad del operador. 	<ul style="list-style-type: none"> - No aptos para documentos de papel quebradizo. - Útiles para documentos que requieren altas resoluciones como mapas, imágenes satelitales, planos.
<p>Escáner para transparencias</p>	<p>Digitalizan varios formatos de película transparente ya sea, negativa, positiva, color o blanco y negro.</p>	<ul style="list-style-type: none"> - Tamaño de escaneado: desde 35mm hasta placas de 9x12cm. - Si se conserva el tamaño original, la imagen tiene una buena resolución. - Escaneado lento. 	<ul style="list-style-type: none"> - Se usa para escanear formatos de película transparente.
<p>Escáner de mano o portátil.</p>	<ul style="list-style-type: none"> - Carecen de un motor para pasar las hojas, la persona debe deslizar el escáner sobre el original. 	<ul style="list-style-type: none"> - Escáner pequeño, de bajo precio y buena resolución. - Son lentos y pocos modelos manejan colores. 	<ul style="list-style-type: none"> - Util para digitalizar: libros encuadernados, artículos periodísticos, facturas e imágenes pequeñas de todo tipo. - Limitado a documentos de cierto ancho (no mayor a 10cm) y sin importar sulargo.

Tabla 3-1 Tipos de Escáner

Elaboración: María Augusta Mateus

Fuente: El escáner y su uso en la digitalización, <http://biblioweb.dgsca.unam.mx/bd/dig/escaner.html>

3.2.2 Parámetros para la digitalización de imágenes

El proceso de digitalización de imágenes depende de tres directrices que hay que analizar:

- ❖ El espacio necesario (y disponible) para su almacenamiento.
- ❖ La calidad de la imagen.
- ❖ La agilidad en la recuperación de la imagen.

Basados en lo indicado anteriormente los parámetros para la digitalización de imágenes son:

1. Número de bits por pixel
2. Resolución de la imagen

Número de bits por pixel

La siguiente tabla, muestra los más comunes tipos de bits por pixel y la imagen que puede obtenerse.

Bits por pixel	Descripción	Resultado de la Imagen
1-bit blanco y negro	2 colores (blanco ó negro) por cada pixel.	- Calidad visual deficiente. - Imagen adecuada para dibujos e imágenes para Web.
24-bit color	16.8 millones de colores posibles por cada punto.	- Mayor calidad. - Imagen realista y perfecta para fotografías. - Archivos de gran tamaño.

Tabla 3-2 Bits por pixel

Elaboración: María Augusta Mateus

Fuente: El escáner y su uso en la digitalización,

<http://biblioweb.dgsca.unam.mx/bd/dig/escaner.html>

Resolución de la imagen

La resolución es el número de píxeles por unidad lineal, está medida en dpi (dots per inch o puntos por pulgada). Mientras más grande el número de dpi, mayor es la información del archivo y el detalle de la imagen mejora.

Las resoluciones que se usan más comúnmente son:

- ❖ 100 dpi: Imágenes pensadas para desplegar en pantalla.
- ❖ 300 dpi: Imágenes pensadas para impresiones, ampliaciones en pantalla o documentos valiosos (manuscritos o ejemplares únicos).

La selección de una resolución adecuada influye directamente al tamaño del archivo, así a mayor número de dpis, mayor tamaño de la imagen pero también mejor calidad. Por ende, según el uso del documento: consulta, difusión o preservación, se debe hacer una valoración entre la calidad del documento y la velocidad de recuperación del mismo.

3.2.3 Formatos de almacenamiento

Los formatos en los que se puede guardar una imagen van desde formatos sin compresión como TIFF, BMP o PICT hasta otros con compresión como GIF y JPEG.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

La compresión hace una selección y reorganización de los datos almacenados de una imagen no comprimida, reduciendo así el tamaño del archivo pero la imagen mantiene su apariencia.

Los formatos más comunes (con y sin compresión) son:

- **TIFF:** es el formato estándar sin compresión más utilizado. Su propósito es describir y almacenar datos de imágenes escaneadas. La riqueza de los datos es requerida para apreciar las propiedades de los escáneres y otros dispositivos similares. Permite obtener imágenes blanco y negro, en escala de grises, paleta de colores y full-color.
- **GIF (Graphics Interchange Format):** ofrece una compresión moderada en 8-bit a color o grises. Fue muy utilizado hasta 1995.
- **JPEG (Joint Photographic Experts Group):** ofrece una excelente compresión de imágenes 8-bit o 24-bit color. Proviene de un grupo de estándares de técnicas conocidas como compresión con pérdidas (lossy compression).

3.2.4 Ventajas y desventajas

Combinando los factores de: profundidad del color, resolución de la imagen y tipo de formato (compresión) se puede obtener diversidad de tamaños y calidad de la imagen, la misma que debe estar de acuerdo al uso que se hará de ella.

El tamaño de la imagen está dado por: $L \times A \times RH \times RV \times \text{bits}$. Donde:

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

L y A son las dimensiones de la imagen.

RH y RV son las resoluciones vertical y horizontal.

Bits: el número de bits dependiendo del color.

Así, el siguiente cuadro muestra la gama de posibilidades que se puede obtener al combinar los factores antes mencionados:

Proceso	Ventajas	Desventajas	Aplicabilidad
Escanear a menor resolución.	Tamaño del archivo en disco y memoria es pequeño.	Calidad menor	Si la imagen sólo se destinará para desplegarla en pantalla.
Reducir el color.	Tamaño del archivo en disco y memoria es pequeño.	Imágenes a color pierden su valor.	Ideal para imágenes en blanco y negro sin muchos matices.
Archivar la imagen en formato comprimido.	Tamaño de archivo en disco baja notablemente. Imágenes de calidad.	Tamaño del archivo en memoria no disminuye.	Excelente para Imágenes de grandes dimensiones y/o a color

Tabla 3-3 Ventajas y desventajas en la digitalización de imágenes

Elaboración: María Augusta Mateus

Fuente: El escáner y su uso en la digitalización,

<http://biblioweb.dgsca.unam.mx/bd/dig/escaner.html>

3.3 *Audio*

3.3.1 *Proceso de digitalización*

Los documentos de audio registran información de sonido en forma digital, es decir, representada por ceros y unos. El proceso de digitalizar al sonido es tomar muestras periódicas de la fuente analógica (sonidos), a las muestras se le asigna valores numéricos que son convertidos a ceros y unos.

3.3.1.1 *Parámetros para la digitalización de audio*

- **Frecuencia de muestreo:** El número de muestras que se recogen por segundo en Hertzios. La calidad del sonido será mayor cuando más muestras se recojan, pero también aumenta el espacio que ocupa en bits. La frecuencia de muestreo limita la frecuencia máxima que se puede recoger de la señal y por ello debe ser algo más del doble de la frecuencia máxima que se registrará.

Soporte	Frecuencia de grabación
DAT	48 kHz
CD-Audio	44.1 kHz
Trackers tradicionales	8.363 kHz
Telefonía	8 kHz

Cuadro 3-1: Frecuencias de grabación más típicas.

Fuente: Coursillo de introducción a la informática musical,
<http://www.salleurl.edu/~is04069/Modders/Docs/sampledoc.html>

- **Bits de profundidad:** Son el número de bits con los que se codifica una muestra. Pueden ser de 8 bits que permite representar $2^8=256$ niveles de amplitud, o de 16 bits que es capaz de representar $2^{16}=65536$ niveles de amplitud.

3.3.1.2 *Soportes digitales de audio*

El disco duro es un soporte digital, pero a más de él están otros soportes portátiles que transportan fácilmente los samples y son:

- A-DAT
- Compact Disc (CD)

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

- Mini Disc
- Digital Compact Cassette

Soporte Digital	Características	Utilidad	Problemas
A-DAT	Cinta de video VHS. Graba 8 pistas a 24 bits y frecuencias de muestreo de 32, 44.1 y 48 kHz.	Se usa mucha en estudios profesionales y caseros por el coste bajo.	Desgaste por el contacto de una cabeza magnética con la cinta. Acceso secuencial. Cintas de video muy voluminosas
CD	Utiliza lectura óptica, que es la reflexión de un láser en el soporte. Almacena 80 min de audio digital a 16 bits, muestreado a 44.1 kHz. Futuro del CD está en el DVD que permite grabar 30 horas de música en él en formato WAV.	Útiles para copias de seguridad y para distribuir maquetas. Pueden convivir audio y datos en el mismo CD.	No se presenta desgaste. Soporte tolerante a ralladas, sobretodo en sentido radial.
Mini Disc	Surgió con la idea de reemplazar al Music Cassete con un soporte digital. Tecnología termo magneto óptica de grabación.	Para grabaciones en sitios que pueden tener vibraciones.	No ofrece tanta calidad como el CD pero permite grabación.
Digital Compact Cassette	Mismo formato que una cinta normal de Music Cassete pero que graba digitalmente. Graba 8 pistas de audio más un canal auxiliar.	Soporte totalmente compatible con Music Cassete.	Soporte sensible a daños. Acceso secuencial.

Tabla 3-4 Soportes digitales de audio.

Elaboración: María Augusta Mateus

Fuente: Fonotrón: Restauración digital del sonido,

<http://www.arrakis.es/~fonotron/dsr.html>

3.3.2 Formatos de almacenamiento

Todo archivo de audio está formado por: cabecera, bits auxiliares y los bytes que representan la onda sonora. Existen infinidad de formatos de archivos de audio, de ellos cabe destacar dos: WAVE RIFF (.WAV) usado en Windows y

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

el AIFF usado en Macintosh. En definitiva, lo que caracteriza a un formato de otro es la manera en la que la cabecera almacena datos útiles como longitud, frecuencia de muestreo, punto de referencia y sobretodo codificación. A un archivo sin cabecera, ni bytes auxiliares se lo llama formato RAW, crudo o sin cabecera. El formato indica cómo se interpretarán los bits de cabecera y la codificación indica como se interpretará la tira de bytes restante.

El formato WAV, guarda documentos con resolución de 8 y 16 bits y a frecuencias de muestreo de 11.025, 22.050, 44.1 y 48 Khz. Los niveles de calidad del sonido, están determinados por el tamaño y la frecuencia de las muestra, así como del software y hardware utilizado para la grabación.

3.3.3 Ventajas y desventajas

Tipo de Resolución	Valor Digital del sonido	Ventajas	Desventajas
8 bits	0 a 255	Menor tamaño del archivo	- Mala calidad del sonido - Albergan ruidos fácilmente distinguidos por el oído humano.
16 bits	0 a 65,536	Grabación de mayor definición y claridad	- Tamaño del archivo aumenta en relación 1 a 2 con respecto a los de 8 bits. - Requiere equipos con mayor memoria y espacio en disco.

Tabla 3-5 Tipos de resolución de audio.

Elaboración: María Augusta Mateus

Fuente: Fonotrón: Restauración digital del sonido,
<http://www.arrakis.es/~fonotron/dsr.html>

3.4 Video

3.4.1 Proceso de digitalización

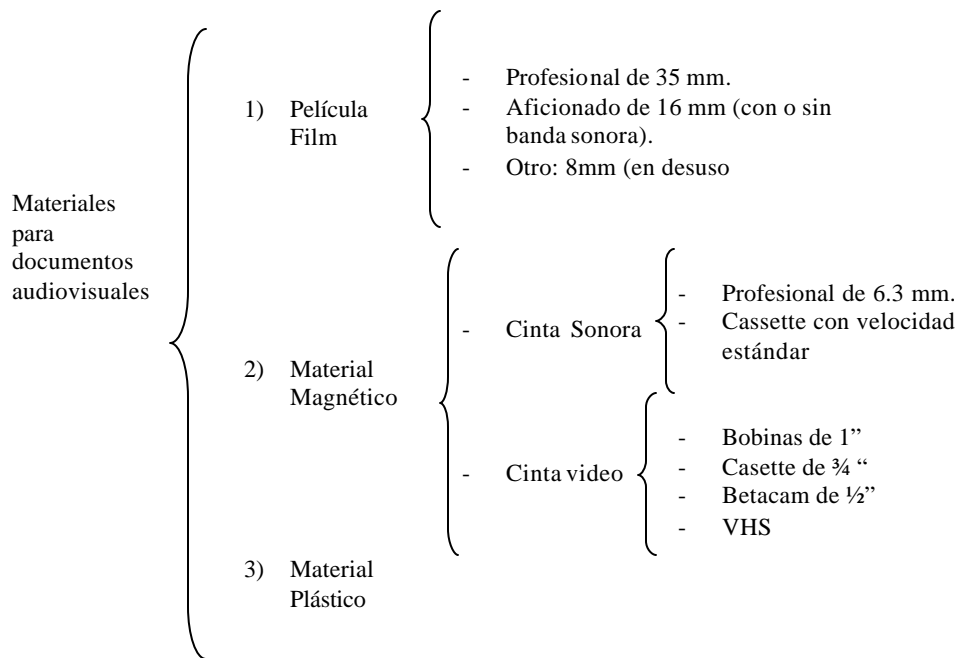
Los documentos de video registran información en imagen y sonido simultáneamente en formato digital.

La digitalización de una señal de video tiene lugar en tres pasos:

1. Muestreo
2. Cuantificación
3. Codificación

Los materiales específicos para la documentación audiovisual son: película film, material magnético y material plástico.

- La película film es una película fotográfica que tiene una capa de poliéster con una emulsión sobre la cinta que actúa por reacción química a la luz y que se fija por el revelado.
- El material magnético tiene una base de poliéster y una capa de óxido de hierro.
- Los materiales plásticos, cubiertos con una capa de aluminio reflectante. La información se fija en forma de muescas en forma de espiral. La lectura de información se hace por la reflexión de un rayo láser contra una superficie.



Cuadro 3 -2 Formatos de presentación de los principales materiales para documentos audiovisuales.
Elaboración: María Augusta Mateus

3.4.2 Formatos de almacenamiento

El video digital es una secuencia de imágenes y audio que son almacenadas y reproducidas en forma digital. Para una buena calidad en el video se necesita un método eficiente de compresión y una línea rápida para la transferencia.

3.4.2.1 Métodos de compresión

Cada cuadro de video tiene información redundante, un adecuado método de compresión sólo considera las diferencias entre un cuadro y otro. MPEG

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

(Moving Picture Experts Group) es un estándar internacional para la compresión de video digital y se basa en buscar la diferencia entre una imagen y otra y registrar sólo los cambios producidos. Por este motivo su compresión puede llegar a ser más alta que la de JPEG de imágenes sin movimiento. Pero MPEG I hace que el acceso a imágenes inmóviles individuales de una serie, sea difícil, ya que sólo el primer cuadro es grabado completamente.

3.4.3 Ventajas y desventajas

Según la Coordinación de Prospección e Innovación Tecnológica de la Dirección General de Servicios de Cómputo Académico, un video de 1 minuto de duración, con una resolución de 160 x 121 dpi y 29.98 cuadros por segundo, con cuadros llave cada 16 cuadros, ocupa un tamaño de 5 Mb. Si a este video le damos características de más alta calidad, el tamaño del archivo puede llegar a pesar 20 Mb. Es por ello que incluir documentos digitales en un sistema puede demandar muchos recursos como equipos poderosos y ancho de banda suficiente para que la transferencia de archivos de estas dimensiones sea rápida.

ANÁLISIS Y DISEÑO DEL SISTEMA

4.1 Análisis

Las empresas deben manejar a diario la documentación recibida o generada dentro de ellas. A más de ello, deben tener la capacidad de encontrar el documento que contenga la información adecuada en el menor tiempo posible y sin importar el lugar físico donde se encuentre. Como vimos en el capítulo 2, los Sistemas de Gestión Documental ayudan al mejoramiento de estos procesos, ya que permiten:

- Consultar documentos de varios tipos.
- Acceder al documento íntegro en formato digital, y
- Simplificar la organización electrónica de documentos.

Muchos sistemas existentes en el mercado se ajustan a las necesidades descritas, e incluso proporcionan nuevas funcionalidades tales como la búsqueda por contenido, sean estas imágenes, texto, video o audio, o replicación de documentos ingresados en distintos sitios, etc.

En el presente trabajo propondremos los siguientes módulos:

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

- **Manejador de Documentos**, que se encargará del almacenamiento en formato digital y la indexación de los datos necesarios para su búsqueda. Así, la acumulación de papel y otros medios disminuirá de gran forma.
- **Interfaz de búsqueda de documentos**, que permitirá encontrar y ver los documentos más relevantes en el menor tiempo posible, y en cualquier lugar; y,
- **Módulo administrativo**, que se creará con el fin de mantener la confidencialidad de los documentos y en general, administrar información relacionada a ellos.

4.1.1 Requerimientos

En la actualidad, los profesionales archiveros (o quien cumpla esta función dentro de la empresa) se encargan de clasificar y archivar los documentos físicos (en papel o cualquier otro medio) que son importantes para una empresa. Debido a que este proceso es manual, muchos documentos se traspapelan y otros se deterioran hasta volverse ilegibles.

Por ello es necesario mejorar los procesos de organización y almacenamiento de los documentos importantes de una empresa, poniendo al servicio del profesional archivero las herramientas informáticas para conseguirlo y asegurar así una rápida y eficaz recuperación del documento por cualquier miembro de la empresa.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Por otro lado, el proceso de búsqueda deberá mostrar los documentos que más se aproximen a la petición del usuario, la cual se basa en el ingreso de términos específicos en la Interfaz de Búsqueda. La rapidez y eficacia de este proceso dependerá de dos factores:

- a) De la información descriptiva con la que el documento fue ingresado.
- b) De las palabras usadas para la búsqueda.

Además, el Módulo Administrativo permitirá manejar y controlar usuarios y sus accesos, documentos y tipos de documentos.

4.1.2 Alcance del sistema:

La funcionalidad del sistema será la siguiente:

- Se accederá al sistema a partir de un usuario y contraseña proporcionado por el Administrador.
- Cada usuario contará con un permiso de tipo sólo lectura (R), lectura y escritura (RW) o ninguno (N). Estos permisos se aplicarán por cada tipo de documento que exista. Así, cuando el usuario A tenga permiso tipo RW sobre el tipo de documento Contratos y permiso tipo R sobre Facturas, podrá ingresar documentos de Contratos y únicamente ver documentos de Facturas.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Manejador de Documentos: Es el encargado del ingreso de documentos y permitirá:

- Almacenar los archivos en el disco del servidor, que estén en los siguientes formatos: TIF, .JPG o .PDF, y que no excedan del tamaño definido de acuerdo al motor de base de datos que se utilice.
- Además permitirá al usuario llenar todos los datos del documento en el momento de su ingreso.
- En esta implementación los archivos se guardarán en un directorio dentro del servidor. Pero también es posible guardarlos directamente en la Base de Datos. Es importante recalcar que el ingreso o eliminación de documentos se deberá realizar únicamente a través de la aplicación para así evitar inconsistencias entre registros en la Base de Datos y archivos en el servidor.

A continuación, el diagrama muestra lo descrito para el Manejador de Documentos:

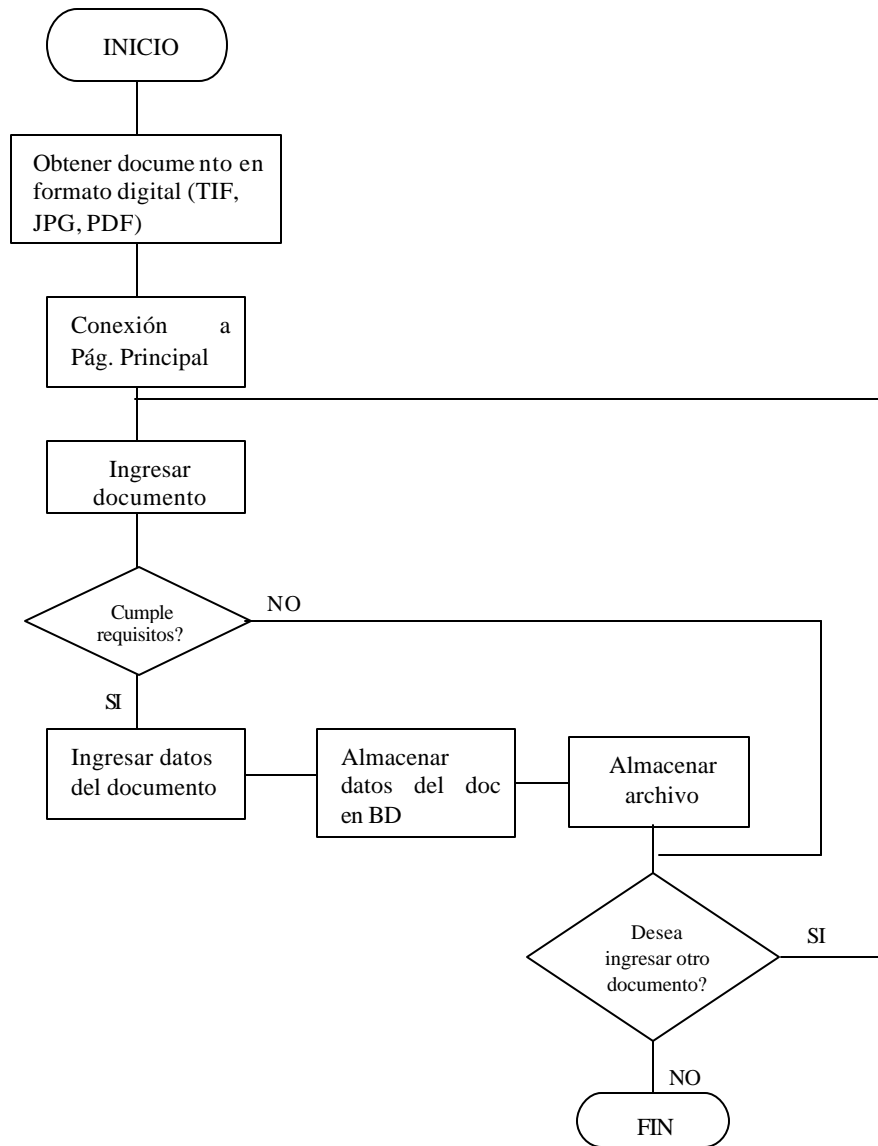


Gráfico 4-1 Diagrama de Procesos para Manejador de Documentos
Elaboración: María Augusta Mateus

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Interfaz de Búsqueda: Es el encargado de buscar documentos y presentarlos al usuario, y permitirá:

- Colocar los términos a buscar en una caja de texto. Se podrá usar operadores booleanos AND u OR para facilitar la búsqueda.
- Mediante una función de evaluación definida se cuantificarán algunos datos de los documentos. Así, si la palabra que estamos investigando está dentro del título, adquirirá un valor, si además está dentro del resumen se sumará otro valor, etc. Con esto, a mayor valor de la función de evaluación, el documento contendrá la información más cercana a lo que se busca, y viceversa.

La función de evaluación es:

$$F.E.= C_1Pr + C_2Pt + C_3Pp + C_4Pc$$

Donde,

Pr: número de palabras que están dentro del resumen.

Pt: número de palabras que están dentro del título.

Pp: número de palabras que están dentro del tipo de documento.

Pc: número de palabras que están dentro de las palabras claves.

C₁: valoración de que una palabra esté contenida en el resumen.

C₂: valoración de que una palabra esté contenida en el título.

C₃: valoración de que una palabra esté contenida en el tipo de documento.

C₄: valoración de que una palabra esté contenida en las palabras claves.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Los valores C_1 , C_2 , C_3 , C_4 con los cuales se trabajará son: 1, 2, 4 y 8 respectivamente, donde se considera que las palabras contenidas en las palabras claves tienen el doble de importancia, que palabras dentro de tipo de documento, y que éstas tienen el doble de importancia que palabras en el título, y a su vez que palabras en el título son el doble de importantes que palabras en el resumen.

Los resultados se presentarán en orden de pertinencia de mayor a menor valor de la función de evaluación, indicando el título del documento, los 200 primeros caracteres del resumen, tipo de documento y tamaño del archivo. Se permitirá al usuario ver los documentos presentados, en función de los derechos que tenga dicho usuario en el sistema.

En el siguiente diagrama se muestra la funcionalidad de la Interfaz de Búsqueda:

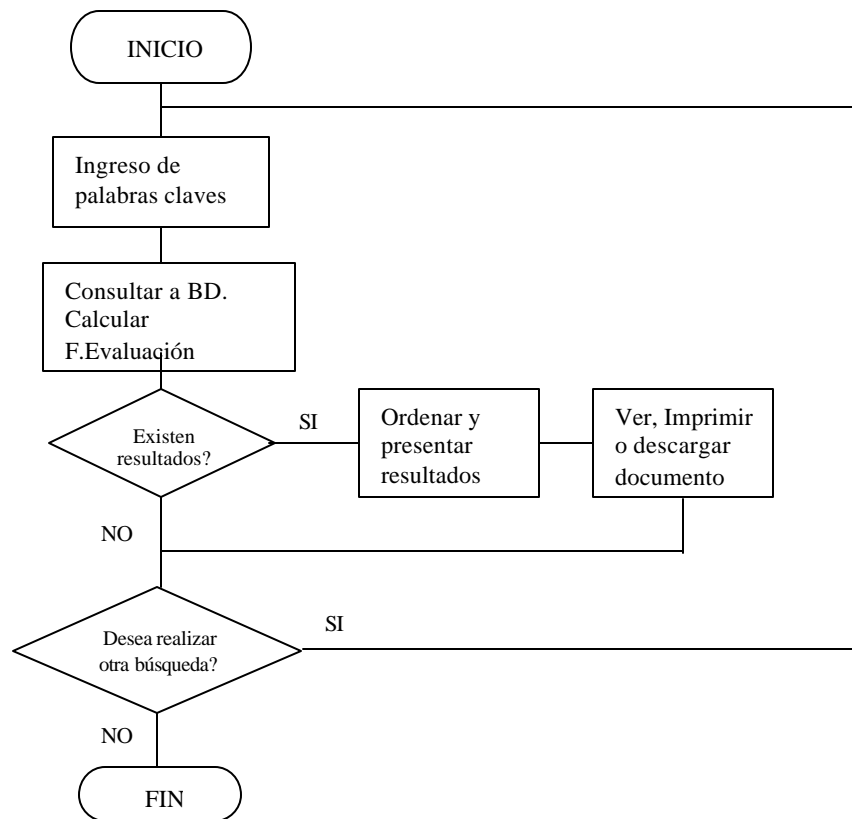


Gráfico 4-2 Diagrama de Procesos para Buscador de Documentos
Elaboración: María Augusta Mateus

Módulo Administrativo: este módulo permitirá:

- Administrar los documentos y sus datos.
- Administrar tipos de documentos, es decir: crear, eliminar o modificar tipos de documentos.
- Administrar usuarios, que se refiere a: crear, eliminar usuarios y cambiar la contraseña de un usuario.

El siguiente diagrama indica las funciones descritas para el Módulo

Administrativo:

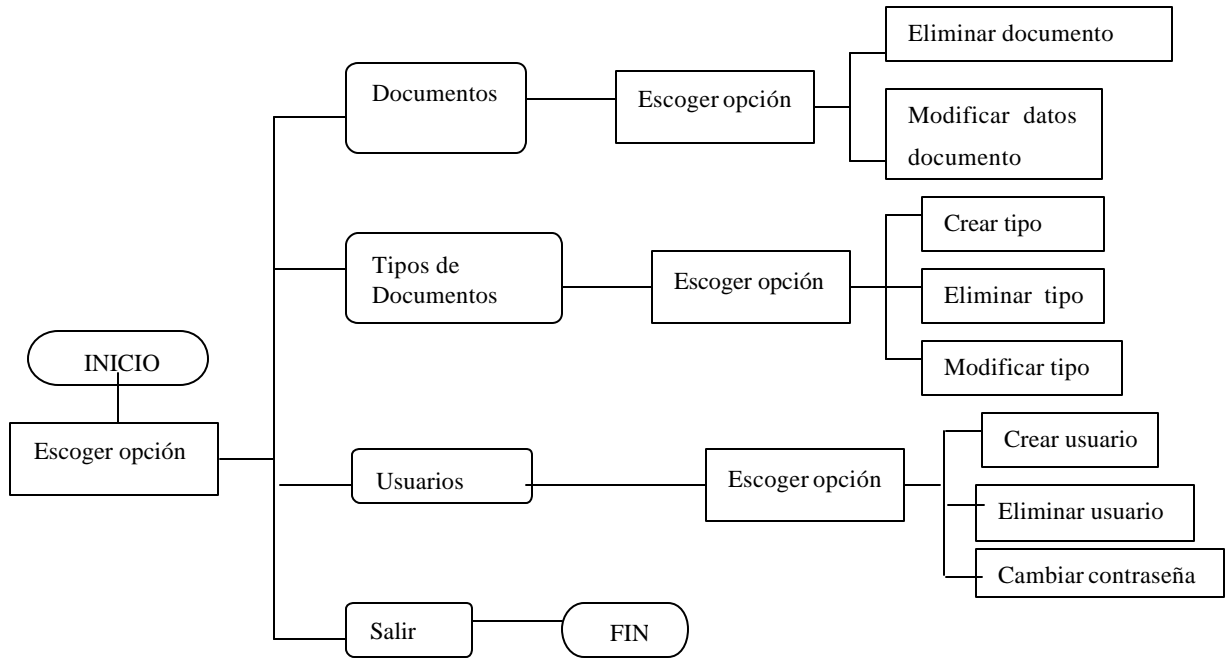


Gráfico 4-3 Diagrama de Procesos para Módulo Administrativo
Elaboración: María Augusta Mateus

4.2 Diseño

Del análisis expuesto anteriormente, se definió el Modelo de Datos (Entidad Relación) siguiente:

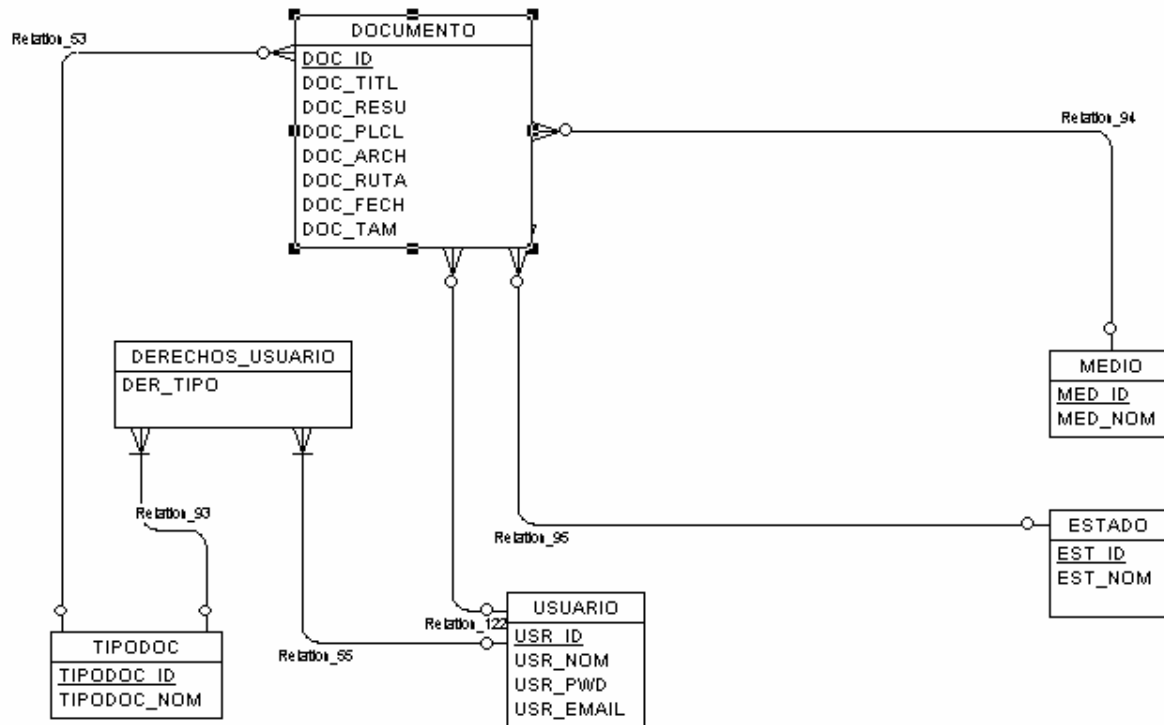


Gráfico 4-4 Modelo E-R
Elaboración: María Augusta Mateus

Como se puede ver en el modelo Entidad Relación presentado, para cada documento se guardará el estado en el que se encuentra que puede ser legible o ilegible, se incluirá el medio físico del documento original como papel, disco compacto, disquete u otro, y además se indicará un identificador único, el título del documento, un breve resumen del documento, un listado de palabras claves, el nombre del archivo digital, la ruta en la cual se grabará el archivo en el servidor, la fecha de ingreso del documento al sistema, y el tamaño en bytes del archivo. Los documentos pertenecerán a un tipo de documento que puede ser factura, contrato, reporte u otro. A estos documentos tendrán acceso los usuarios, quienes

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

estarán identificados por un identificador único, nombre, contraseña y e-mail. Por cada tipo de documento se asignarán derechos de lectura, lectura y escritura o ninguna, lo que significa que si un usuario tiene un derecho para un tipo de documento tendrá derecho para todos los documentos de ese tipo.

El Diccionario de Datos correspondiente a este modelo, se presenta en el **Anexo 1.**

DESARROLLO, IMPLEMENTACIÓN Y PRUEBAS

5.1 Desarrollo del Sistema

La aplicación cuenta con tres módulos: búsqueda de documentos, ingreso de documentos y módulo de administración. Los mismos que fueron desarrollados en paginas asp (80% del la aplicación).

Para el ingreso de documentos se utilizó y se adecuó a las necesidades el código en lenguaje VB script proporcionado por Carlos de la Orden (pseudónimo) del sitio web *www.aspfacil.com*. El mencionado código permite subir archivos de varios formatos a un directorio en el servidor y grabar información del archivo como su nombre y tamaño.

El **Anexo2** al final del presente documento muestra el código utilizado para subir archivos al servidor.

Para la búsqueda de documentos se desarrolló una serie de funciones (en la aplicación) y procedimientos almacenados (en la base de datos) los mismos que se explican a continuación:

La función *sacarterminos* que separa los términos de búsqueda, filtrando por signo + o espacio en blanco si hubieren.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Para inicializar a cero el campo *Doc_aux* de todos los registros en la tabla Documento se ejecuta el procedimiento almacenado *sp_upd_cero*.

Se determina si el operador usado fue espacio en blanco o signo (+) y según ello se ejecuta el procedimiento *sp_calc_c* o *sp_calc_c_and* respectivamente enviando como parámetros los términos de búsqueda y el número de términos. Este procedimiento almacenado *sp_calc_c* o *sp_calc_c_and* calcula los valores C_1, C_2, C_3, C_4 indicados en la función de evaluación que se detalló en el capítulo 4.

Se suman los valores C_1, C_2, C_3, C_4 de cada documento y con ese valor se actualizan el campo *Doc_aux* de cada documento.

Se obtiene el sql de la vista creada entre *Tipodoc* y *Documento*, llamada *v_doc_tipo*. Además se filtran por los permisos que el usuario tenga sobre el tipo de documento y se ordenan los documentos por el campo *Doc_aux* que contiene el valor de la función de evaluación para cada documento.

Para comprender mejor el algoritmo de búsqueda y cálculo de la función de evaluación, el **Anexo 3** al final del presente documento muestra las funciones y procedimientos almacenados implementados.

5.2 Implementación del Sistema

En esta sección se realiza una breve descripción de los recursos de software necesarios para implementar el prototipo del “Sistema de Manejo de Documentos” y se detalla el funcionamiento general de la aplicación.

5.2.1 Recursos de Software

Servidor

Windows 2000 Server

Microsoft SQL Server 2000 Standard Edition

Internet Information Server 5.0

Cliente

Internet Explorer 5 o superior

Acrobat Reader 5.0

5.2.2 Recursos de Hardware

Servidor

1 Procesador de 512 KHz

128 Mb de RAM

Espacio en Disco Necesario 400 Mb (Base de Datos + Aplicación)

5.2.3 Funcionamiento de la aplicación

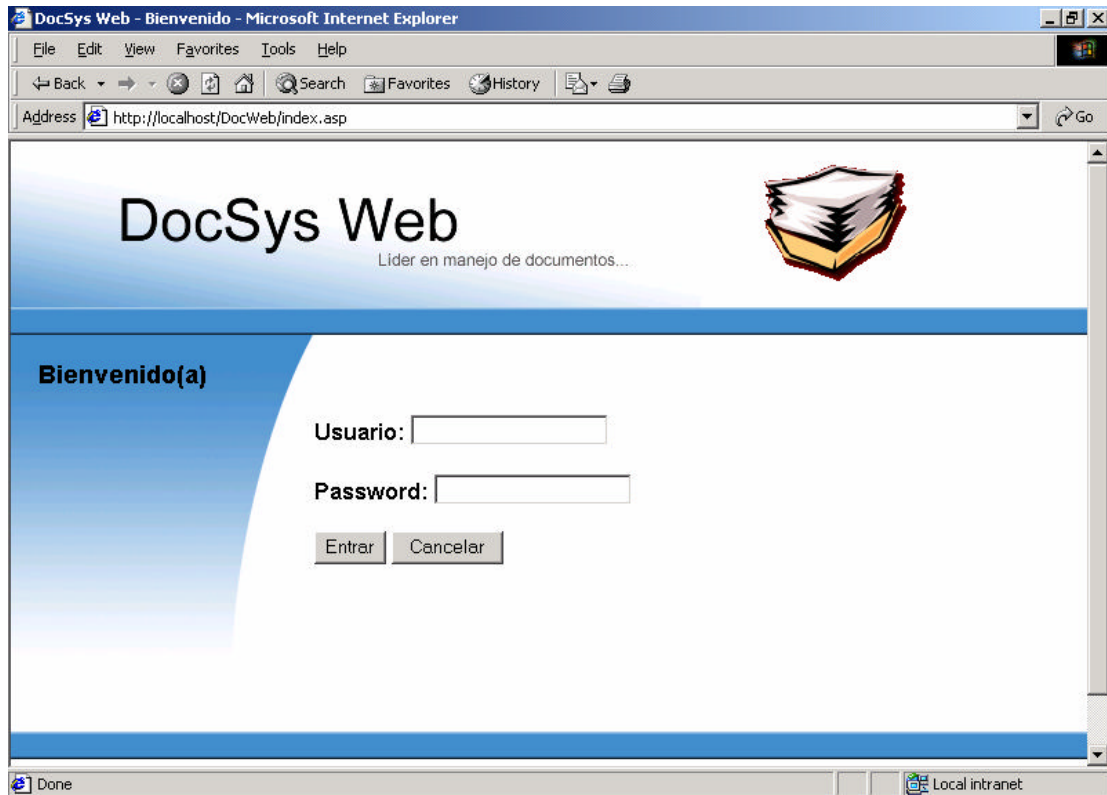
Esta aplicación fue realizada con Macromedia Dreamweaver Ultradev 4.0 utilizando principalmente lenguaje Visual Basic Script. Contiene además código HTML y varias secciones en lenguaje Java Script. Posee una interfaz amigable y de fácil uso para el usuario final.

Como se menciona en los requerimientos de software, el usuario final necesita únicamente de un Navegador (Internet Explorer) para acceder a ella, además del usuario y password proporcionado por el Administrador del Sistema.

El link de acceso a la aplicación en mención es:
<http://localhost/DocWeb/index.asp> que desplegará la siguiente pantalla:

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador



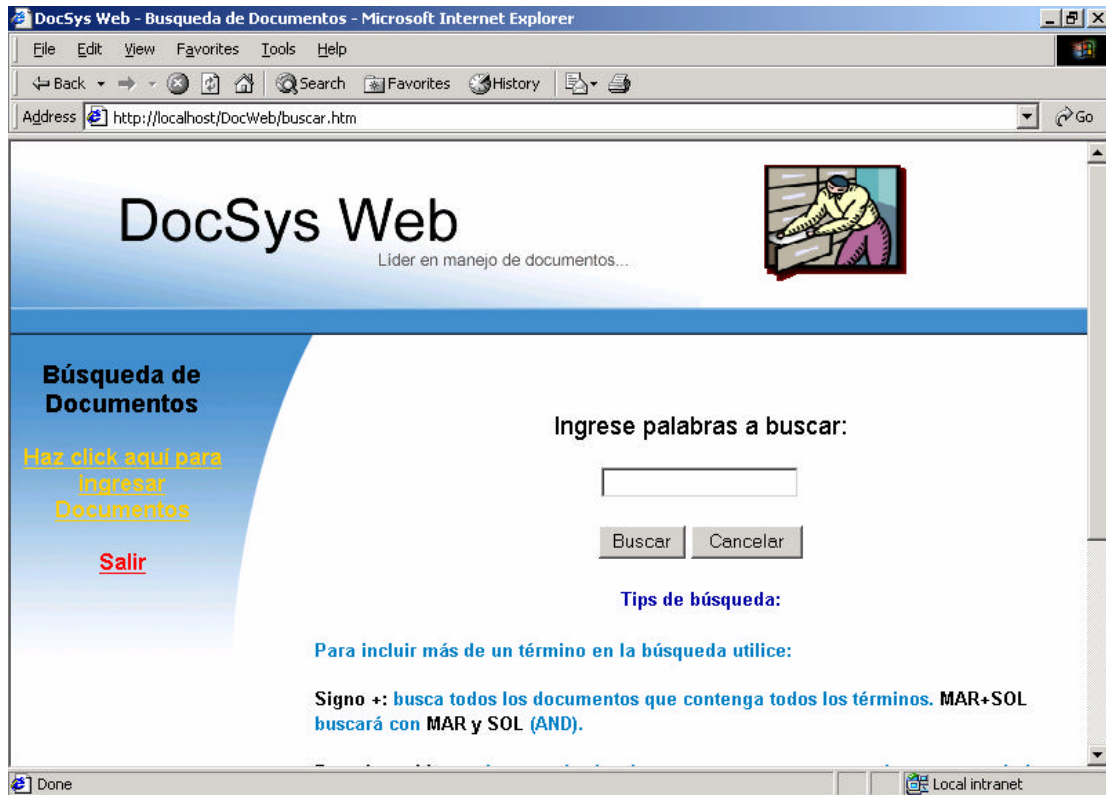
Al ingresar el usuario y password correspondiente, y luego de presionar Entrar; aparecerá la pantalla de búsqueda.

En esta pantalla se ingresarán los términos de búsqueda, con la ayuda de los tips de búsqueda usted llegará al documento más rápidamente.

Además puede ingresar documentos presionando el link correspondiente a la izquierda.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador



En esta pantalla se ingresarán los términos de búsqueda, con la ayuda de los tips de búsqueda usted llegará al documento más rápidamente.

Además puede ingresar documentos presionando el link correspondiente a la izquierda.

Si los términos de búsqueda fueron ingresados y se presionó Buscar, en pocos segundos se desplegarán los resultados que satisfagan su búsqueda en la siguiente pantalla:

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Ingrese palabras a buscar:

Id	Título	Resumen	Tipo	Tamano	Archivo	Descarga
6	MANUAL REFERENCIA MYSQL 1	MANUAL DE MYSQL PARA DESARROLLO DE BASE DE DATOS EN WEB	TESIS	2.82	Ver	Descargar
7	WEB DATABASE STEP BY STEP GUIDE	GUIA PARA CREACION DE APLICACIONES WEB CON BASE DE DATOS MYSQL	TESIS	0.05	Ver	Descargar
24	REPORTE DE AVANCES DE PROYECTO EQUIPOS ANTIDERRAME	PROYECTO ANTIDERRAME JUNIO 2003 LAGO AGRIO	MANUALES	0.12	Ver	Descargar
28	COMPENDIO GENERAL DE EQUIPOS DE PLANTA PAPALLACTA	LISTA DE EQUIPOS EN PLANTA PAPALLACTA CON EL DETALLE DE ESPECIFICACIONES TECNICAS	MANUALES	7.01	Ver	Descargar

Si desea ver o descargar cualquiera de los documentos utilice los links a la derecha de cada documento para este propósito.

Además puede hacer una nueva búsqueda al dar click en el link que lo indica, con lo cual regresará a la pantalla de ingreso de palabras de búsqueda

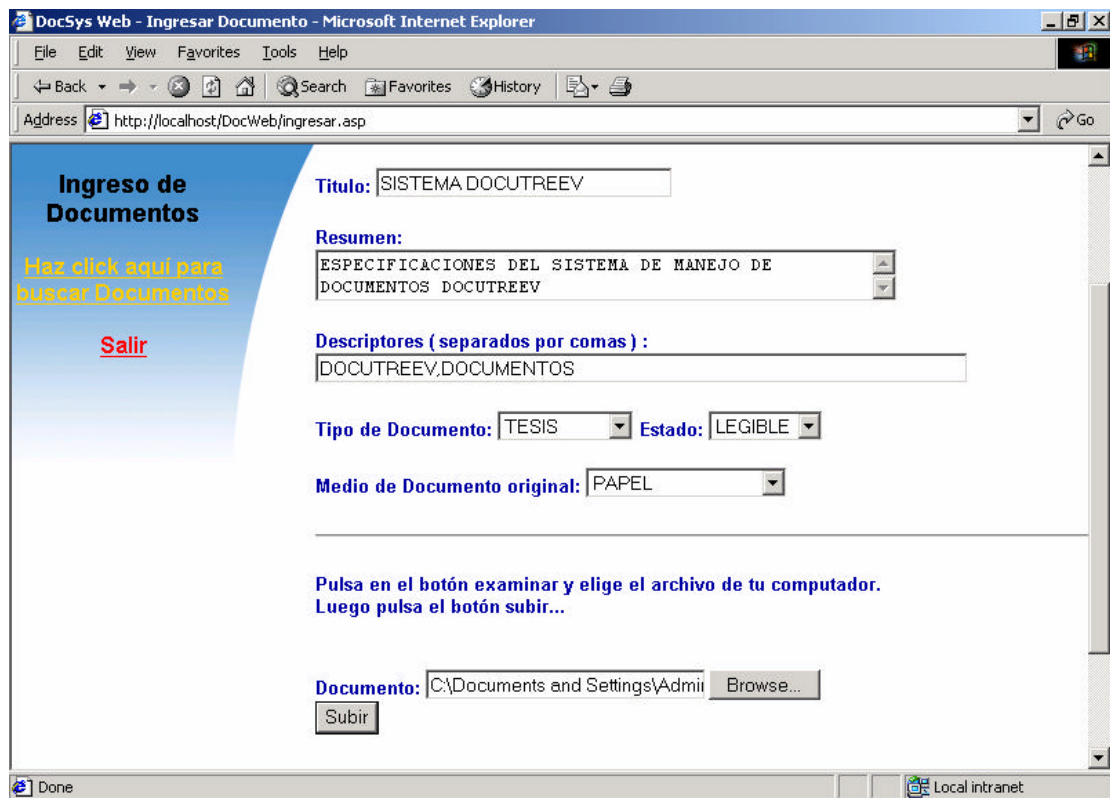
Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

En la pantalla de ingreso de nuevos documentos usted distinguirá dos secciones:

La primera donde se ingresan los datos del documento como título, resumen, descriptores, tipo de documento, estado y medio físico del documento original.

Y la sección dos que permite subir el archivo. Tal como se muestra a continuación:



The screenshot shows a web browser window titled "DocSys Web - Ingresar Documento - Microsoft Internet Explorer". The address bar shows "http://localhost/DocWeb/ingresar.asp". The page content is as follows:

Ingreso de Documentos

[Haz click aquí para buscar Documentos](#)

[Salir](#)

Título: SISTEMA DOCUTREEV

Resumen: ESPECIFICACIONES DEL SISTEMA DE MANEJO DE DOCUMENTOS DOCUTREEV

Descriptores (separados por comas) : DOCUTREEV,DOCUMENTOS

Tipo de Documento: TESIS **Estado:** LEGIBLE

Medio de Documento original: PAPEL

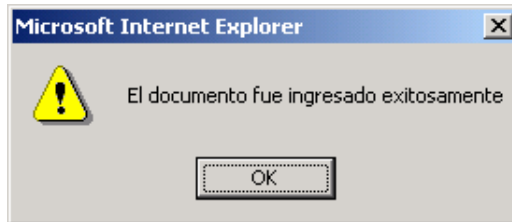
Pulsa en el botón examinar y elige el archivo de tu computador. Luego pulsa el botón subir...

Documento: C:\Documents and Settings\Admi...

Luego de ingresar los datos y el archivo correspondientes, y al presionar Subir, usted recibirá una notificación de que el documento fue añadido al sistema.

Desarrollo de prototipo de un sistema de manejo de documentos

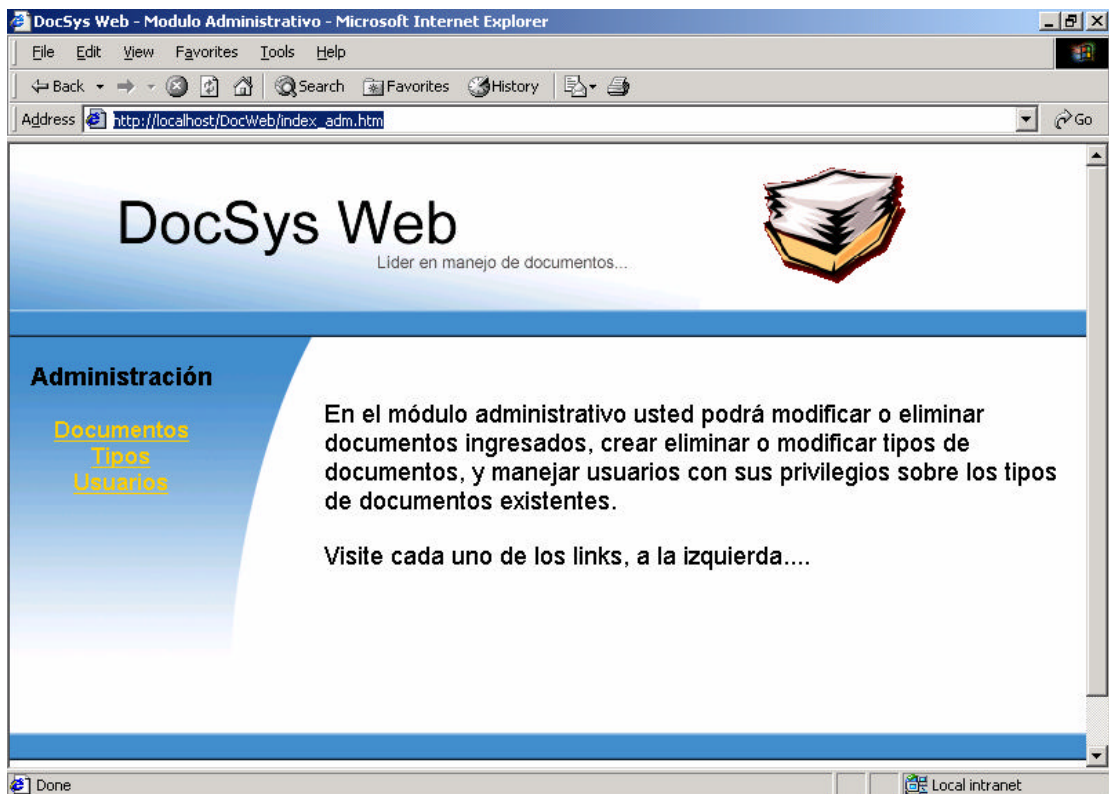
Pontificia Universidad Católica del Ecuador



Modulo de Administración del Sistema

Para acceder al módulo de Administración de la aplicación utilice el siguiente link: http://localhost/DocWeb/index_admin.htm

Luego de lo cual se desplegará la siguiente pantalla:



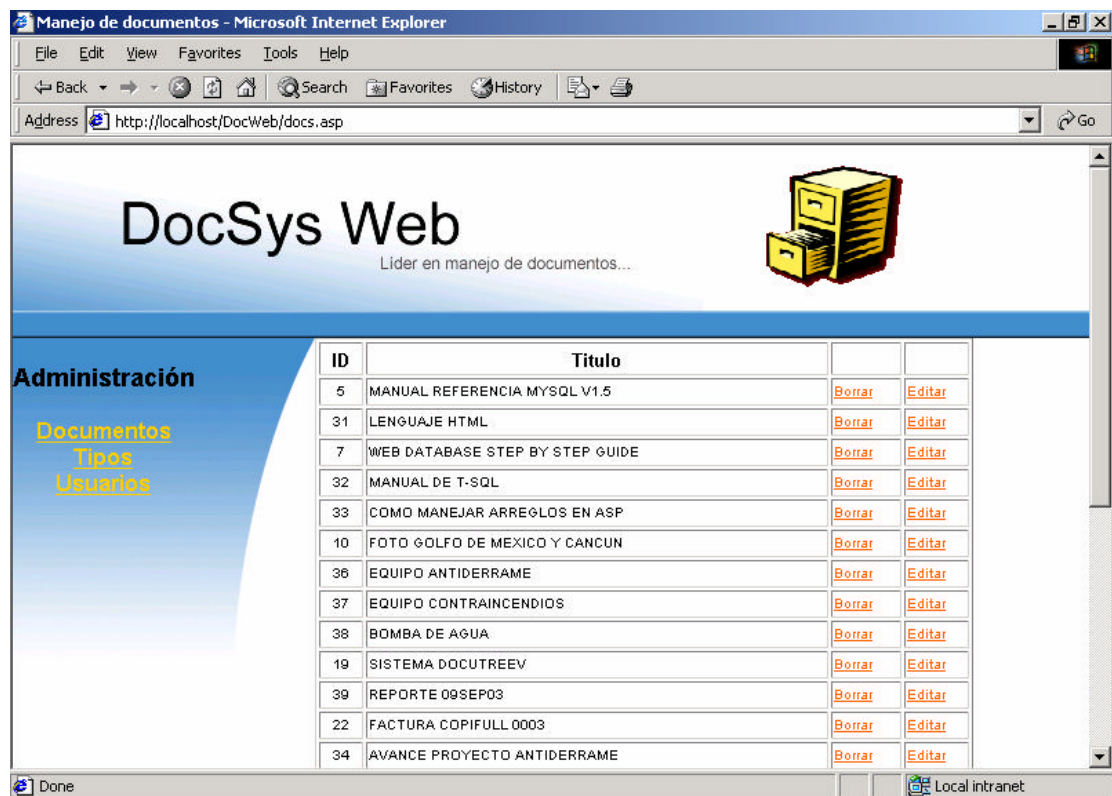
A continuación revisaremos brevemente lo que cada uno de los links a la izquierda permite realizar:

Desarrollo de prototipo de un sistema de manejo de documentos

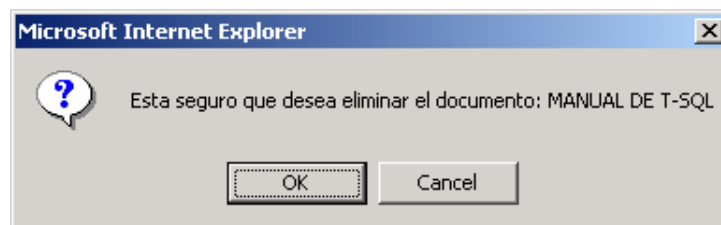
Pontificia Universidad Católica del Ecuador

Documentos

Al dar click en Documentos usted recibirá la siguiente pantalla, con la lista de documentos y la opción de borrarlo o editarlo, según el link que elija.

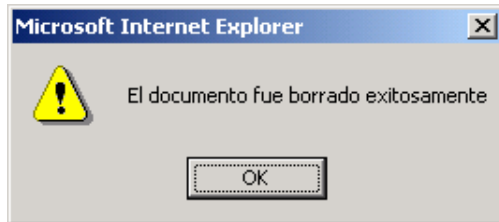


Al elegir Borrar, aparece un mensaje de confirmación del documento que se desea borrar y si se acepta entonces el documento (incluido el archivo físico) serán eliminados del sistema. Una vez borrado el documento, el usuario recibirá una notificación de dicha acción.

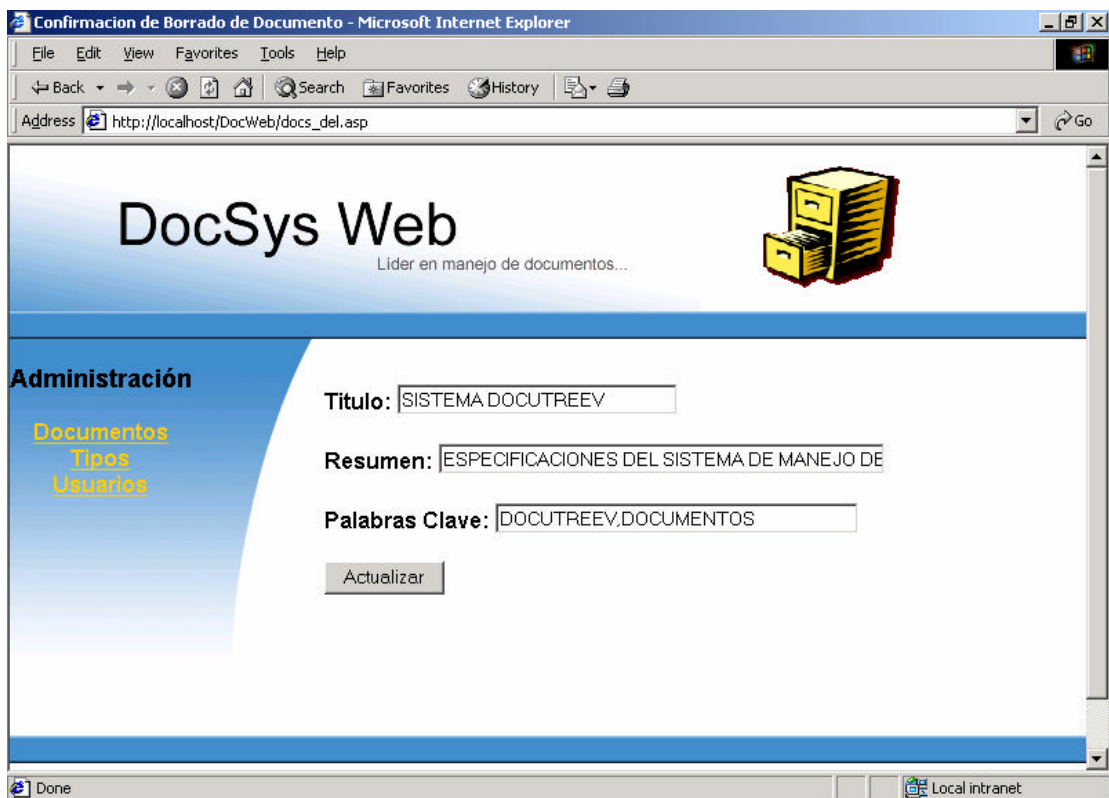


Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador



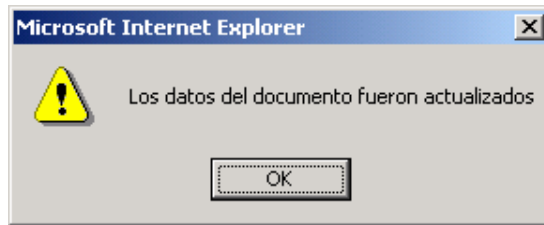
Si por el contrario, eligió Editar aparecerá la siguiente pantalla con el Título, Resumen y Palabras Claves del documento al cual desea hacer cambios.



Al hacer los cambios necesarios y presionar Actualizar, aparecerá un mensaje que le indicará que los cambios se grabaron.

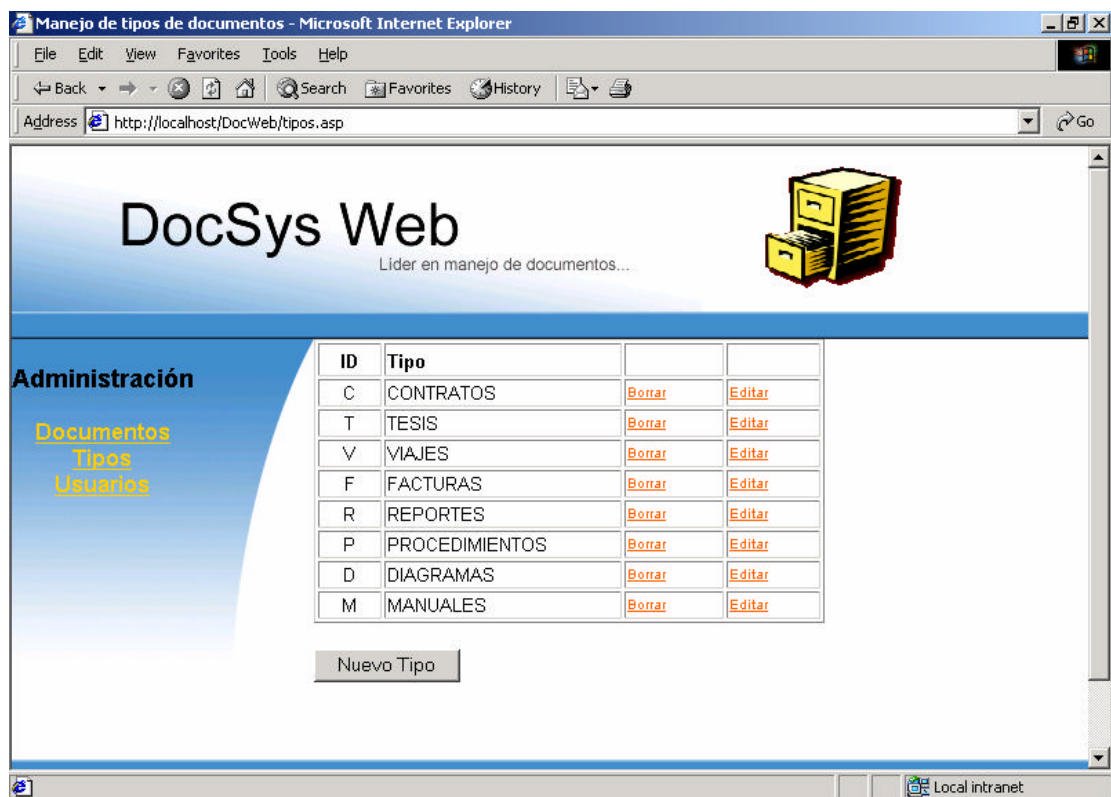
Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador



Tipos

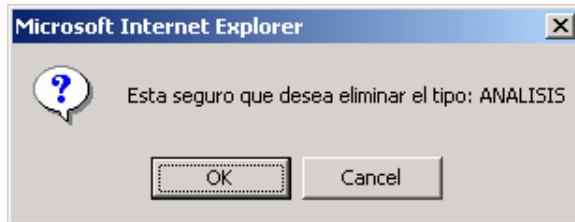
Al dar click en Tipos usted recibirá la siguiente pantalla, con la lista de tipos de documentos y la opción de borrarlos o editarlos, según el link que elija. Además un botón Nuevo Tipo para ingresar un nuevo tipo de documento.



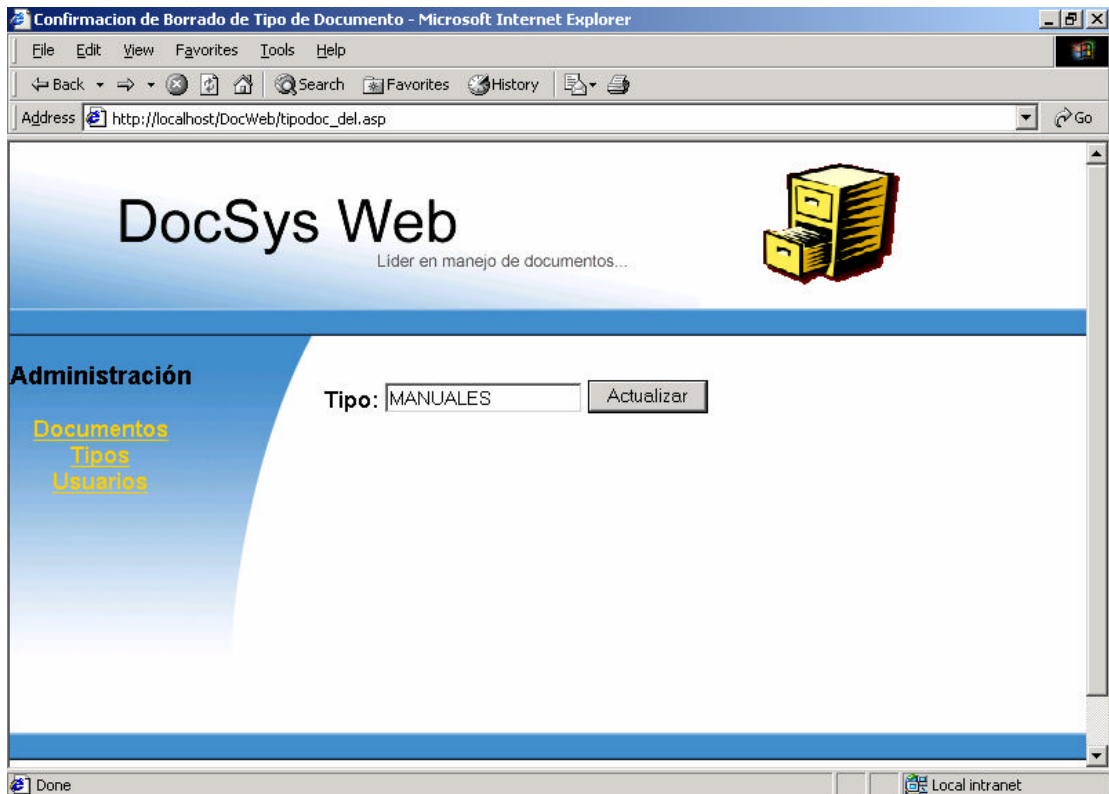
Al elegir Borrar, el tipo de documento será eliminado del sistema. Luego de pedir la confirmación del usuario para borrarlo. Una vez borrado el tipo de documento, el usuario recibirá una notificación de dicha acción.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador



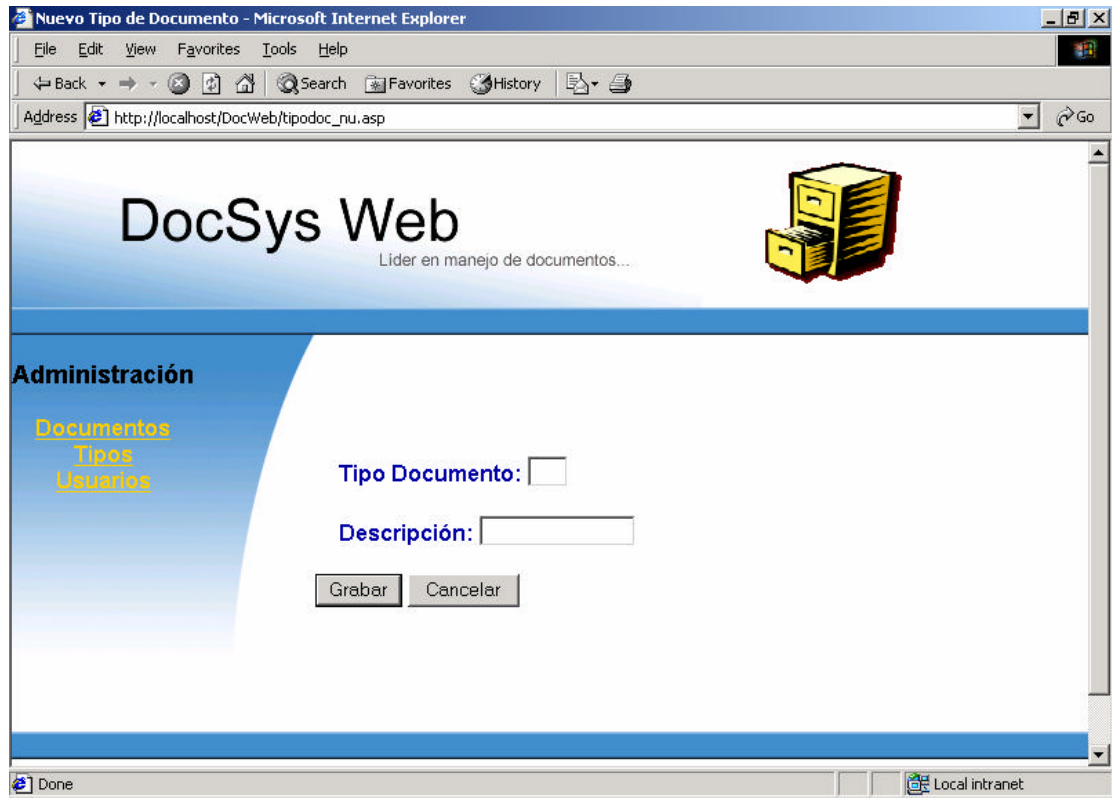
Si eligió Editar aparecerá la siguiente pantalla con el Tipo de Documento al que desea hacer cambios. Luego de presionar Actualizar, los cambios se grabarán.



Por el contrario, al presionar Nuevo Tipo la siguiente pantalla le permitirá ingresar un nuevo Tipo de Documento:

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

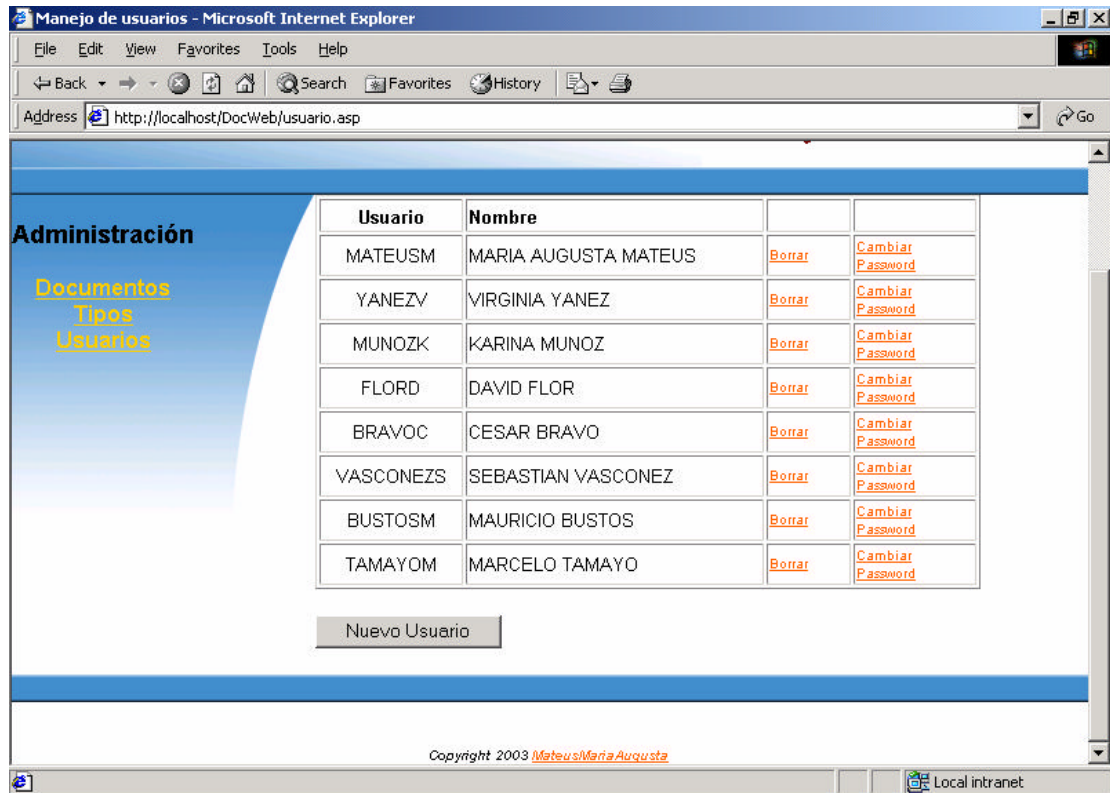


Usuarios

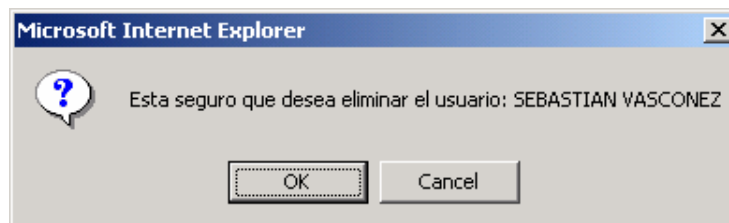
Al dar click en Usuario usted recibirá la siguiente pantalla, con la lista de usuarios y la opción de borrarlos o cambiar su contraseña, según el link que elija. Además un botón Nuevo Usuario para ingresar un nuevo usuario.

Desarrollo de prototipo de un sistema de manejo de documentos

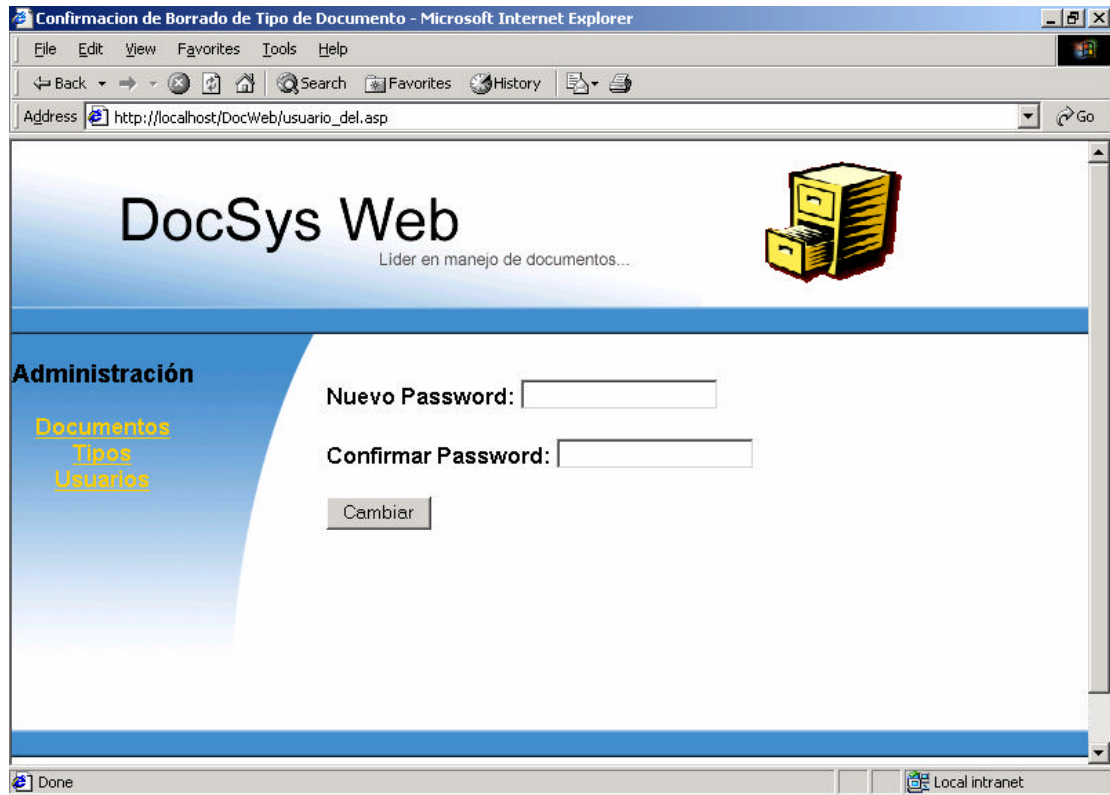
Pontificia Universidad Católica del Ecuador



Al elegir Borrar aparecerá un mensaje de confirmación de borrado, al aceptarlo el usuario será eliminado del sistema. Una vez borrado el usuario, se recibirá una notificación de dicha acción.



Si eligió cambiar de password, recibirá una pantalla para ingresar el nuevo password y confirmarlo. Luego presione cambiar para guardar los cambios.



Si por el contrario, usted presiona Nuevo Usuario la siguiente pantalla le permitirá ingresar un nuevo Usuario. Esta pantalla consta de dos secciones, la primera con los datos generales del usuario como Usuario, Password, Nombre, E-mail. Y la segunda sección con los tipos de documentos existentes y los permisos que deberá tener este usuario sobre cada tipo de documento. Al presionar Grabar el nuevo usuario se grabará.

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

Nuevo Usuario - Microsoft Internet Explorer

Address http://localhost/DocWeb/usuario_nu.asp

Administración

[Documentos](#)
[Tipos](#)
[Usuarios](#)

Usuario:

Password:

Nombre:

E-mail:

Permisos por Tipo de documento

Tipo	Permiso (N, R o RW)
CONTRATOS	<input type="text" value="N"/>
FACTURAS	<input type="text" value="N"/>
MANUALES	<input type="text" value="N"/>
PROCEDIMIENTOS	<input type="text" value="N"/>
REPORTES	<input type="text" value="N"/>
TESIS	<input type="text" value="N"/>
VIAJES	<input type="text" value="N"/>

Done Local intranet

5.3 Pruebas

Para propósitos de pruebas del módulo de Ingreso de Documentos, de entre los formatos permitidos en la aplicación (JPG, PDF y TIF), se escogió archivos de diversos tamaños para subirlos. Los tiempos de respuesta obtenidos son los siguientes:

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

No.	Tamaño	Formato	Tiempo (seg)
1	106 KB	TIF	7.12
2	100 KB	PFD	6.28
3	228 KB	JPG	16.22
4	1133 KB	JPG	84.97
5	124 KB	PDF	8.25
6	141 KB	TIF	9.47
7	473 KB	JPG	34.39
8	100 KB	PDF	6.42
9	7180 KB	PDF	502.69
10	491 KB	JPG	35.54

Tabla 5-1 Tiempos de respuesta en subir documentos
Elaboración: María Augusta Mateus

Como se puede apreciar el tiempo óptimo de ingreso de un documento es independiente del formato del archivo, sin embargo depende del tamaño del mismo.

Con el usuario de la aplicación BustosM, que posee permisos de lectura sobre los tipos de documentos Diagramas, Viajes, Tesis y Reportes. Se ha realizado la búsqueda con las siguientes frases y los tiempos de búsqueda obtenidos con 30 documentos, se presentan a continuación:

Cadena de búsqueda	Tiempo (seg)	Num. Docs.
reportes+papallacta	0.26	3
Reportes	2.10	10
Reportes papallacta	1.51	4
Diagrama equipo	1.86	7
Factura copifull 2003	2.52	10
Contraincendio	1.39	2

Tabla 5-2 Tiempos de respuesta en buscar documentos
Elaboración: María Augusta Mateus

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

De los resultados obtenidos se puede concluir que a la aplicación le toma mayor tiempo encontrar el documento, mientras la cadena de búsqueda es más compleja.

Los procesos de descarga o vista del contenido de un documento en general son satisfactorios, pero su velocidad mejorará si la capacidad de memoria del equipo es adecuada.

CONCLUSIONES Y RECOMENDACIONES

6.1 Conclusiones

- Sí es posible manejar documentos mediante una base de datos relacional, ello permite asociar un documento con otro tipo de información no necesariamente documental dentro del manejo automatizado de un sistema de información.
- Como se pudo demostrar en las pruebas realizadas, el rendimiento de la aplicación durante el ingreso de archivos es proporcional al tamaño del archivo que se desea subir,. Por tal razón archivos comprimidos, tales como formatos PDF y JPG mejoran el tiempo de ingreso.
- El cálculo de la función de evaluación, necesaria en el proceso de búsqueda de documentos es eficiente si lo realiza la base de datos mediante procedimientos almacenados, ya que hacerlo en la aplicación implica que el objeto Recordset que es el que contiene los registros traídos de la base de datos deba hacer repetidas peticiones según la cantidad de datos contenidos en cada registro.
- En principio el motor de base de datos utilizado fue MySQL, que es un software gratis, accesible y de fácil manejo, pero se lo reemplazó por dificultades en el Control de Integridad, Generación de Vistas, Tablas Temporales, y la Implementación de Procedimientos Almacenados. En su

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

lugar se utilizó Microsoft SQL Server 2000, el cual a pesar de ser un motor de base de datos comercial es conocido por mi persona.

- Los valores 1, 2, 4 y 8 para palabras en Resumen, Título, Tipo y Descriptores respectivamente, permiten establecer un orden adecuado de importancia a los campos que describen un documento, ya que un resumen puede contener palabras de relleno (preposiciones, artículos, etc.) o derivadas, haciendo perder el enfoque de la búsqueda.
- Las empresas deben contar con un sistema como el presentado para manejo documental, ya que permite organizar los documentos, protegerlos de posibles desastres que ocasionen pérdidas de información y ahorrar espacio, tiempo y papel. Todo ello contando además con políticas de respaldo y planes de contingencia.
- Sí es posible lograr oficinas sin papeles que manejen toda la información en forma digital. Siempre y cuando las seguridades sean adecuadas, y complementándolo con tecnologías como firmas digitales, correo electrónico, comercio electrónico y otros en situaciones que lo requieran.
- Nos encontramos en el inicio de una nueva era en el manejo de la información, el futuro son los documentos de audio y video, que se constituirán en las principales fuentes de información para cualquier tipo de empresa ya que transmiten información en forma exacta y completa.

6.2 *Recomendaciones*

- Para mejorar el rendimiento del ingreso de documentos en lo que se refiere al cálculo del tamaño del archivo se deben combinar funciones de sistema operativo con las utilizadas en el presente trabajo.
- Futuras versiones del sistema podrían desarrollar los siguientes aspectos:
Utilización de formatos para texto, audio y video como los descritos en el capítulo 3, ya que son frecuentemente usados en la actualidad, y en el futuro podrían convertirse en una parte común de los datos manejados en una empresa.
Y, la implementación de técnicas de OCR para la búsqueda por contenido en imágenes, puesto que gran parte de los documentos de una empresa están en papel y muchas veces es necesario obtener el contenido de ellos.
- A raíz del trabajo realizado se propone los siguientes temas de disertación: integración de sistemas de manejo documental con otros sistemas como: contables, de nómina, etc. Tecnologías en el procesamiento de audio y video. Y, esquemas de replicación de documentos desde diversas áreas geográficas.
- El presente trabajo podría ser de gran utilidad en la Universidad, donde grandes centros de documentación como la Biblioteca y Secretarías General y de la Facultad, pongan al alcance de los estudiantes

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

documentación personal, libros, extractos de disertaciones, etc. en forma digitalizada.

- De las materias recibidas durante la carrera se han aplicado para el presente tema de disertación principalmente Procesamiento de Imágenes, Bases de Datos e Ingeniería del Software y considero que los conocimientos proporcionados a partir de las mencionadas materias fueron beneficiosos y útiles, además recomiendo incluir nuevas tecnologías para cada una de ellas.

ANEXOS

Anexo 1:

Diccionario de Datos

Tabla	Definición de la Tabla	Atributos	Definición de Atributos	Tipo de Datos
Usuario	Usuarios del sistema	Usr_id	Id del usuario	Varchar(9)
		Usr_nom	Nombre del usuario	Varchar(30)
		Usr_pwd	Password del usuario	Varchar(9)
		Usr_e-mail	E-mail del usuario	Varchar(30)
Tipodoc	Tipo de documento	Tipodoc_id	Id del tipo de documento	Varchar(1)
		Tipodoc_nom	Nombre del tipo de documento	Varchar(15)
Estado	Estado del documento físico	Est_id	Id del estado	Varchar(1)
		Est_nom	Nombre del estado	Varchar(15)
Medio	Medio físico del documento	Med_id	Id del medio físico	Varchar(1)
		Med_nom	Nombre del medio físico	Varchar(15)
Derechos_usuario	Permisos por tipo de documento para cada usuario	Usr_id	Id del usuario	Varchar(9)
		Tipodoc_id	Id del tipo de documento	Varchar(1)
		Der_tipo	Tipo de permiso sobre el tipo de documento. Puede ser R (Read), RW (Read Write), N (None)	Varchar(2)
Documento	Documentos ingresados al sistema	Doc_id	Id del documento	Integer
		Doc_titl	Título del documento	Varchar(200)
		Doc_res	Descripción del contenido del documento	Varchar(250)
		Doc_plcl	Descriptor del documento	Varchar(200)
		Doc_arch	Nombre del archivo con extensión	Varchar(30)

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

		Doc_ruta	Lugar donde se guarda el archivo en servidor	Varchar(200)
		Doc_fech	Fecha de ingreso del documento al sistema	DateTime
		Doc_tam	Tamaño del archivo en Megabytes.	Float
		Usr_id	Id del usuario que ingresó el documento	Varchar(9)
		Tipodoc_id	Id del tipo de documento	Varchar(1)
		Med_id	Id del medio físico	Varchar(1)
		Est_id	Id del estado del documento	Varchar(1)

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
'Permite hacer, por ejemplo: Response.Write(upload.Form("nombre"))
Public Property Get Form(campo)
    if eltosForm.Exists(campo) then
        Form = eltosForm.Item(campo)
    else
        Form = ""
    end if
End Property
'-----
Public Sub Upload()
'Inicia el proceso. Debe llamarse ANTES DE HACER CUALQUIER OTRA COSA

Dim byteDatos, strControl
Dim iPosInicio, iPosFin, iPos, byteLimite, posLimite
Dim iPosFich, iPosLim

byteDatos = Request.BinaryRead(Request.TotalBytes)
iPosInicio = 1
iPosFin = InStrB(iPosInicio, byteDatos, str2byte(chr(13)))
if (iPosFin-iPosInicio) <= 0 then
'terminamos, no hay nada que leer
    Exit Sub
end if
'extraemos el limite de principio y fin de los datos (p.e. ---- 2323g237623)
byteLimite = MidB(byteDatos, iPosInicio, iPosFin-iPosInicio)
posLimite = InStrB(1, byteDatos, byteLimite)

'terminamos cuando la posición del próximo límite sea igual
'a la del límite final, que lleva "-" detrás.
do until posLimite = InStrB(byteDatos, byteLimite & str2byte("-"))

    iPos = InStrB(posLimite, byteDatos, str2byte("Content-Disposition"))
    iPos = InStrB(iPos, byteDatos, str2byte("name=")) 'nombre del control en <FORM>
    iPosInicio = iPos + 6 'me salto 6 caracteres -> name="
    iPosFin = InStrB(iPosInicio, byteDatos, str2byte(chr(34))) 'busco las comillas de cierre
    'y tengo el nombre del control!
    strControl = byte2str(MidB(byteDatos, iPosInicio, iPosFin-iPosInicio))
    'busco ahora los datos en sí del control
    iPosFich = InStrB(posLimite, byteDatos, str2byte("filename="))
    posLimite = InStrB(iPosFin, byteDatos, byteLimite)

'¿ fichero o campo del formulario?
if iPosFich <> 0 and iPosFich < PosLimite then
    'es un fichero, creo un nuevo objeto fichero y lo añado a Ficheros
    Dim oFichero, strNombre, strForm
    set oFichero = new Fichero

    iPosInicio = iPosFich + 10 'me salto 10 caracteres -> filename="
    iPosFin = InStrB(iPosInicio, byteDatos, str2byte(chr(34)))
    strNombre = byte2str(MidB(byteDatos, iPosInicio, iPosFin-iPosInicio))
    'quito la ruta inicial
    oFichero.Nombre = Right(strNombre, Len(strNombre)-InStrRev(strNombre,
"\"))"

    iPos = InStrB(iPosFin, byteDatos, str2byte("Content-Type:"))
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
iPosInicio = iPos + 14 'me salto Content-Type y un espacio!!
iPosFin = InStrB(iPosInicio, byteDatos, str2byte(chr(13))) 'busco el retorno de
carro
oFichero.TipoContenido = byte2str(MidB(byteDatos, iPosInicio, iPosFin-
iPosInicio))

iPosInicio = iPosFin + 4 'me salto los 3 retornos de carro que lleva!!!
iPosFin = InStrB(iPosInicio, byteDatos, byteLimite)-2 'dos caracteres atrás
oFichero.Datos = MidB(byteDatos, iPosInicio, iPosFin -iPosInicio)
if oFichero.Tamano > 0 then 'lo añado a la colección Ficheros!
    Ficheros.Add strControl, oFichero
end if
else
'es un campo del formulario
iPos = InStrB(iPos, byteDatos, str2byte(chr(13)))
iPosInicio = iPos + 4
iPosFin = InStrB(iPosInicio, byteDatos, byteLimite)-2
'extraigo el valor del control del formulario!
strForm = byte2str(MidB(byteDatos, iPosInicio, iPosFin-iPosInicio))
if not eltosForm.Exists(strControl) then
    eltosForm.Add strControl, strForm
else
    eltosForm.Item(strControl) = eltosForm.Item(strControl)+","&strForm
end if
end if
'saltamos al siguiente límite
iPosLimite = InStrB(iPosLimite+LenB(byteLimite), byteDatos, byteLimite)
loop

End Sub
'-----
Private Function str2byte ( str )
Dim i, strbuf
for i = 1 to Len(str)
    strbuf = strbuf & ChrB(AscB(Mid(str, i, 1)))
next
str2byte = strbuf
End Function
'-----
Private Function byte2str ( bin )
Dim i, bytebuf
for i = 1 to LenB(bin)
    bytebuf = bytebuf & Chr(AscB(MidB(bin, i, 1)))
next
byte2str = bytebuf
End Function
'-----
End Class

##### Clase Fichero!!! #####

Class Fichero
'-----
Public Nombre
Public TipoContenido
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
Public Datos
'-----
Public Property Get Tamano()
    Tamano = LenB(Datos)
End Property
'-----

Public Sub Guardar(ruta)
Dim oFSO, oFich
Dim i

if ruta = "" or Nombre = "" then Exit Sub
if Mid(ruta, Len(ruta)) <> "\" then      ""
    'añado la ultima barra a la ruta
    ruta = ruta & "\"                  ""
end if

set oFSO = Server.CreateObject("Scripting.FileSystemObject")
if not oFSO.FolderExists(ruta) then Exit Sub
set oFich = oFSO.CreateTextFile(ruta & Nombre, true)

for i = 1 to LenB(Datos)
    oFich.Write Chr(AscB(MidB(Datos, i, 1)))
next

oFich.Close
set oFSO = nothing
End Sub
'-----

Public Sub GuardarComo(nombrefichero, ruta)
Dim oFSO, oFich, i

if ruta = "" or nombrefichero = "" then Exit Sub
if Mid(ruta, Len(ruta)) <> "\" then      ""
    'añado la ultima barra a la ruta
    ruta = ruta & "\"                  ""
end if

set oFSO = Server.CreateObject("Scripting.FileSystemObject")
if not oFSO.FolderExists(ruta) then Exit Sub
set oFich = oFSO.CreateTextFile(ruta & nombrefichero, true)

for i = 1 to LenB(Datos)
    oFich.Write Chr(AscB(MidB(Datos, i, 1)))
next

oFich.Close
set oFSO = nothing

End Sub
'-----

Public Sub GuardarBD (byRef field)
if LenB(Datos) = 0 then Exit Sub
field.AppendChunk Datos
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

End Sub

End Class

'-----

%>

Anexo 3:

Algoritmo de búsqueda de documentos

(Cálculo de la Función de Evaluación)

Función Sacarterminos

```
<%  
function Sacarterminos(cadena1)  
dim I  
Dim array_texto  
dim SacarAux1  
while InStr(cadena1," ")  
    cadena1=Replace(cadena1," "," ")  
wend  
  
'Lo que exista con + reemplazamos por espacio en blanco, para efectos de sacar los terminos  
while InStr(cadena1,"+")  
    cadena1=Replace(cadena1,"+"," ")  
wend  
  
'Colocamos en el arreglo todas los terminos uno a uno  
array_texto = split(cadena1," ")  
  
Sacarterminos=array_texto  
end function  
  
function GeneraSql(cadena,campos)  
if len(cadena)>0 then  
    generaSql="Select * from V_DOC_TIPO Where " & Sacar(cadena,campos) & " and doc_aux!=0  
and tipodoc_id in (select tipodoc_id from derechos_usuario where der_tipo IN ('RW','R') and  
usr_id = "&strUsuario&"") order by doc_aux desc"  
else  
    Response.Write "No hay criterios"  
end if  
end function  
%>
```

Procedimiento Almacenado SP_CALC_C

```
CREATE PROCEDURE SP_CALC_C  
(@PALABRA1 nvarchar(20),  
@PALABRA2 nvarchar(20),  
@PALABRA3 nvarchar(20),  
@NUM int)  
  
AS  
  
-- CALCULAMOS LOS COEFICIENTES SEGUN EL CAMPO DONDE ESTA LA PALABRA  
IF @NUM = 1
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
BEGIN
SELECT
  D.DOC_ID,
  C1=CASE
    WHEN D.DOC_RESU LIKE @PALABRA1 THEN 1
  ELSE 0
  END,
  C2=CASE
    WHEN D.DOC_TITL LIKE @PALABRA1 THEN 2
  ELSE 0
  END,
  C3=CASE
    WHEN D.TIPODOC_NOM LIKE @PALABRA1 THEN 4
  ELSE 0
  END,
  C4=CASE
    WHEN D.DOC_PLCL LIKE @PALABRA1 THEN 8
  ELSE 0
  END
INTO #TEMP
FROM
V_DOC_TIPO D;
END

IF @NUM = 2
BEGIN
  SELECT
    D.DOC_ID,
    C1=CASE
      WHEN D.DOC_RESU LIKE @PALABRA1 OR D.DOC_RESU LIKE @PALABRA2
    THEN 1
    ELSE 0
    END,
    C2=CASE
      WHEN D.DOC_TITL LIKE @PALABRA1 OR D.DOC_TITL LIKE @PALABRA2 THEN
    2
    ELSE 0
    END,
    C3=CASE
      WHEN D.TIPODOC_NOM LIKE @PALABRA1 OR D.TIPODOC_NOM LIKE
    @PALABRA2 THEN 4
    ELSE 0
    END,
    C4=CASE
      WHEN D.DOC_PLCL LIKE @PALABRA1 OR D.DOC_PLCL LIKE @PALABRA2 THEN
    8
    ELSE 0
    END
  INTO #TEMP2
  FROM
  V_DOC_TIPO D;
END

IF @NUM = 3
BEGIN
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
SELECT
  D.DOC_ID,
  C1=CASE
    WHEN D.DOC_RESU LIKE @PALABRA1 OR D.DOC_RESU LIKE @PALABRA2 OR
D.DOC_RESU LIKE @PALABRA3 THEN 1
    ELSE 0
  END,
  C2=CASE
    WHEN D.DOC_TITL LIKE @PALABRA1 OR D.DOC_TITL LIKE @PALABRA2 OR
D.DOC_TITL LIKE @PALABRA3 THEN 2
    ELSE 0
  END,
  C3=CASE
    WHEN D.TIPODOC_NOM LIKE @PALABRA1 OR D.TIPODOC_NOM LIKE
@PALABRA2 OR D.TIPODOC_NOM LIKE @PALABRA3 THEN 4
    ELSE 0
  END,
  C4=CASE
    WHEN D.DOC_PLCL LIKE @PALABRA1 OR D.DOC_PLCL LIKE @PALABRA2 OR
D.DOC_PLCL LIKE @PALABRA3 THEN 8
    ELSE 0
  END
INTO #TEMP3
FROM
V_DOC_TIPO D;
END

EXECUTE SP_UPD_SUMA @NUM
GO
```

Procedimiento Almacenado SP_UPD_SUMA

```
CREATE PROCEDURE SP_UPD_SUMA @NUM int

AS

DECLARE @Docu int,@Suma int

-- CURSOR PARA RECORRER LA TABLA QUE TIENE LA SUMA DE LOS COEF POR
DOC_ID
IF @NUM=1
BEGIN
  DECLARE DOC_CURSOR CURSOR FOR

  SELECT DOC_ID,SUM(C1+C2+C3+C4) AS SUMA FROM #TEMP
  GROUP BY DOC_ID
  ORDER BY DOC_ID

  OPEN DOC_CURSOR

  FETCH NEXT FROM DOC_CURSOR INTO @Docu ,@Suma

  WHILE @@FETCH_STATUS = 0 BEGIN
  -- UPDATE CON EL VALOR TOTAL DE SUMA
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
EXECUTE SP_UPD_DOC_AUX @Suma,@Docu

FETCH NEXT FROM DOC_CURSOR
INTO @Docu,@Suma
END
-- CERRAMOS CURSORES
CLOSE DOC_CURSOR
DEALLOCATE DOC_CURSOR
-- BORRAMOS TABLA TEMPORAL
DROP TABLE #TEMP
END

IF @NUM=2
BEGIN
DECLARE DOC_CURSOR CURSOR FOR

SELECT DOC_ID,SUM(C1+C2+C3+C4) AS SUMA FROM #TEMP2
GROUP BY DOC_ID
ORDER BY DOC_ID

OPEN DOC_CURSOR

FETCH NEXT FROM DOC_CURSOR INTO @Docu ,@Suma

WHILE @@FETCH_STATUS = 0 BEGIN
-- UPDATE CON EL VALOR TOTAL DE SUMA
EXECUTE SP_UPD_DOC_AUX @Suma,@Docu

FETCH NEXT FROM DOC_CURSOR
INTO @Docu,@Suma
END
-- CERRAMOS CURSORES
CLOSE DOC_CURSOR
DEALLOCATE DOC_CURSOR
-- BORRAMOS TABLA TEMPORAL
DROP TABLE #TEMP2
END

IF @NUM=3
BEGIN
DECLARE DOC_CURSOR CURSOR FOR

SELECT DOC_ID,SUM(C1+C2+C3+C4) AS SUMA FROM #TEMP3
GROUP BY DOC_ID
ORDER BY DOC_ID

OPEN DOC_CURSOR

FETCH NEXT FROM DOC_CURSOR INTO @Docu ,@Suma

WHILE @@FETCH_STATUS = 0 BEGIN
-- UPDATE CON EL VALOR TOTAL DE SUMA
EXECUTE SP_UPD_DOC_AUX @Suma,@Docu

FETCH NEXT FROM DOC_CURSOR
```

```
INTO @Docu,@Suma
END
-- CERRAMOS CURSORES
CLOSE DOC_CURSOR
DEALLOCATE DOC_CURSOR
-- BORRAMOS TABLA TEMPORAL
DROP TABLE #TEMP3
END
GO
```

Procedimiento Almacenado SP_CALC_C_AND

```
CREATE PROCEDURE SP_CALC_C_AND
(@PALABRA1 nvarchar(20),
@PALABRA2 nvarchar(20),
@PALABRA3 nvarchar(20),
@NUM int)
AS
-- CALCULAMOS LOS COEFICIENTES SEGUN EL CAMPO DONDE ESTA LA PALABRA
IF @NUM = 1
BEGIN
SELECT
D.DOC_ID,
C1=CASE
WHEN D.DOC_RESU LIKE @PALABRA1 THEN 1
ELSE 0
END,
C2=CASE
WHEN D.DOC_TITL LIKE @PALABRA1 THEN 2
ELSE 0
END,
C3=CASE
WHEN D.TIPODOC_NOM LIKE @PALABRA1 THEN 4
ELSE 0
END,
C4=CASE
WHEN D.DOC_PLCL LIKE @PALABRA1 THEN 8
ELSE 0
END
INTO #TEMP
FROM
V_DOC_TIPO D;
END
IF @NUM = 2
BEGIN
SELECT
D.DOC_ID,
C1=CASE
WHEN D.DOC_RESU LIKE @PALABRA1 AND D.DOC_RESU LIKE @PALABRA2
THEN 1
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
ELSE 0
END,
C2=CASE
WHEN D.DOC_TITL LIKE @PALABRA1 AND D.DOC_TITL LIKE @PALABRA2
THEN 2
ELSE 0
END,
C3=CASE
WHEN D.TIPODOC_NOM LIKE @PALABRA1 AND D.TIPODOC_NOM LIKE
@PALABRA2 THEN 4
ELSE 0
END,
C4=CASE
WHEN D.DOC_PLCL LIKE @PALABRA1 AND D.DOC_PLCL LIKE @PALABRA2
THEN 8
ELSE 0
END
INTO #TEMP2
FROM
V_DOC_TIPO D;
END

IF @NUM = 3
BEGIN
SELECT
D.DOC_ID,
C1=CASE
WHEN D.DOC_RESU LIKE @PALABRA1 AND D.DOC_RESU LIKE @PALABRA2
AND D.DOC_RESU LIKE @PALABRA3 THEN 1
ELSE 0
END,
C2=CASE
WHEN D.DOC_TITL LIKE @PALABRA1 AND D.DOC_TITL LIKE @PALABRA2 AND
D.DOC_TITL LIKE @PALABRA3 THEN 2
ELSE 0
END,
C3=CASE
WHEN D.TIPODOC_NOM LIKE @PALABRA1 AND D.TIPODOC_NOM LIKE
@PALABRA2 AND D.TIPODOC_NOM LIKE @PALABRA3 THEN 4
ELSE 0
END,
C4=CASE
WHEN D.DOC_PLCL LIKE @PALABRA1 AND D.DOC_PLCL LIKE @PALABRA2
AND D.DOC_PLCL LIKE @PALABRA3 THEN 8
ELSE 0
END
INTO #TEMP3
FROM
V_DOC_TIPO D;
END

EXECUTE SP_UPD_SUMA @NUM

GO
```

Procedimiento Almacenado SP_UPD_CERO

```
CREATE PROCEDURE SP_UPD_CERO AS
```

```
UPDATE DOCUMENTO SET DOC_AUX=0  
GO
```

```
CREATE PROCEDURE SP_UPD_DOC_AUX @Suma smallint, @Doc int  
AS
```

```
-- ACTUALIZAMOS EL VALOR AUX CON EL VALOR DE SUMA DE COEF.  
UPDATE DOCUMENTO SET DOC_AUX = DOC_AUX+@Suma WHERE DOC_ID=@Doc  
GO
```

Vista V_DOC_TIPO

```
CREATE VIEW dbo.V_DOC_TIPO  
AS
```

```
SELECT D.DOC_ID, D.DOC_TITL, D.DOC_RESU, D.DOC_PLCL, D.DOC_ARCH,  
D.DOC_RUTA, D.DOC_FECH, D.DOC_TAM, D.USR_ID,  
T.TIPODOC_ID,T.TIPODOC_NOM,  
D.MED_ID, D.EST_ID, D.DOC_AUX  
FROM dbo.DOCUMENTO D INNER JOIN  
dbo.TIPODOC T ON D.TIPODOC_ID = T.TIPODOC_ID
```

Llamada desde la interfaz de búsqueda

```
Dim Conexion,Tabla,miTabla  
Dim I,J,sum
```

```
Dim Temp  
Dim a  
strUsuario = Ucase(session("Usuario"))
```

```
'Asignamos los terminos a buscar a la variable cadena  
cadena2=Ucase(Request.form("palabra"))  
cadena_def=Ucase(Request.form("palabra"))
```

```
'Si no se ingreso nada retorne todos los documentos  
if Ucase(Request.form("palabra"))="" then  
Temp="Select * from V_DOC_TIPO Where tipodoc_id in (select tipodoc_id from  
derechos_usuario where der_tipo IN ('RW', 'R') and usr_id = '&strUsuario&') order by doc_id"
```

```
else  
'Verificamos si la cadena tiene blancos o +, y si tiene blancos y + lo reemplaza por blancos
```

```
'Generamos la lista de terminos que tiene la cadena  
Dim array  
array =Sacarterminos(cadena2)
```

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

```
'Limpiamos los valores doc_aux
Set Conect = Server.CreateObject("ADODB.Connection")
conect.Open "dsn=DOCDB;UID=DBUSER;PWD=ecuadordb"
'Ejecutas el store procedure para actualizar doc_aux a cero
sp0="SP_UPD_CERO "
conect.execute sp0

'Cerramos los objetos que usamos
Conect.Close
set Conect = nothing

'Creamos la conexion para el store procedure que calculara la F. evaluacion
Set Conec = Server.CreateObject("ADODB.Connection")
conec.Open "dsn=DOCDB;UID=DBUSER;PWD=ecuadordb"
'Si la cadena tiene espacio en blanco o + tomara caminos diferentes

if InStr(cadena_def," ")>0 then
    strOp=""
else
    strOp="_AND"
End if

'response.Write "Es:" & strOp & "."
if UBound(array)=0 then
    sp1="SP_CALC_C"&strOp&" ("&array(0)&"',",",1)"
    conec.execute sp1
    Conec.Close
    set Conec = nothing
End if

if UBound(array)=1 then
    sp1="SP_CALC_C"&strOp&" ("&array(0)&"', '&array(1)&"',",2)"
    conec.execute sp1
    Conec.Close
    set Conec = nothing
End if

if UBound(array)=2 then
    sp1="SP_CALC_C"&strOp&" ("&array(0)&"', '&array(1)&"', '&array(2)&"',3)"
    conec.execute sp1
    Conec.Close
    set Conec = nothing
End if

'Generamos el query que distingue los documentos encontrados y los filtra por los derechos del
usuario logeado
Temp="Select * from V_DOC_TIPO Where doc_aux!=0 and tipodoc_id in (select tipodoc_id
from derechos_usuario where der_tipo IN ('RW', 'R') and usr_id = '"&strUsuario&"') order by
doc_aux desc"

End if
```

BIBLIOGRAFÍA

LIBROS

- ÉTHIER Guylène, “*Introducción al manejo de documentos*”, Editorial Trillas, México, 2000
- DOSSETT Patti, “*Handbook of special librarianship and information work*”, Editorial Aslib, 1992.
- KAUFFMAN John, “*Beginning ASP databases*”, Wrox Press, 1999.
- GUAJUN Lu, “*Multimedia Database Management*”, Artech House Boston, 1999

SITIOS WEB

- “*Sistemas Informáticos de Tratamiento y Recuperación de Información Documental*”, <http://www.tramullas.com/nautica/documatica/4-1.html>
- “*La digitalización de documentos*”, <http://www.puj.edu.co/vice/academica/documentos/digitalizacion.ppt>
- “*SGBDD-Mini-COMO*”, <http://imigueldiaz.eresmas.com/SGBDD-Mini-COMO.htm>
- “*Gestor de Bases de Datos Documentales*” <http://www.ucm.es/info/multidoc/multidoc/revista/cuadern5/ceseda.htm>
- “*No Paper Archivo electrónico de documentos*”, <http://www.grupdes.com/prnopape.htm>

Desarrollo de prototipo de un sistema de manejo de documentos

Pontificia Universidad Católica del Ecuador

- “*Biblioteca Digital Universitaria*”,
http://www.bibliodgsca.unam.mx/tesis/tes7c1lg/sec_1.htm
- “*Docuware*”, <http://www.summan.com/Productos/Docuware/Elproblema.htm>
- “*Modern Information Retrieval*”,
<http://www.sims.berkeley.edu/~hearst/irbook/10/node3.html>
- “*El escáner y su uso en la digitalización*”,
<http://biblioweb.dgsca.unam.mx/bd/dig/escaner.html>
- “*Cursillo de introducción a la informática musical*”,
<http://www.salleurl.edu/~is04069/Modders/Docs/sampledoc.html>
- “*ASP Facil*”, <http://www.aspfacil.com>
- “*ASP Tutor*”, <http://www.asptutor.com>
- “*Programación en castellano*”, <http://programacion.com>
- “*Centros de documentación audiovisual*”,
<http://www.geocities.com/crachilecl/docaudio.htm>
- “*Restauración Digital del Sonido*”,
<http://www.arrakis.es/~fonotron/dsr.html>
- “*Canal Sur TV Archivo y Documentación*”,
http://www.canalsur.es/PROGRAMAS/television/archivo_documentacion.html
- “*Los servicios de documentación sonora ante el reto digital*”,
<http://www.ucm.es/info/multidoc/multidoc/revista/num8/marivi.html>
- “*Zyindex*”, <http://www.systemsautomation.com/zyindex.html>
- “*LaserFiche, Document Imaging for Real World*”,
<http://www.laserfiche.com/products/platform.html>
- “*Search Tools*”, <http://www.searchtools.com>