

**CENTRO DE INVESTIGACION Y DE ESTUDIOS AVANZADOS
DEL INSTITUTO POLITECNICO NACIONAL**

**UNIDAD ZACATENCO
DEPARTAMENTO DE INGENIERIA ELECTRICA
SECCION COMPUTACION**

SGD-Web: Sistema de Gestión de Documentos en la Web

Tesis que presenta la
Lic. Sais Flores Hernández

Para obtener el grado de
Maestro en Ciencias

En la especialidad de
Ingeniería Eléctrica con Opción en Computación

Director de tesis:
Dr. José Oscar Olmedo Aguirre

México, D.F., mayo del 2004

Resumen

Un sistema de gestión documental es parte del sistema empresarial de información desarrollado con el propósito de almacenar, recuperar, modificar e intercambiar documentos. Los sistemas de este tipo se han diseñado para coordinar y controlar todas aquellas funciones y actividades específicas que afectan la creación, recepción, almacenamiento, acceso y preservación de documentos, salvaguardando sus características estructurales y contextuales, y garantizando su autenticidad e integridad.

La tecnología *Web* aporta elementos que permiten gestionar documentos, pero desafortunadamente no proporciona los medios para definir el orden de ejecución, la sincronización y el intercambio de información entre los procesos administrativos de un sistema de gestión documental. Por tal razón, es necesario contar con mecanismos que permitan diseñar y construir procesos administrativos con reglas bien definidas que permitan integrar recursos humanos y tecnológicos para el trabajo en grupo.

El sistema SGD-Web desarrollado en ésta tesis proporciona los medios necesarios para conducir la gestión de documentos que requiere la colaboración de varios participantes. Las reglas de colaboración establecen los documentos que se deben transferir entre los repositorios de documentos de los participantes. Los repositorios se han implementado como servidores WebDAV, una extensión de HTTP que hace de la *Web* un medio capaz de almacenar documentos. SGD-Web facilita la colaboración entre los participantes siguiendo la metáfora del escritorio de oficina organizado por bandejas de documentos. El sistema SGD-Web ha sido probado con una aplicación que define el proceso de inscripción de un alumno a un programa de estudio. El proceso de inscripción involucra a un grupo de trabajo que participa con la aprobación del documento de inscripción.

Las principales aportaciones de este trabajo consisten en implementar las operaciones básicas de un sistema de *workflow* orientado a documentos y en facilitar la conducción de actividades para el intercambio de documentos entre los participantes. La importancia de las contribuciones consiste en aumentar la productividad al simplificar el desarrollo de un gran número de procesos administrativos de gestión documental que son comunes en departamentos y corporaciones.

Abstract

A document management system is part of the enterprise information system developed in order to store, recover, modify and exchange documents. Systems of this type have been designed to coordinate and to control the specific functions and activities that affect the creation, reception, storage, access and preservation of documents, preserving their structural and contextual characteristics, and guaranteeing their authenticity and integrity.

The Web technology provides elements that allow to manage documents, but unfortunately it does not provide the techniques to define order of execution, synchronization and exchange of information between the administrative processes of document management system. For such a reason, it is necessary to consider mechanisms that allow to design and to generate administrative processes with well defined rules that allow to integrate human and technological resources for workgroup.

The SGD-Web system developed in this thesis provides the essential means to coordinate the management of documents that requires the collaboration of among participants. The collaboration rules determine the documents to be transferred among the repositories of documents. The repositories have been implemented as WebDAV servers, an extension of HTTP that makes the Web able to store documents. The paradigm SGD-Web makes collaboration among the participants easy, by organizing trays of documents and providing the suitable facilities following the desktop metaphor. The SGD-Web system has been tested with an application for the registry of a student into a training program. The registry involves a work group that participates with the approval of the registry document.

The main contributions of this thesis consist of implementing the basic operations of a workflow system oriented to documents and in providing the sequence of activities for document exchange among the participants. The relevance of the contributions consists of increasing the productivity, in a way that simplifies the development of a great number of administrative processes of document management that are common in departments and corporations.

Contenido

Capítulo 1	Introducción.....	1
1.1	Área de estudio	1
1.2	Motivación.....	2
1.3	Antecedentes	2
1.4	Objetivo.....	5
1.5	Alcance.....	5
1.6	Organización de la tesis	6
Capítulo 2	Gestión documental	9
2.1	Aspectos generales.....	9
2.2	Características	11
2.3	Workflow y su correlación con gestión documental.....	13
2.4	World Wide Web y la gestión documental	21
2.5	Sinopsis.....	25
Capítulo 3	Protocolo WebDAV	27
3.1	Conceptos básicos de WebDAV	27
3.2	Requerimientos para colaboración	30
3.2.1	<i>Soporte para todo tipo de contenido.....</i>	<i>30</i>
3.2.2	<i>Soporte de control de concurrencia.....</i>	<i>30</i>
3.2.3	<i>Soporte para propiedades.....</i>	<i>31</i>
3.2.4	<i>Soporte para vínculos independientes del tipo de contenido.....</i>	<i>32</i>
3.2.5	<i>Recuperación para edición de fuente no procesada</i>	<i>32</i>
3.2.6	<i>Manipulación de espacio de nombre</i>	<i>33</i>
3.2.7	<i>Soporte para colecciones</i>	<i>33</i>
3.3	Características	33
3.3.1	<i>Propiedades</i>	<i>34</i>
3.3.2	<i>Bloqueo</i>	<i>35</i>
3.3.3	<i>Manejo de espacio de nombre</i>	<i>37</i>
3.4	Métodos.....	38
3.5	Implementaciones	45
3.5.1	<i>Servidores</i>	<i>45</i>
3.5.2	<i>Clientes</i>	<i>46</i>
3.6	Sinopsis.....	46
Capítulo 4	Soluciones actuales.....	49
4.2	Sistemas de automatización de workflow con gestión documental.....	49
4.3	BSCW system (<i>Basic Support for Cooperative Work</i>).....	50
4.4	DreSS	52
4.5	Sinopsis.....	53
Capítulo 5	Caso de estudio: inscripción escolar	55
5.1	Descripción del proceso	55
5.2	Identificación de entidades de workflow en el proceso	56

5.3	La gestión de documentos dentro del proceso.....	60
5.4	Discusión	62
5.5	Sinopsis.....	63
Capítulo 6 Diseño e implementación de SGD-Web		65
6.1	Diseño de la solución	65
6.2	Implementación de la solución.....	70
6.2.1	<i>Arquitectura de la aplicación.....</i>	71
6.2.2	<i>Creación de estructura de directorios y almacenamiento del documento.....</i>	72
6.2.3	<i>Movimiento del documento</i>	74
6.2.4	<i>Descripción de clases</i>	77
6.4.	Sinopsis.....	81
Capítulo 7 Uso del sistema: SGD-Web		83
7.1.	Requerimientos del sistema	83
7.2.	El Servidor <i>Web</i>	83
7.3.	Uso del sistema.....	85
7.4.1.	<i>Menú del alumno.....</i>	86
7.4.2.	<i>Tutor de estudios</i>	92
7.4.3.	<i>Coordinador académico.....</i>	94
7.4.4.	<i>Secretaría.....</i>	97
7.4.5.	<i>Servicios escolares.....</i>	98
7.5.	Sinopsis.....	102
Capítulo 8 Conclusiones		105
8.1.	Trabajo futuro.....	107
Referencias		109

Lista de Figuras

Figura 2. 1: Operaciones con los documentos	11
Figura 2. 2: Relación entre las entidades en un workflow	15
Figura 2. 3: Operación de la etiqueta <a>	25
Figura 5. 1: Flujo de información en el proceso de inscripción.....	57
Figura 5. 2: Meta modelo de la definición del proceso	57
Figura 6. 1: Funcionalidad y estructura del repositorio de un proyecto basado en WebDAV	67
Figura 6. 2: Estructura de directorio creada para un participante.....	67
Figura 6. 3: Diseño de la simulación del escritorio del participante	68
Figura 6. 4: Movimiento del documento de un participante a otro.	69
Figura 6. 5: Manipulación de los recursos.....	70
Figura 6. 6: Arquitectura de la aplicación.	72
Figura 6. 7: Miembros de una colección.....	73
Figura 6. 8: Diagrama de clases.....	78
Figura 7. 1. Directorio de publicación de Tomcat.....	84
Figura 7. 2. Opción que inicializa el servicio de Tomcat.....	85
Figura 7. 3. Ventana que indica el servicio inicializado de Tomcat.....	85
Figura 7. 4. Página de inicio del sistema.	86
Figura 7. 5. Menú del alumno.	87
Figura 7. 6. Pantalla donde introducirá el número de matrícula el alumno	88
Figura 7. 7. Datos del documento de inscripción.	88
Figura 7. 8. Notificación de creación de documento.....	89
Figura 7. 9. Documentos a enviar por el alumno a su tutor de estudios.	89
Figura 7. 10. Escritorio del alumno, sin documentos disponibles.	90
Figura 7. 11. Documentos rechazados al alumno.	91
Figura 7. 12. Muestra el documento en el expediente del alumno.....	91
Figura 7. 13. Menú del tutor de estudios.....	92
Figura 7. 14. Escritorio del tutor de estudios. autoriza o rechaza documentos.....	93
Figura 7. 15. Escritorio del tutor de estudios, después de autorizar un documento.....	93
Figura 7. 16. Documentos rechazados al tutor de estudios.	94
Figura 7. 17. Menú del coordinador académico.....	95
Figura 7. 18. Documentos a ser revisados por el coordinador académico.	96
Figura 7. 19. Escritorio del coordinador académico, sin documentos a revisar.	96
Figura 7. 20. Documentos rechazados al coordinador académico.	97
Figura 7. 21. Menú de la secretaria.	98
Figura 7. 22. Escritorio de la secretaria.....	99
Figura 7. 23. Documentos rechazados al coordinador académico.	99

Figura 7. 24. Documentos aprobados. La secretaria envía al expediente del alumno.....	100
Figura 7. 25. Menú de servicios escolares.	101
Figura 7. 26. Escritorio de servicios escolares.	101

Capítulo 1 Introducción

1.1 Área de estudio

La *World Wide Web* (*Web*), bien conocida por ser la aplicación de mayor trascendencia de Internet, ha tenido un crecimiento muy rápido en los últimos años. Permite el acceso a documentos multimedia, enlazados entre sí, los cuales se almacenan en servidores dispersos alrededor del mundo. La disponibilidad ilimitada de información, en cualquier momento y en cualquier lugar, es una realidad.

Entre las ventajas que podemos obtener con el uso de la *Web* en la empresa se encuentran el mejoramiento de la organización, recuperación y publicación de información así como de operaciones electrónicas (venta electrónica, transacciones electrónicas, etc.). Si se aprovecha las ventajas que la *Web* tiene para el funcionamiento interno de la empresa, hablamos específicamente de una Intranet.

Para definir los objetivos de la empresa así como su funcionamiento interno, la creación y administración de los documentos que intervienen deben regirse por las políticas y procedimientos de la empresa.

El manejo de la información de la empresa a través de documentos, involucra la gestión de los mismos, es decir, las operaciones que se pueden hacer con documentos: la obtención del mismo, el almacenamiento, la publicación, la revisión y aceptación por diferentes miembros de la empresa, para lo cual es necesario contar con un sistema de gestión documental.

Un sistema de gestión documental es parte del sistema empresarial de información desarrollado con el propósito de almacenar y recuperar documentos. El sistema debe estar diseñado para coordinar y controlar todas aquellas funciones y actividades específicas que afectan la creación, recepción, almacenamiento, acceso y preservación de los documentos, salvaguardando sus características estructurales y contextuales, garantizando su autenticidad e integridad.

La tecnología *Web* aporta muchos elementos que permiten gestionar documentos más eficientemente pero, desafortunadamente, no proporciona la estructura organizativa ni conceptual necesaria para conducir el sistema de gestión documental. Por tal razón, es necesario contar con procesos diseñados bajo reglas

bien establecidas que involucren recursos humanos o tecnológicos en una forma de trabajo en grupo conocido como *workflow*. Un *workflow* define también el orden de invocación de las tareas así como su sincronización y el flujo de información entre ellas.

1.2 Motivación

Gestión documental e Internet, el propósito y el medio, son dos conceptos estrechamente relacionados. La utilización de entornos Internet en las empresas ofrece grandes oportunidades para la implantación de una gestión documental eficiente, pero siempre que se haya realizado un esfuerzo previo en el diseño y planificación del sistema, acorde con los objetivos y necesidades de la empresa. Sin embargo no se debe limitar un sistema de gestión documental local de ámbito empresarial, es decir, no se limitará a una Intranet ya que es posible exportar e importar información de y hacia otras fuentes incluyendo otras corporaciones.

Hoy en día el almacenamiento físico de archivos y documentos en papel es cada vez más difícil debido a la carencia de espacio en los lugares de trabajo. Grandes volúmenes de papel se van acumulando, haciendo problemático el acceso fácil y rápido a la información requerida. Esto causa pérdida de tiempo y de dinero en insumos para implementar lugares apropiados de archivo.

Por ésta razón es valioso contar con un sistema que administre la información contenida en los documentos, es decir, con un sistema de gestión documental que integre las estrategias de las diferentes áreas de la empresa, logrando mayor eficiencia y eficacia en el servicio a través de una buena comunicación y difusión de la información. El desarrollo de tal sistema de gestión documental usando la infraestructura de la *Web* es el propósito de éste trabajo.

1.3 Antecedentes

Es indudable que los frutos de la colaboración son mayores que aquellos conseguidos por el trabajo individual, una cuestión que es especialmente importante cuando se trata de realizar tareas dentro de una organización. Gracias al software para el trabajo en grupo es factible la mejora del rendimiento general de todo el proceso productivo y su mayor aportación es hacer posible que diferentes personas puedan trabajar compartiendo la misma información y cooperar estrechamente en el desarrollo de proyectos. Estas herramientas no fueron diseñadas para un entorno

tan amplio como el alcance mundial que ofrece la *Web* sino más bien para las *Intranets* corporativas. Dentro de este contexto, las funciones básicas que se consiguen llevar a cabo son las siguientes:

1. Ayudar a que dos o más personas trabajen juntas.
2. Compartir conocimientos y experiencias.
3. Automatizar sus actividades.
4. Ayudar a crear una memoria de la organización.
5. Enlazar grupos de trabajo dispersos geográfica y temporalmente.

A estas herramientas se les denomina colectivamente como *groupware*. El *groupware* es una herramienta poderosa que permite compartir toda clase de conocimiento relativo a una organización y facilita el movimiento y control de la información que se manipula constantemente. Los módulos que integran son, principalmente:

- Calendario y planificación.
- Videoconferencia.
- Sistemas de reunión electrónica.
- Pizarra electrónica y conferencia de datos.
- Conversación (*chat*).
- Correo electrónico.
- Conferencia y grupos de noticias.
- Almacenes de conocimiento.
- Escritura en grupo y edición compartida.
- Flujo de trabajo (*workflow*).

El *groupware* aporta soluciones para cada uno de los momentos o situaciones que puedan producirse durante la evolución de esta colaboración. De tal forma, integran sistemas que facilitan su realización, coordinación y comunicación de eventos, noticias, nuevas ideas, etc. Los últimos dos puntos son fundamentales cuando se trata de compartir operaciones y documentos, por lo que su implementación es esencial. El intercambio de información, su gestión y control, la notificación de toda clase de actualizaciones, utilización del conocimiento almacenado referente a la organización, etc., todo ello constituye el núcleo central del trabajo en los sistemas *groupware*.

La mayor parte de estos modelos incorporan funciones *workflow* para la automatización de las fases y tareas del trabajo en grupo. Mediante esta técnica es posible dirigir la información y controlar automáticamente su flujo a través de todos los niveles organizativos de una empresa o institución por medio de grupos de

reglas y formularios. *Workflow* y *groupware* trabajan conjuntamente para la consecución de objetivos comunes y, aunque ambos conceptos son diferentes, se complementan a la hora de ofrecer sus servicios. De cualquier forma, cada uno de ellos ha sufrido un desarrollo paralelo, y si bien en sus inicios se consideraron parte de un mismo concepto, su evolución independiente ha propiciado su diferenciación.

Este proceso necesita de la interacción del participante con el sistema, debido a la creación, revisión que debe hacer del documento. Llegando al fin del proceso cuando todos los participantes han llegado a un acuerdo y en conjunto contribuyen a la creación del documento, ya que la información se envió de un participante a otro. En este caso hablamos de un *workflow* de colaboración continuo. El *workflow* se dice que es de colaboración porque diferentes personas colaboran para lograr el objetivo del proceso, la inscripción del alumno, y se dice que es continuo porque la información no se remite a otro participante hasta que el participante que actualmente tiene el documento no finaliza su actividad. A diferencia de un *workflow* de producción, en donde el objetivo es administrar un gran número de tareas similares y optimizar su productividad mediante la automatización de tareas tan simples como sean posibles, la intervención de un usuario es mínima, solo para administrar excepciones. El *workflow* de colaboración no se enfoca en automatizar un gran número de tareas, sino en los objetivos que tienen en común un grupo de personas, las cuales interactúan entre sí y con el sistema mismo, sistema que es muy haga uso de la *Web*.

Actualmente la comunicación entre los participantes de un proceso, se puede llevar a cabo comenzando por la manera más simple y tradicional como la comunicación oral (como teléfono), el ya conocido e-mail, hasta llegar a la videoconferencia o a un sistema de colaboración en la *Web*. Estos sistemas de colaboración en la *Web* han sido implementados utilizando diferentes formas de comunicación. Algunos hacen uso de mensajería para poder enviar la información de un participante a otro, acceder a un repositorio ubicado en un lugar específico para tomar de ahí la información y poder realizar sus tareas correspondientes. Las herramientas clásicas como el correo electrónico y las listas de distribución no son, sin embargo, capaces de ofrecer mayor soporte para realizar trabajo en grupo a través de la *Web*. En este punto es donde se están desarrollando las aplicaciones *groupware* para promover la colaboración a través de la *Web*.

Entre estas aplicaciones podemos encontrar BSCW que ofrece un soporte idóneo para la creación de zonas en donde se pueden compartir documentos comunes. RedIRIS ofrece BSCW, entre otros servicios, para su aplicación al colectivo de Comunidades Virtuales Científicas (CVCs) y desarrollado por el GMD-FIT

(German National Research Center for Information Technology). Finalmente, DreSS es un conjunto de herramientas que permiten autores depositar y/o actualizar documentos en un servidor *Web*, sin la necesidad de autenticarse en el servidor *web*, y sin la necesidad de acceder al sistema de archivos de las computadoras. Este sistema usa un servidor *Web* como repositorio de documentos para múltiples autores y que este repositorio puede ser construido usando un navegador y servidor *web*.

En este trabajo se propone desarrollar un sistema para la gestión de documentos el cual obedece un esquema básico de actividad que consiste en mover documentos entre los repositorios de documentos de los participantes del grupo de trabajo. El trabajo usa el protocolo WebDAV el cual extiende el protocolo HTTP 1.1 definiendo nuevos métodos para hacer posible la colaboración en la *Web*. El método *move* del protocolo WebDAV permite mover documentos de un repositorio origen a un repositorio destino pero tiene el inconveniente que dichos repositorios deben estar ubicados en la misma computadora. Para remediar ésta deficiencia, en éste trabajo se proponen mecanismos alternativos para garantizar el movimiento de documentos hacia repositorios que residen en otras computadoras.

1.4 Objetivo

El objetivo de esta tesis es diseñar y construir un sistema de gestión documental con base en el protocolo WebDAV que facilite la colaboración entre grupos de trabajo mediante el intercambio de documentos electrónicos.

Para su realización este objetivo se puede descomponer en los siguientes objetivos específicos:

- a) Permitir la interoperabilidad remota, asíncrona y la publicación colaborativa de documentos.
- b) Utilizar conjuntamente los métodos *GET*, *PUT* y *DELETE* para mover documentos entre servidores habilitados con el protocolo WebDAV.
- c) Optimizar en la medida que la estructura de la institución y/o empresa lo permita, de los procesos que involucren la interrelación entre los participantes, para la gestión de documentos.

1.5 Alcance

El sistema construido en éste trabajo requiere una conexión a Internet para el envío de documentos pero una vez depositado en el repositorio, la revisión para su autorización o rechazo por el participante no depende de la conexión a Internet. Por tal razón, se realiza la implementación usando tres métodos importantes del protocolo WebDAV *Get*, *Put* y *Delete*, logrando así depositar el documento en las computadoras de los participantes, sin importar donde se encuentran ubicados geográficamente.

La implementación supone que los participantes colaboran en forma secuencial. Con ello se garantiza que aquella actividad en la que se encuentre el proceso en un momento dado será regida por el participante del cual depende que continúe o que se regrese, es decir, nadie más podrá afectar en ese momento la integridad de la información. Este tipo de procesos existe en muchas empresas y es de gran importancia contar con un sistema de este tipo para llevar control y garantizar que se cumplan las reglas establecidas y así evitar problemas de alteración de la información

La implementación del sistema se limita al mecanismo por el cual los documentos se mueven del repositorio de un participante a otro. Este sistema debe usarse con otros ya que no ofrece mecanismos para la administración de firmas electrónicas, de edición o modificación del documento, de notificación de la existencia de documentos o de detección de la finalización de un proceso.

1.6 Organización de la tesis

El documento se organiza en seis capítulos, de los cuales se da una breve descripción a continuación:

En el capítulo 2, denominado Gestión documental, se aborda la problemática actual del manejo de documentos en papel, y como se resuelve ésta a través de un sistema de gestión documental. Se mencionan las características de un sistema de este tipo y como se relaciona y complementa con la tecnología de *workflow* y de la *Web*.

El capítulo 3, titulado Protocolo WebDAV, se mencionan los requerimientos que debe cumplir un protocolo con soporte colaborativo en la *Web*, y como WebDAV da soporte a dichos requerimientos.

El capítulo 4, nombrado: Soluciones Actuales, se hace una descripción del estado

del arte de dos soluciones que han originado soluciones de automatización de *workflow*, a los procesos de negocios en diferentes áreas de la industria.

El capítulo 5, denominado: Caso de estudio, donde se realiza el planteamiento general del problema y se da una descripción del caso de estudio utilizado, se describe los participantes de este proceso, así como las tareas que cada uno realiza.

El capítulo 6 titulado: Diseño e implementación de la gestión de documentos, se describen los métodos del protocolo utilizados y clases implementadas para lograr el movimiento de documentos en la *Web*.

El capítulo 7 denominado: uso del sistema, donde se explica a través de un manual de usuario el uso del sistema y la forma como los participantes interactúan con él.

El capítulo 8 son las conclusiones, obtenidas al realizar este trabajo. Así como también se mencionan los trabajos futuros o las mejoras del mismo.

Capítulo 2 Gestión documental

La cantidad de documentos en papel que maneja una persona, en su lugar de trabajo, es inmensa, y si multiplicamos esto por el número de empleados de dicha empresa, el número de papeles manejado se incrementa considerablemente, y con esto, se incrementan los costos, así como el tiempo de búsqueda de la información, por lo cual la gestión de documentos resulta ineficaz e ineficiente.

En el presente capítulo se nombrarán los problemas que se presentan al manejar los documentos de manera manual, explicando porque resulta ineficaz e ineficiente la gestión de los mismos, así como también las operaciones que se realizan con los documentos, para comprender cada una de las características con las que debe cumplir un sistema de gestión documental automatizado, así como las ventajas que se obtienen con el mismo.

Una vez descrito un sistema de gestión documental, se explicarán las herramientas, tecnologías y los conceptos que se usan hoy en día para implementación de un sistema de gestión documental.

2.1 Aspectos generales

En una empresa se denominará *corporativo* a su entorno de trabajo más general y se distinguirá del que está restringido a un grupo o parte de ella, que será denominado entorno *departamental*. El manejo de documentos es muy grande, ya que tanto a nivel departamental como corporativo, existen procesos que se llevan a cabo por varias personas de forma conjunta y que involucran manejar o transferir varios documentos.

Lo que se busca es que corporativa y departamentalmente, la gestión documental en la empresa cumpla con sus necesidades. No obstante, muchas veces la gestión de documentos resulta ineficaz e ineficiente. La ineficacia equivale a mal servicio, es decir, está enfocada al cliente, por lo cual, al tener un sistema documental manual, llamado *papeleo*, ocasiona retrasos, pérdidas materiales, molestias y complicación en los procesos. Por otro lado la ineficiencia equivale a costes excesivos, se dirige a la forma de actuar del servidor. Por lo tanto, la gestión documental ineficiente provoca colectiva e individualmente esfuerzos excesivos y gastos elevados. La

mayoría de las corporaciones prestan atención a la eficiencia y no a la eficacia. Pero lo que es cierto es que al mejorar la eficacia y eficiencia se brinda mejor servicio a menor coste.

Hoy en día se cuenta con las tecnologías de la información y las comunicaciones las cuales proporcionan los medios para resolver el problema de gestionar documentos en una empresa. Podemos decir que la gestión documental automatizada tiene como objetivo proporcionar apoyo a la utilización de los documentos para todo tipo de procesos en la empresa, ya sea corporativa o departamental, en donde un documento puede ser cualquier tipo de objeto que contenga información en forma de datos, texto, o imagen fija.

Es bueno saber el volumen de documentos que se maneja en la empresa, así como que documentos utilizamos, pero es importante conocer que hacemos con los documentos para determinar las funciones a cumplir por un sistema de gestión documental.

Las operaciones que se realizan con los documentos son:

- **Producción.** Es todo el proceso que lleva a la creación de un documento, en nuestro lugar de trabajo, en papel o en un archivo electrónico. Las operaciones que involucra el obtener los documentos son:
 - Crear o utilizar los existentes
 - Actualizarlos
 - Buscarlos
 - Modificarlos

- **Análisis.** Su principal propósito es comprender el contenido del documento para utilizarlo de forma adecuado, así como clasificarlo y relacionarlo de manera correcta con otros documentos. Las operaciones comprendidas en el análisis son:
 - Leerlos
 - Subrayarlos
 - Marcarlos
 - Comentarlos
 - Resumirlos

- **Gestión.** La administración y el control de las actividades de una empresa se pueden llevar a cabo interpretando correctamente las indicaciones establecidas en los documentos. Las operaciones a realizar son:

- Decretarlos
 - Remitirlos
 - Controlarlos
- **Publicación.** se selecciona de los documentos la información de la empresa que se necesita y convenga divulgar al exterior, por ejemplo, en la *Web*.
 - Archivarlos
 - Distribuirlos
 - Visualizarlos
 - Imprimirlos
 - Reverenciarlos

Estas son las operaciones realizadas por los documentos a gestionar (figura 2.1), los cuales contienen la información utilizada por la empresa, en un entorno departamental o corporativo.

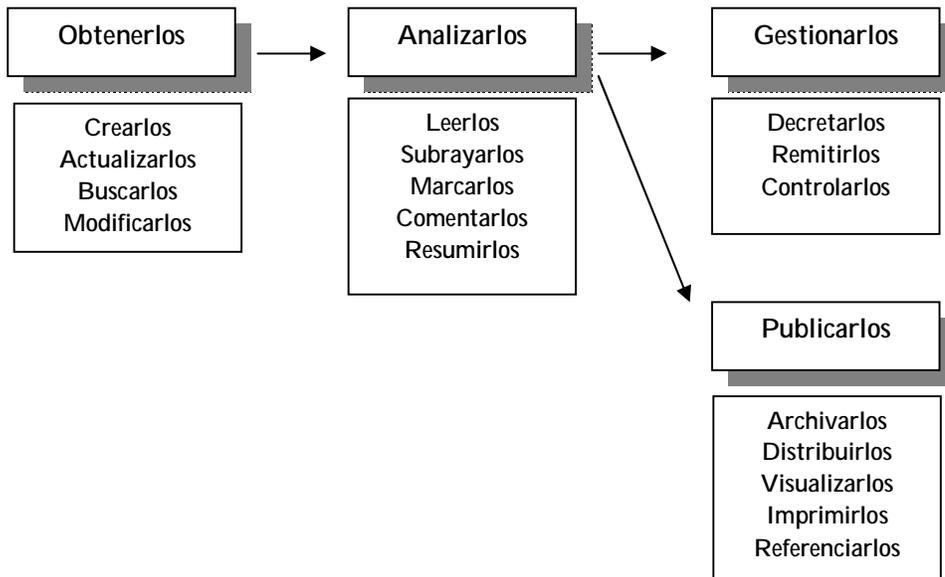


Figura 2. 1: Operaciones con los documentos

2.2 Características

Un sistema de gestión documental deberá satisfacer varios aspectos relacionados que involucran desde la información de: cómo y dónde se almacena, quiénes y para qué la utilizan. Por lo tanto podemos decir que un sistema de gestión documental debe contar con las siguientes características:

- Corporativo. La información que se maneja en los documentos, es donde se basa la funcionalidad de la empresa, así como se crean nuevas aplicaciones y procedimientos.
- Escalabilidad. La consideración de la naturaleza corporativa demandará un rápido crecimiento del sistema.
- Interoperabilidad. Debe ser integrado con subsistemas de correo electrónico, de esta forma se puede realizar una comunicación en la empresa de la información de los documentos y las operaciones que se realicen con ellos.
- Publicación. Diseminar la información interna.
- Integración. Integrar fuentes externas, como son los documentos que son enviados por terceros vía fax, CD-ROM, o papel.
- Importación. Contar con un sistema de importación masiva de documentos, para toda aquella información de archivos históricos o de grandes cantidades.
- Almacenamiento. Almacenar los documentos masivamente, que garantice su fácil acceso para cualquier usuario autorizado, así como su recuperación en un tiempo mínimo del mismo.
- Edición. Contar con herramientas que optimicen la creación de documentos, así como de metodologías que apoyen la creación y modificación de los mismos.
- Control. Para almacenar y distribuir documentos, el sistema deberá permitir el control de la comparación de las distintas versiones de los documentos, de tal forma que contemple el conflicto que se pueda dar entre el control de acceso y el control de versiones.
- Clasificación. Agrupar documentos, los cuales deberán ser clasificados por diferentes criterios y a la vez estar enlazados por medio de hipervínculos dentro de los mismos documentos y hacia otros documentos que se relacionen.
- Replicación. Debe permitir replicar los documentos entre diferentes servidores que incluso se encuentren en ubicaciones diferentes y enviar automáticamente mensajes (correo electrónico) para notificar a los usuarios de la existencia de documentos relevantes.
- Búsqueda. Así mismo tiene que contar con un mecanismo de búsqueda rápida, que permita obtener los documentos en el menor tiempo posible y de forma eficiente. El mecanismo de búsqueda deberá ser capaz de realizar

dicha operación, por distintos criterios de búsqueda, es decir, por palabras claves o descriptores.

- **Recuperación.** Al buscar un documento, y querer visualizarlo, se debe contar con la recuperación instantánea de cualquier documento, para presentarse en pantalla, y a la vez presentar en forma instantánea todos los documentos que se encuentren enlazados con el solicitado.
- **Seguridad.** Se debe garantizar la existencia y almacenamiento adecuado de los documentos, por lo cual, es importante contar con un mecanismo que garantice la seguridad para que valide al usuario que accesa la información. Además al registrar el documento, la accidental o intencionada eliminación del mismo, así como también el acceso simultáneo para su lectura, actualización.

Todas estas consideraciones son las que deben contener un sistema de gestión documental, pero para que cuente con ellas, es importante contar o apoyarse con subsistemas tales como correo electrónico, un subsistema de base de datos documental, un subsistema de OCR, subsistemas de voz y video.

Como se mencionó antes, un sistema de gestión documental permite administrar todos los documentos, desde el nivel departamental hasta el nivel corporativo. Dichos documentos son parte de un proceso de negocio o de varios procesos, procesos que involucran a personal que llevan a cabo determinadas actividades regidas por reglas. Es importante conocer que son cada uno de estas entidades y como interactúan en un sistema de gestión documental, por lo cual en la siguiente sección abordaremos conceptos de *workflow* y su correlación con el sistema de gestión documental.

2.3 Workflow y su correlación con gestión documental

El *workflow* es un proceso administrativo en el cual el flujo de trabajo es ejecutado por recursos humanos y/o aplicaciones de tecnología de la información (IT) y cuyas reglas de ejecución son marcadas por los procesos del negocio. De acuerdo al WfMC [19] un *workflow* se define como:

La automatización de procesos de negocios, total o parcial, durante el cual los documentos, información o tareas se pasan de un participante a otro, de acuerdo a un conjunto de reglas de procedimientos.

La automatización de un proceso de negocio está dada en una definición de procesos, en la cual identifica las diferentes actividades, reglas y datos usados para manejar el *workflow*. Muchas instancias del proceso se operan con un conjunto específico de datos.

El *workflow* está asociado con la reingeniería de procesos de negocios, la cual se ocupa de la valoración, análisis, modelado, definición e implementación operacional de los procesos de negocios de una organización.

Un *workflow* involucra diferentes terminologías que tienen y conservan relaciones entre sí. Esto se ilustra en la figura 2.2. Como se muestra en la figura, el inicio del proceso de negocio es la actividad a realizar. Un proceso de negocio se define como:

Un conjunto de uno o más procedimientos ligados o actividades en las cuales colectivamente realizan objetivos o metas del negocio, normalmente en el contexto de una estructura organizacional definiendo reglas funcionales y relaciones [19].

Un proceso de negocio se asocia con objetivos operacionales y relaciones de negocios. Como se muestra en la figura 2.2, un *proceso de negocio* se da en una definición de proceso, cuya definición dice:

La representación de un proceso de negocio que haga posible la manipulación automática, tal como el modelado o la representación por un sistema manejador de workflow. La definición de proceso consiste de una red de actividades y sus relaciones, criterios para indicar el comienzo y la terminación del proceso, e información acerca de las actividades individuales, tales como participantes, asociados a aplicaciones IT y datos, etc. [19].

Puede involucrar interacciones entre participantes, durante largos periodos. Puede consistir de actividades.

Un proceso generalmente consiste de muchas actividades las cuales están lógicamente relacionadas en términos de sus contribuciones para la realización del proceso de negocio. Una actividad es la unidad más pequeña de trabajo la cual se programa en un motor de *workflow* durante el proceso (es decir, usando transiciones y pre/post condiciones), aunque una actividad puede quedar asignada en diferentes trabajos (a un participante). Una actividad se clasifica como "manual" o "automatizada".

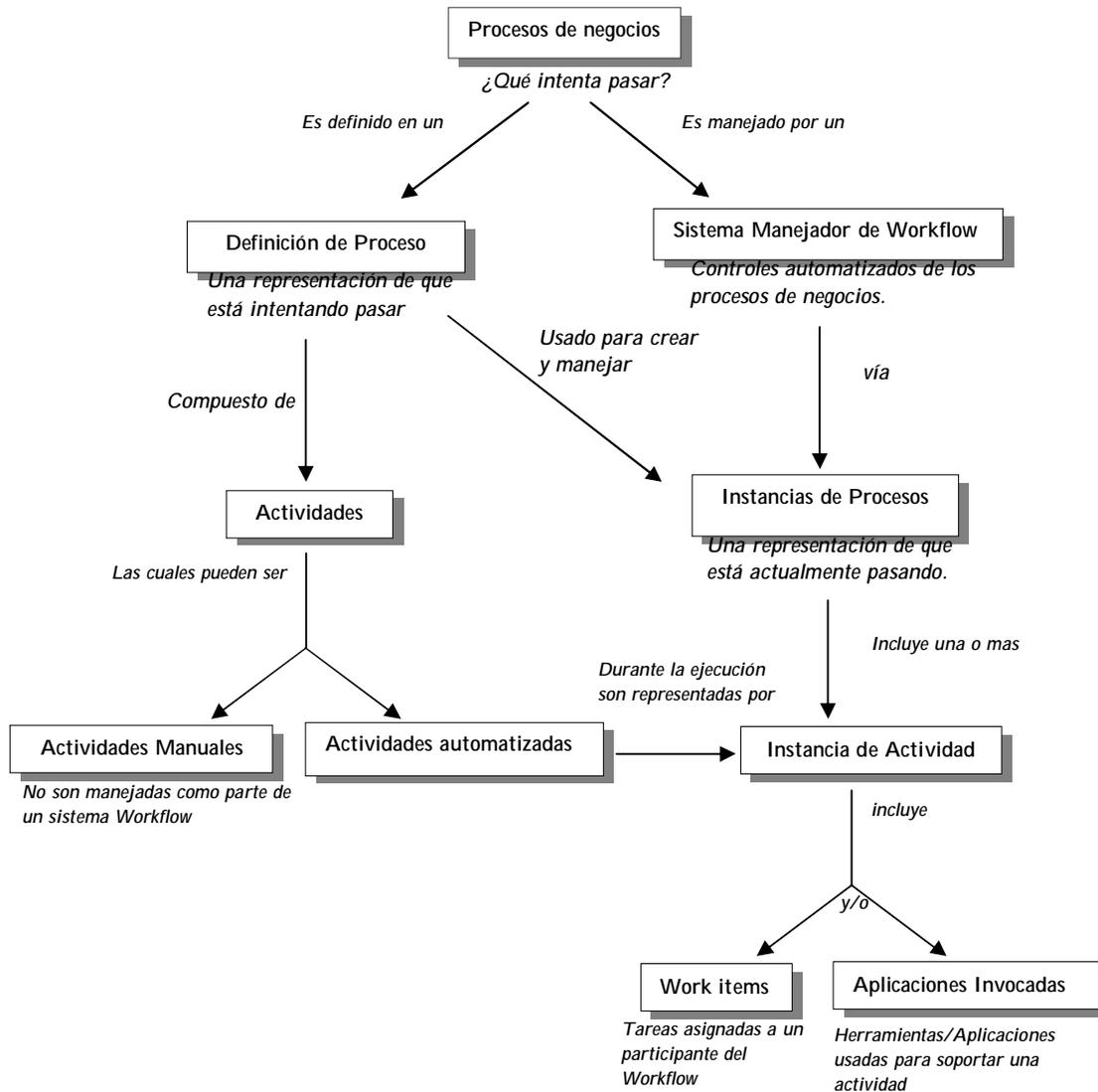


Figura 2. 2: Relación entre las entidades en un Workflow

Una actividad manual se define como, *una actividad dentro de un proceso de negocio la cual no es capaz de automatizarse y yace fuera del marco de un WMS. Tales actividades podrían ser incluidas dentro de una definición de proceso, por ejemplo, para apoyar la modelación de procesos, pero no formar parte de un workflow resultante [19].*

Por otro lado una actividad automatizada es: *una actividad capaz de automatizarse usando un sistema manejador de workflow (WMS) para conducir la actividad durante la ejecución del proceso de negocio del cual forma parte [19].*

Durante el proceso de ejecución, una actividad automatizada (o *workflow*) se maneja por el WMS. Esto podría resultar en:

- Una aplicación invocada comienza a activarse directamente por el WMS (no con los participantes).
- Uno o más trabajos comienzan asignarse a un participante del *workflow*, con soporte de herramientas o aplicaciones invocadas y manejadas por el WMS.
- Uno o más trabajos comienzan a asignarse a un participante del *workflow* para procesos independientemente del WMS, con la terminación del trabajo que se notifica al WMS por el participante (en un sistema de *workflow* esto puede algunas veces describirse como un trabajo ejecutado manualmente).

Como se muestra en las relaciones entre los conceptos básicos de un *workflow*, a partir de un proceso de negocio, se tiene una bifurcación, donde en una se define el proceso mismo en la otra se maneja el proceso de negocio, es decir, el sistema manejado de *workflow*.

Actualmente ya se cuenta con software que hace uso del concepto básico del *workflow* y su motor de desarrollo se centra en esto. Por lo cual un sistema manejador de *workflow* se define como:

Un sistema que define, crea y maneja la ejecución de workflows a través del uso de software, corriendo uno o mas motores de workflow, el cual se habilita para interpretar el proceso, interactuar con participantes, y donde se requiera, invocar el uso de herramientas y aplicaciones TI [19].

En sí un sistema manejador de *workflow* consiste de componentes de software para almacenar e interpretar definiciones de procesos, crear y manejar instancias de *workflow* para ser ejecutadas, y controlar sus interacciones con participantes y aplicaciones. Tales sistemas también proveen funciones administrativas y de supervisión, por ejemplo permite reasignación de trabajo, más auditoria y manejo de la información en el sistema global.

Algunas de las definiciones sobre sistemas de manejo del *workflow* se dan a continuación, las cuales están enfocadas en áreas de consultoría, de investigación y del sector industrial.

El software de manejo de workflow es un sistema proactivo el cual maneja el flujo de trabajo entre participantes, de acuerdo a procedimientos definidos que consisten de un número de tareas. Coordina participantes (usuarios y sistemas), junto con los recursos apropiados, los cuales pueden ser accesibles directamente por el sistema o fuera de línea, para lograr objetivos definidos. La coordinación involucra el paso de tareas de participante a participante en correcta secuencia, asegurando que todos los requerimientos y contribuciones se lleven a cabo, tomando acciones por preestablecidas cuando sea necesario [4].

En esta definición dada por Hales y Lavery, se identifican tres componentes de los sistemas de *workflow* que tiene como aspecto central "tarea", los cuales son: el trabajo organizado como un conjunto de tareas ordenadas, participantes quienes ejecutan estas tareas, y los recursos que son necesarios para lograr estas tareas.

Reinwald (1994), define un sistema manejador del *workflow* de la siguiente manera:

Un sistema de manejo de workflow es un sistema activo que maneja el flujo de procesos de negocios desempeñado por múltiples personas. Esto coloca los datos correctos, a la gente correcta, con las herramientas correctas en el tiempo correcto [4].

En esta definición tenemos la restricción de que los procesos de negocios se ejecutan por personas, y no abarca otros recursos (TI), menciona tres componentes adicionales: datos, herramientas y tiempo.

Un sistema manejador de *workflow* controla de manera automatizada los aspectos de un proceso de negocio, a través de instancias de procesos, proceso definido anteriormente precisamente para crear y manejar dichas instancias. Por lo tanto, una instancia del proceso es: *la representación de lo que actualmente está pasando en el proceso [19].*

Una instancia de proceso se crea, manipula y (eventualmente) finaliza por un WMS, de acuerdo con la definición del proceso. Cada instancia de proceso representa una ejecución individual del proceso, usando sus propios datos, y con la cual es (normalmente) capaz de llevar control independiente y comprobar como avanza hacia su terminación. Representa la unidad de trabajo con respecto a un proceso de negocio el cual pasa a través de un WMS (por ejemplo, el proceso de un reclamo de seguro, o la producción de un diseño de ingeniería). Cada instancia del

proceso muestra un estado interno, el cual representa su progreso hacia su término y su estado con respecto a sus actividades.

Esta instancia de proceso puede incluir una o varias instancias de actividades, que representan durante la ejecución las actividades automatizadas. Una instancia de actividad es: *La representación de una actividad dentro de la ejecución de un proceso, por ejemplo, dentro de una instancia de proceso* [19].

Cada instancia de actividad representa una sola llamada de una actividad, relacionada exactamente a una instancia de proceso y usa los datos asociados con la instancia del proceso. Varias instancias de actividades podrían ser asociadas con una instancia del proceso, donde actividades paralelas existen dentro del proceso, pero una instancia de actividad no puede asociarse con más de una instancia del proceso.

Esta instancia de actividad incluye elementos de trabajo y/o invocación a aplicaciones en donde un elemento de trabajo es la tarea asignada a un participante del *workflow*. Un participante de *workflow* es *un recurso el cual realiza el trabajo representado por una instancia de actividad del workflow*. Este trabajo se manifiesta generalmente como uno o más elementos de trabajos asignados al participante del *workflow* mediante la lista de trabajo [19].

Un elemento de trabajo es *la representación del trabajo para ser procesado (por un participante de workflow) en el contexto de una actividad dentro de una instancia de proceso* [19].

Una actividad genera uno o más elementos de trabajo los cuales juntos constituyen la tarea para llevarse a cabo por el usuario (participante del *workflow*) dentro de esta actividad. En ciertos casos podría ser completamente manipulado por una aplicación solicitada la cual puede operar sin un participante de trabajo.

Los elementos de trabajo son normalmente presentados por el usuario a través de una lista de trabajo, la cual mantiene detalles de los elementos de trabajo asignados a un usuario y a un manejador de lista de trabajo, el cual interactúa con la lista de trabajo en beneficio del usuario.

La lista de trabajo está definida como: *una lista de elementos de trabajo asociados con un participante del workflow (o en algunos casos con un grupo de participantes de workflow quienes podrían compartir una lista de trabajo en común)*. La lista de trabajo forma parte de la interfaz entre un arranque de *workflow* y el manipulador

de la lista de trabajo.

Generalmente, un manejador de lista de trabajo requerirá elementos de trabajo desde un arranque de *workflow* en orden para crear tal lista. Esto algunas veces se hace a través de un mecanismo de consulta.

Entendamos como manejador de lista de trabajo lo siguiente: *un componente de software que manipula la interacción entre el usuario (o grupo de usuarios) y la lista de trabajo mantenida por un motor de workflow. Esto permite a los elementos de trabajo ser pasados desde un WMS a usuarios y notificaciones de terminación u otras condiciones de trabajo para ser pasadas entre el usuario y el WMS.*

Un manejador de lista de trabajo podría asignarse como un componente del software manejador de *workflow*, o ser desarrollado como una aplicación cliente. Un manejador de lista de trabajo podría comunicarse con varios sistemas *workflow* consolidando elementos de trabajo de un usuario en una sola lista de tareas para presentación al usuario. Este principio será extendido para incluir otros recursos de información externos tales como elementos de correo en una bandeja.

Las funciones posibles que podrían realizarse por el manejador de la lista de trabajo incluyen la selección un elemento de trabajo, reasignación de un elemento de trabajo, notificación del término de un elemento de trabajo, invocación de una herramienta o aplicación cliente como parte del procesamiento de un elemento de trabajo.

A lo largo de todos estos conceptos o terminología empleado en *workflows* hemos hablado precisamente de realización de actividades, que se encuentra dentro de un proceso completamente definido, y que dichas actividades son realizadas por participantes. Pero cuanto tiempo se realizan estas actividades, y bajo que condiciones se realizan o que condiciones se tienen que cumplir para realizarlas, concluir las y dar inicio a otras. El concepto de interacción nos ayuda comprender un poco esto. Una interacción está definida como: *un ciclo de actividades del workflow que envuelven la ejecución repetitiva de una o mas actividades hasta una condición es encontrada.*

Una condición sabemos que es una expresión lógica que debe cumplirse para dar lugar a que alguna acción se lleve a cabo. Para el caso de un *workflow*, hablaremos de pre-condición y post-condición, mismas que deben evaluarse antes de que una instancia del proceso o actividad se dé o concluya. Por lo tanto definidos como pre-condición, lo siguiente:

Una expresión lógica la cual puede evaluarse por un motor de workflow para decidir si una instancia del proceso o actividad en un proceso puede comenzar.

Una o más pre-condiciones pueden definirse como un criterio para una actividad particular o instancia. La pre-condición puede referir a un dato relevante en la expresión y puede también preguntar las variables tales como fecha u hora. Las pre-condiciones son definidas en la definición del proceso.

Una post-condición se define como: *una expresión lógica la cual puede evaluarse por un motor de workflow para decidir si una instancia del proceso o actividad en una instancia será completada.*

Una o más post-condiciones pueden ser definidas como criterio de terminación para una actividad particular o instancia del proceso. Tales condiciones pueden formar parte de una iteración, en la cual uno o más actividades son repetitivamente ejecutadas hasta que la post-condición definida es encontrada.

La post-condición puede referir a un dato relevante en la expresión y puede también preguntar por variables tales como fecha u hora. Las post-condiciones son definidas en la definición del proceso.

Ésta es la forma en como se relaciona todas las entidades de un *workflow*, relación y entidades que son funcionales independientemente del tipo de escenario en donde se requiera implementar un sistema de *workflow*, como puede ser el industrial, administrativo, etc. Para lo cual se cuentan ya con diferentes tipos de *workflow*, con características definidas.

Existen diferentes tipos de *workflow*, dentro de los cuales podemos mencionar:

- **Producción.** Cuyo objetivo es manejar grandes números de tareas similares, y optimizar la productividad.
- **Administrativo.** La característica más importante de un *workflow* administrativo es la facilidad para definir el proceso. Pueden existir diferentes instancias de procesos corriendo concurrentemente las cuales tienden a involucrar un gran número de empleados.
- **Colaborativo.** Se enfoca en equipos de trabajo con metas comunes. Los grupos varían desde pequeños, equipos orientados a proyectos, a gente dispersa geográficamente con intereses comunes que demandan el uso de Internet y la *Web*.

- Ad-hoc. Un sistema *ad-hoc* permite a los usuarios crear definiciones de procesos fácilmente conforme se conozca el contexto de su ejecución.

Un *workflow* como hemos visto, involucra primeramente la definición de un proceso, las actividades que se encuentran dentro del proceso, donde pueden ser manuales y automatizadas y la interacción entre ellas, así como la intervención de participantes que llevan a cabo dichas actividades. Dentro de estas actividades el intercambio de información es crucial porque es la entidad principal a manejar dentro de un proceso de negocio.

Esta información se incluye dentro de documentos, los cuales se manejan por cada uno de los miembros de la corporación. Documentos que también son las entidades principales para un sistema de gestión. Documentos que son creados primeramente por uno o varios recursos de un sistema de *workflow*, y gestionados a otros recursos. Estas operaciones con los documentos pueden realizarse en colaboración de varios participantes. En la elaboración de un documento manejado en un proceso de negocio, lo ideal es utilizar un sistema de *workflow* con posibilidad de colaboración, donde el uso de Internet es el medio apropiado para llevar a cabo las actividades.

En la siguiente sección se abarcará el tema de Internet, principalmente la *World Wide Web*, y los elementos con los que se cuenta para llevar a cabo la representación de la información.

2.4 World Wide Web y la gestión documental

Colaborar se define como la participación de personas para lograr un fin. Desde los inicios del hombre se ha visto que esta forma de trabajo para la realización de las actividades ha contribuido al desarrollo de todos los aspectos de la vida y de la civilización. En todas partes y en todas las áreas podemos apreciar como el hombre al trabajar en conjunto ha logrado la creación de grandes obras arquitectónicas, de grandes descubrimientos y de grandes desarrollos científicos y tecnológicos que resultan en mejores formas de vida.

Internet surgió como un proyecto del gobierno para conseguir la información de las redes de radio y las basadas en los satélites del Departamento de Defensa de los EEUU. Esta primera red, conocida como ARPANET, consistía de cuatro computadoras conectadas por modem permitiendo el envío de la información de uno a otro a través de las líneas telefónicas. ARPANET cambió su nombre a

Internet y su uso se extendió a las compañías que podían conectar sus computadoras a pesar de los diversos sistemas operativos con que trabajaban. Hoy en día Internet ha crecido enormemente y miles de personas alrededor del mundo se comunican por este medio.

Internet permite el intercambio de información entre computadoras. Actualmente existen diferentes servicios que permiten realizar este intercambio, entre los cuales se puede mencionar: telnet, ftp (protocolo de transferencia de archivos), gopher y http (protocolo de transferencia de hipertexto) el cual es uno de los servicios más importantes. Por medio de http se pueden leer e interpretar archivos de una máquina remota. Dichos archivos pueden contener texto, imágenes, sonido, video. Este protocolo se utiliza en la *Web*.

A través de la *World Wide Web* se puede visualizar la información de un sitio a través de documentos de hipermedia. Estos documentos pueden hacer referencias unos a otros a través de enlaces, los cuales solo permiten acceder al documento solicitado para su lectura. Si se desea acceder al documento para su elaboración en colaboración y posteriormente enviarlo a otro participante, los enlaces actuales están limitados a realizar esto, debido a que solo manejan los métodos *Get* y *Post* del protocolo http. *WebDAV (World Wide Web Distributed Authoring and Versioning)* es una extensión de http/1.1 que define nuevos métodos para colaborar en la *Web*.

Dentro de las características con las que debe cumplir un sistema de gestión documental es el tener enlaces dentro del mismo documento y hacia otros documentos que se relacionen entre sí. Por lo tanto, se necesita el uso de enlaces o vínculos para navegar entre documentos que pueden involucrar diferentes tipos de medios.

Un enlace o vínculo está formado por un par de anclas: fuente y destino. El ancla se refiere a una parte del documento. El ancla fuente es de donde el vínculo parte y el ancla destino es el punto a donde el vínculo arriba, es decir, la información que el vínculo debe mostrar. Esto puede darse en el mismo documento y/o documentos diferentes. Esta forma de organizar la información la cual incluye vínculos a otros documentos se conoce como hipertexto. Ted Nelson define al hipertexto [7], como:

Una combinación de lenguaje natural con la capacidad de computadoras para despliegue dinámico, interactivo, ramificado de un texto no lineal, el cual no puede ser impreso en una página convencional.

Los vínculos pueden usarse para varias funciones, como las siguientes:

- Pueden conectar una referencia al mismo documento.
- Pueden conectar un comentario o anotación al texto acerca del cual se escribe.
- Puede proveer organización de información (para instancias, establece las relaciones entre dos piezas de texto o entre una tabla de contenido y sus secciones).
- Pueden conectar dos piezas de texto sucesivas, o una pieza de texto y todos sus sucesores inmediatos.
- Pueden conectar entradas en una tabla o figuras para descripción o para otras tablas o figuras.

La forma de ir de un punto a otro en un documento es por medio de vínculos, proporcionados por la *Web*, pero como se muestra esta información contenida en el documento para posteriormente lograr esta navegación entre la información.

La *Web* hace uso de lenguajes de marcado para la representación de la información al usuario. Un lenguaje de marcado se define como:

Un lenguaje de marcado o de anotación construye un conjunto de reglas que definen todo aquello que es parte de un documento digital pero que no pertenece al texto del mismo.

Un archivo de un lenguaje de marcado cualquiera lleva lo mismo que cualquier comunicación usual:

- el mensaje (la información que vemos y queremos transmitir), y
- metainformación (información acerca del mensaje)

Los lenguajes de marcado pueden ser la solución ideal en situaciones como:

- intercambiar mi información entre diferentes computadoras, plataformas y software,
- filtrar y buscar en mi información en formas muy precisas o sofisticadas,
- publicar la información en diferentes medios (visual electrónico, papel, CD ROM, auditivo, etc.), o
- reutilizar o reprocessar mi información durante mucho tiempo.

En general los lenguajes de marcado siguen una sintaxis sustentada en el uso de

etiquetas; una etiqueta que indica el principio de un elemento y otra el final del mismo.

El SGML (*Standard Generalized Markup Language*) es la especificación que dio origen de todos los lenguajes de marcado.

Es un estándar internacional (de la ISO) para la definición de métodos independientes de dispositivos y plataformas de representación de textos en forma electrónica.

Es también una estructura que permite guardar información que facilita la manipulación, búsqueda precisa y de arreglo de grandes repositorios de información.

Dentro de los lenguajes de marcados utilizados en la *World Wide Web* existen HTML, XHTML, XML.

HTML (*HyperText Markup Language*) es un lenguaje muy sencillo que permite describir hipertexto, es decir, texto presentado de forma estructurada y agradable, con hiperenlaces que conducen a otros documentos o fuentes de información relacionadas, y con inserciones multimedia (gráficos, sonido, etc.).

Un enlace o hipervínculo se puede ver como una parte básica de un elemento hipertexto, ya que a través de éste se puede navegar entre documentos, es decir, un enlace es una conexión de un recurso *Web* a otro. La forma en como se logra esta navegación dentro de un documento *Web* es por medio de la etiqueta `<a>` (o bien `<A>` ya que no hay distinción entre mayúsculas y minúsculas) como se ilustra a continuación:

```
<a href = "URL" >Texto a visualizar </a>
```

Dicha etiqueta utiliza los siguientes elementos:

- El atributo `href` en el cual se especifica la dirección del ancla destino a través de un URL.
- El *texto a visualizar* es el texto que se desplegará en la página *Web*, es decir, lo que el usuario visualizará.

El ancla destino puede especificarse en el elemento `<a>`, por medio del atributo *"name"* o por algún otro elemento, utilizando el atributo *"id"* como se ilustra a continuación.

```
<H1>Contenido</H1>  
<P><A href="#seccion1">Introducción</A><BR>  
<A href="#seccion2">Capitulo 1 </A><BR>
```

```

<A href="#seccion3">Capitulo 2</A><BR
...el resto de la tabla de contenido...
...el cuerpo del documento...
<H2><A name="seccion1">Introducción</A></H2>
...sección 1...
<H2><A name="seccion2">Capitulo 1</A></H2>
...sección 2...
<H3><A name="seccion3">Capitulo 2</A></H3>
...sección 3...

```

Con la etiqueta `<A>` podemos hacer referencia tanto a un documento distinto como a una sección del mismo documento. Al especificar el URL, lo que se hace es solicitar una copia del recurso que se encuentra en la ubicación indicada para visualizarla en la máquina del cliente (un navegador de Internet como MS Explorer o Netscape Navigator) (Figura 2.3).

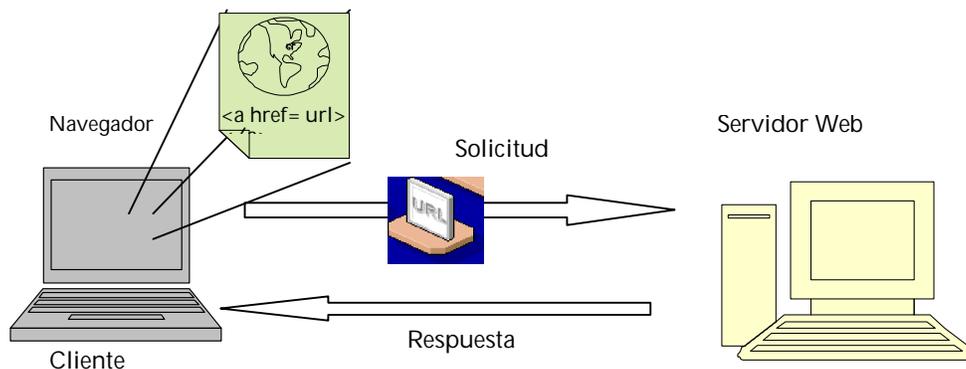


Figura 2. 3: Operación de la etiqueta `<a>`

Como vemos el *World Wide Web* provee de un ambiente bastante poderoso para la gestión documental bajo un entorno de colaboración. Ya que cuenta con un lenguaje con el cual podemos mostrar y vincular información, así como también al ser accesible desde diferentes partes del mundo.

2.5 Sinopsis

En el presente capítulo se ha descrito la problemática a resolver con un sistema de Gestión documental, así como también las características de un sistema como tal. Se han definidos conceptos de *workflow* que permitirán completar un sistema de gestión documental. Al realizar un sistema de gestión documental en la *Web*, es importante también conocer acerca del *World Wide Web* y la facilidad que nos brinda al contar con un lenguaje de marcado como HTML, con el cual se puede

satisfacer algunos requerimientos de un sistema de gestión documental.

Al implementar un sistema de gestión documental en la *Web*, es importante contar con un protocolo de red, que permita además de la transferencia de información para facilitar la colaboración. En el siguiente capítulo se hablara de WebDAV, el protocolo *Web* que permite la colaboración en la *Web*.

Capítulo 3 Protocolo WebDAV

En el capítulo anterior se expusieron los conceptos de gestión documental y sus características, así como los conceptos de *workflow* y *World Wide Web*, los cuales permiten integrar un sistema de gestión documental en la *Web*. Ahora hablaremos del protocolo de red que permite la colaboración en la *Web*.

WebDAV es un estándar que describe como, a través de la extensión del protocolo HTTP 1.1, pueden realizarse acciones de gestión de archivos tales como escribir, copiar, eliminar o modificar.

No sólo se trata de escribir archivos en una ubicación utilizando HTTP. WebDAV también ofrece la posibilidad de mover o copiar archivos en el servidor, modificar sus propiedades, su nombre, sus características de seguridad, o sus niveles de acceso etc.

WebDAV es una extensión al protocolo HTTP 1.1 que sigue el lema: *makes the Web writable*. El objetivo es convertir a la *Web* en un espacio de colaboración, en donde varias personas distintas puedan participar conjuntamente en la elaboración de documentos.

En este capítulo se dan primeramente las contribuciones del protocolo WebDAV precisamente para cumplir con los requerimientos que demanda sistemas de colaboración. Se desatacan las características del protocolo WebDAV que permite llevar a cabo la colaboración mediante escritura, bloqueo, movimiento y copia de recursos. Posteriormente se describen los métodos con los que cumple sus características y por último se describen las aplicaciones servidores y clientes de WebDAV.

3.1 Conceptos básicos de WebDAV

El protocolo WebDav (*Web-based Distributed Authoring and Versioning*) es una extensión de http/1.1 que define nuevos métodos para ofrecer colaboración en la *Web*. Para usar la nueva funcionalidad WebDAV, una aplicación cliente puede tener acceso directo al protocolo y puede estar disponible para pasar ciertos parámetros dados por el usuario.

En general, el *World Wide Web* tiene un alto potencial como plataforma para sistemas *groupware* distribuidos. Los autores crean e intercambian documentos localmente y actualizan estos a un servidor *Web*, donde sus colaboradores pueden accederlos. Sin embargo, los navegadores y servidores *Web* los cuales implementan el protocolo *http* están muy limitados para proveer escritura de documentos *Web*. Por esta razón, existen sistemas *groupware* basados en *Web* (el sistema BSCW) que se implementan usando *scripts* del lado del cliente y del servidor. Estos *scripts* se usan para implementar las operaciones básicas necesarias para este tipo de aplicaciones.

El protocolo *WebDAV* fue desarrollado como una extensión de *http/1.1* por el grupo de trabajo *WebDAV* del *Internet Engineering Task Force* (IETF) y define nuevos métodos los cuales ofrecen servicios adicionales para editar y administrar archivos en servidores *Web* remotos en una forma estructurada. Usando *WebDAV*, la *Web* enriquece la infraestructura para aplicaciones que ofrecen colaboración, donde los métodos *WebDAV* se usan para implementar las operaciones básicas necesarias (copiar, borrar, actualizar, asignar meta datos a documentos, etc.).

El protocolo permite a usuarios colaborar para crear sus contenidos directamente a un servidor *http*, permitiendo al *Web* verse no sólo como un medio de solo lectura, sino también como un medio de escritura.

Uno de los escenarios para el protocolo *WebDAV* es la elaboración de un artículo académico por un equipo de autores dispersos geográficamente. Al comienzo del proyecto, un miembro del proyecto utiliza un mecanismo específico del servidor para crear un nombre de usuario y contraseña para cada colaborador. Luego crea el documento compartido con un procesador de palabras que habilita *WebDAV* y lo guarda en un sitio con un URL. Después otorga acceso de escritura a cada colaborador, comunica el URL del documento al equipo usando mail o teleconferencia. Los autores negocian su sección de edición del documento usando email, teléfono u otras formas de comunicación. Durante cada edición, un autor carga el documento desde el URL, causando el bloqueo del mismo. Entonces trabaja en el editor, guardando directamente de regreso al URL. Cuando se concluye, los autores finalmente guardan en el depósito dado por la dirección del URL, entonces desbloquean el documento. En todo este tiempo, otros autores pueden cargar el documento directamente desde el URL usando su procesador de palabras (solo lectura, si está bloqueado) o en su navegador *Web*, para ver el estado actual. Pueden proporcionar URL a colegas interesados en el artículo y crear hipervínculos al artículo en otras páginas *Web*.

Las principales contribuciones del protocolo WebDAV, incluyen los siguientes puntos:

- Actualizar aplicaciones que no dan soporte a la colaboración. Existen muchas aplicaciones que no soportan autoría con colaboración. El protocolo WebDAV provee a estas aplicaciones una forma para facilitar autoría con colaboración sin modificar dramáticamente la aplicación.
- Protocolo orientado a la no conexión para control de concurrencia. El protocolo WebDAV provee bloqueo compartido o exclusivo de escritura, para control de concurrencia. Para lograr colaboración en Internet, donde las conexiones de red puede ser desconectadas arbitrariamente, y cada conexión abierta consume recursos del servidor, WebDAV implementa independencia del bloqueo y la conexión de red.
- Soporte para manipulación remota de espacio de nombre de documentos. El protocolo WebDAV provee operaciones de copiar, mover, y brinda soporte para listado de los miembros de una colección *Web*, similar a un directorio listado en un sistema de archivos. Los sistemas de autoría con colaboración necesitan mecanismos para navegar en el documento a construir y mantener una agrupación lógica del documento en la colección.
- Satisfacción simultánea de requerimientos. WebDAV cuenta con características importantes para la adopción de una tecnología de colaboración remota, como son, conexión segura, robusta, autenticación, internacionalización e independencia de ambientes. Mientras WebDAV se ha beneficiado al adoptar tecnologías existentes tal como TLS (*Transport Layer Security*), *Digest Authentication* y XML, la identificación de los requerimientos y la integración de estas tecnologías para satisfacer los mismos es única.
- Extensibilidad. En el diseño del protocolo WebDAV se ha reconocido que no posee todas las características que un sistema de autoría necesita. Así la extensibilidad será una característica crítica. El protocolo fue diseñado con una buena definición de la característica del espacio de nombre que permitirán adicionar nuevas características.

3.2 Requerimientos para colaboración

Los siguientes requerimientos son los que debe contar un protocolo de red para autoría con colaboración en la *Web*.

3.2.1 Soporte para todo tipo de contenido

Debido a que en la *Web* los documentos pueden contener imágenes, textos o videos es importante que el protocolo de autoría trate todo tipo de contenido de manera uniforme.

No es posible que un protocolo de autoría en la *Web* brinde únicamente operaciones para recursos con cierto tipo de contenido específico, como HTML, ya que no permitiría a los colaboradores integrar, o vincular recursos de otro tipo de contenido. Además hoy en día muchos tipos de contenido están en constante evolución, que para asegurar la estabilidad del protocolo es necesaria una separación entre el protocolo mismo y el formato de los recursos operados por el protocolo.

3.2.2 Soporte de control de concurrencia

La *Web* es un sistema multiusuario y, por lo tanto, un protocolo de autoría *Web* deberá controlar el acceso concurrente de múltiples autores.

Las herramientas de autoría *Web* primitivas, se enfrentan con el problema de pérdida de actualizaciones, el cual ocurre cuando dos o más autores de una página *Web* guardan una página después de otra, sin primero mezclar adecuadamente los cambios. Aunque HTTP tiene soporte para detectar pérdida de actualización a través del uso de identificadores únicos asociados con el estado del documento, el soporte no fue provisto para prevenir pérdida de actualización en primer lugar.

WebDAV usó al principio bloqueo completo de recurso como mecanismo de control de concurrencia. Enfocándose entonces a decidir las propiedades de bloqueo:

- Independencia de bloqueo. La operación de bloqueo deberá ser independiente de otras operaciones. Por ejemplo, debe permitir recuperar el recurso bloqueado (tal es el caso de *Get* con bloqueo). La primera motivación para este requerimiento es mantener una separación entre el bloqueo y las otras

operaciones http.

- **Bloqueo multi-recurso.** El protocolo deberá soportar una operación atómica para bloquear múltiples recursos en el mismo servidor. Los documentos muchas veces están formados por diferentes tipos de datos, puestos separadamente pero que son vistos todos como un solo recurso. Por lo tanto una herramienta de autoría necesita la habilidad para prevenir pérdida de actualización en todos los componentes de los documentos multi-partes, así el documento puede considerarse como una unidad. Un bloqueo multi-recurso garantiza esto. La restricción atómica se refiere específicamente a que si dos personas están tratando de bloquear el mismo grupo de recursos, solamente uno logrará el bloqueo.
- **Bloque de escritura.** El protocolo se requiere solo para soportar un bloque de escritura ya que no es necesario bloqueo de lectura. En la *Web*, la funcionalidad preestablecida es que los recursos se pueden leer, aunque esto puede estar regulado por el control de acceso. Por tal razón, http no requiere que un navegador *Web* haga bloqueo para leer un recurso, como es el caso con un bloqueo tradicional de base de datos.
- **Localización de bloqueo.** Esto deberá ser posible para descubrir si un recurso está bloqueado. Estos permiten a una interfaz de usuario construirse, la cual indicará si un recurso está bloqueado.

3.2.3 Soporte para propiedades

Un recurso *Web* contiene una cantidad significativa de información que no está necesariamente almacenada en el mismo recurso, por lo cual es importante que un protocolo de autoría *Web* sea capaz de crear, modificar, leer y borrar de manera arbitraria propiedades en todos los recursos *Web*, independientemente del tipo de contenido.

Muchos recursos *Web* tienen asociados información descriptiva, como es el autor, título, publicador, derechos de copia, contenido, etc. Mientras HTML, con sus etiquetas Meta provee una forma para insertar esta información en los documentos HTML, muchos tipos de recursos *Web* no tienen la capacidad para almacenar las propiedades en sus formatos. Para hacer esto, sus capacidades y lugar de almacenamiento varía con el tipo de contenido, haciéndolos difíciles de usar.

Es importante mantener la consistencia de las propiedades, de tal manera que se garantiza la veracidad de la información acerca del recurso. Por ejemplo, para un recurso la propiedad *autor*, afirma quien es el creador del recurso y está dado en la propiedad *valor*, pero si un nuevo autor comienza a editar el recurso, la propiedad debe ser actualizada o ésta será inconsistente. Un servidor *Web* puede automáticamente mantener la consistencia del valor de la propiedad que éste puede computar, tal como una propiedad que almacena la longitud del recurso. WebDAV condiciona las propiedades definidas por el servidor como “propiedades latentes” y propiedades las cuales son almacenadas sin procesamiento y sin mantener consistencia como “propiedades dormidas”.

3.2.4 Soporte para vínculos independientes del tipo de contenido

Como los recursos *Web* tienen una gran variedad de relaciones entre ellos, un protocolo de autoría *Web* deberá proveer operaciones para crear, modificar, leer y eliminar hipervínculos entre los recursos *Web* independientes del tipo de contenido. Los vínculos pueden usarse para capturar una gran variedad de relaciones, tales como tablas de contenido, índices o glosarios.

3.2.5 Recuperación para edición de fuente no procesada

El original de un recurso dado puede recuperarse por algún autor con autorización para edición. Existen muchos casos donde el original almacenado en el servidor no corresponde a la respuesta en un método *Get* de http. Los casos son directivas incluidas en el servidor y originales de lenguajes de marcado (XML) las cuales son convertidas al vuelo a HTML. Existen otros casos, tal como la conversión automática de imágenes *bitmap* dentro de diferentes variantes de tipos de medios *bitmap* (GIF, JPG) y convierten automáticamente en un medio nativo de la aplicación dentro de HTML. Como un ejemplo de este último caso, un procesador de palabra puede almacenar su tipo de medio nativo en un servidor, el cual convierte automáticamente éste a HTML. Una operación *GET* deberá recuperar el HTML. Recuperando el original se debe recuperar el formato nativo del procesador de palabras.

3.2.6 Manipulación de espacio de nombre

Puesto que los recursos necesitan copiarse o moverse a un sitio *Web*, un protocolo de autoría *Web* deberá contar con facilidades para copiar y mover recursos nombrados.

Existen varios problemas que no puede resolverse con un simple renombramiento. Por ejemplo, las características del ambiente del espacio de nombre entre servidores podrían no ser uniformes y las políticas para manejar las diferencias en las fechas de actualización podrían variar a través del espacio de nombre. Los recursos *Web* existen en un espacio de nombre identificado por un URL y un protocolo de autoría *Web* necesita proveer soporte para manipular este ambiente.

3.2.7 Soporte para colecciones

Un protocolo de autoría *Web* deberá permitir la creación y eliminación de colecciones, la inserción y eliminación de miembros en una colección y el listado de sus miembros.

Las colecciones son recursos que pueden contener a otros recursos incluyendo otras colecciones. Usando esta organización se pueden agrupar recursos que ayudan a reducir las dificultades para administrar los recursos en la autoría de documentos. Las colecciones también se usan para navegar en documentos grandes, visitando colecciones de archivos a través de las opciones de "Guardar como" y "Abrir" de la herramienta usadas. Los protocolos de autoría *Web* necesitan ofrecer este estilo de navegación.

3.3 Características

Esta sección describe en detalle las mejores características del protocolo WebDAV *propiedades, bloqueo, manejo de* espacio de nombre, que fueron diseñadas para satisfacer los requerimientos de autoría dada en la sección previa.

3.3.1 Propiedades

Una propiedad WebDAV se define por un par nombre y valor. El nombre de la propiedad es un URI y el valor de la propiedad se expresa como elementos XML. Se tiene dos métodos para manipular propiedades, PROPFIND y PROPPATCH.

El método PROPFIND se usa para recuperar propiedades y soporta tres operaciones: recupera nombre y valor de todas las propiedades, recupera nombres y valores seleccionados o recupera solamente la propiedad *nombre*. Cuando se aplica a una colección PROPFIND se le puede indicar que actúe recursivamente en la colección y sus miembros. Recursividad quiere decir que si el recurso es una colección que cuenta con miembros e incluso otras colecciones, la operación llevada a cabo por PROPFIND se realiza una y otra vez a lo largo del contenido del recurso. Es importante resaltar que cuando se usa recursividad, solamente una solicitud se envía, y solamente una respuesta se devuelve conteniendo información de la propiedad para todos los recursos afectados.

El método PROPPATCH puede redefinir o remover múltiples propiedades en un recurso en una operación atómica. Esto es, que cualquiera de los comandos de redefinir y remover en PROPPATCH será satisfactoriamente aplicado o no en el orden. Como PROPFIND, PROPPATCH puede también aplicarse recursivamente en todas los recursos de la colección.

Consideraciones de diseño para las propiedades

El diseño adoptado para las propiedades sigue la noción de permitir “vínculos” para relacionar directamente meta datos. Así las propiedades pequeñas se registran directamente dentro del valor de la propiedad, mientras que valores grandes se acceden indirectamente a través de un URL. Este diseño es conveniente para pequeños datos, tales como nombre de autor.

El método PROPFIND permite a propiedades de múltiples valores recuperarse una vez para satisfacer los requerimientos de la consulta. La regla adicional de WebDAV es que algunos elementos XML no entendidos deberán ignorarse permitiendo a la propiedad *valores* ser extendida mientras se mantenga compatible. XML también permite a las propiedades contener datos con menor complejidad.

Los URI's se usan para nombrar propiedades WebDAV debido a que diferentes grupos y aplicaciones crearán nuevas propiedades. Los URI's forman un espacio de

nombre controlados que proveen soporte para extensiones tanto centralizadas como descentralizadas que todavía garantizan igualdad. Registrando un nuevo esquema de URI dará los beneficios de un espacio de nombre centralmente controlado, mientras que usando URL's permite a una organización con un dominio crear nuevas propiedades.

3.3.2 Bloqueo

El control de concurrencia del protocolo WebDAV permite bloquear escritura. El método LOCK soporta dos tipos de bloqueo: "*bloqueo de escritura exclusiva*" y "*bloqueo de escritura compartida*". Debido a que diferentes tipos de autores pueden escribir a un recurso *Web*, un bloqueo de escritura garantiza que nadie más que el propietario del recurso bloqueado escriba en el mismo.

Un bloqueo de escritura exclusivo previene que otros usuarios escriban en el recurso bloqueado y que sólo el propietario del bloqueo lo haga. Hay otro tipo de bloqueo de escritura el compartido que permite que varios propietarios tengan acceso al recurso.

En la tabla 3.1, el estado actual del bloqueo se especifica en la primera columna a la izquierda y la solicitud de bloqueo se nombra en el primer renglón (superior). La intersección de columna y renglón indica el resultado de la solicitud del bloqueo que será o no otorgada dependiendo del estado actual del bloqueo.

Estado actual del bloqueo	Solicitud de bloqueo compartido	Solicitud de bloqueo exclusivo
Ninguno	Otorgado	Otorgado
Bloqueo compartido	Otorgado	No otorgado
Bloqueo exclusivo	No otorgado	No otorgado

Tabla 3. 1: Tabla de compatibilidad de bloqueo

Digest Authentication, una extensión para http especificado en Franks et al (1997), se usa para la autenticación de los autores. Este esquema usa combinaciones múltiples para encriptar tanto el nombre de usuario como su contraseña y es substancialmente mejor que el esquema de autenticación en herramientas existentes

de autoría remota.

El mecanismo de bloqueo de WebDAV está orientado a la no conexión. Cuando un bloqueo es otorgado, el servidor regresa al cliente un identificador único. Como la señal de bloqueo se almacena, el cliente no necesita permanecer conectado. Cuando el cliente subsecuentemente quiere ejecutar operaciones de escritura en el recurso bloqueado, éste reinicia la conexión a la red. Luego envía la autenticación y la señal de bloqueo al servidor junto con la solicitud de escritura. La autenticación asegura que el solicitante es quien dice ser y la señal de bloqueo asegura que el software del autor es enterado del bloqueo.

La propiedad DAV:lockdiscovery permite a un cliente localizar la señal de bloqueo para un bloqueo que esté activo. Las señales de bloqueo son de conocimiento público. La posesión de una señal de bloqueo no otorga permisos a un cliente para escribir a un recurso bloqueado. Esto es la combinación de ser autenticado de forma correcta y demostrar conocimientos del bloqueo, es decir, si se pasa la señal de bloqueo, se garantiza la habilidad para escribir en un recurso bloqueado.

Varios recursos pueden bloquearse con una sola solicitud de bloqueo ejecutando un bloqueo recursivo sobre una colección y sus hijos, recibiendo una sola señal de bloqueo en la respuesta que representa bloqueo entero.

Consideraciones de diseño para el bloqueo

Las conexiones Internet son poco confiables: una aplicación debe estar preparada para la pérdida de conexión en algún momento. Para lograr un ambiente robusto, un protocolo debe diseñarse para manejar la pérdida de conexión. Http fue diseñado como un protocolo "sin estado", donde el estado no está asociado con una conexión de red. Es más robusto que el acceso, cuando la conexión se pierde solamente afecta una sola solicitud y es más eficiente, entonces no se requiere estado de conexión para establecer una conexión cada vez que se crea. Estos también eliminan la necesidad de una operación de "resincronización" en el protocolo para recuperar el estado de conexión después de una interrupción.

El control de concurrencia en el protocolo WebDAV fue diseñado para satisfacer los objetivos del protocolo independientemente del tipo de contenido, permitiendo un ambiente de Internet robusto, y ofreciendo una fácil adopción para aplicaciones sin colaborar como hojas de cálculo, procesadores de palabras y editores de textos.

Adicionar tipo de contenido al protocolo, trae consigo el problema del número y variación de la cantidad de tipo de contenidos. Existen actualmente 19 text/* y sobre 150 application/* tipos de MIME en Internet. Además, algunos tipos de contenido, requerirán que el protocolo distinga entre y proveer soporte para múltiples versiones de estos tipos de contenidos.

El bloqueo WebDAV no requiere una conexión permanente a la red, permitiendo a colaboradores desconectarse completamente una vez que un recurso se bloqueo y descarga en la máquina local. El bloqueo de escritura exclusivo WebDAV prevé la necesidad para unir y disponer de documentos durante el bloqueo.

Finalmente, los bloqueos WebDAV tienen la granularidad de un recurso *Web* entero, el cual se ajusta fácilmente a conceptos tales como hojas de cálculo, documentos de procesador de palabra o imágenes. Muchas aplicaciones pueden convertirse fácilmente usando WebDAV en lugar de que ellas adopten un modelo propio de bloqueo.

3.3.3 Manejo de espacio de nombre

WebDAV cuenta con cinco métodos para la manipulación de espacios de nombres, DELETE, MKCOL, COPY, MOVE y PROPFIND. A continuación describiremos cada uno de ellos brevemente.

El método DELETE, toma un solo argumento, el *nombre del recurso* a ser eliminado. El protocolo WebDAV extiende el método DELETE para que pueda ser aplicado recursivamente a través de una colección. El método MKCOL toma un solo argumento, el *URL* donde una nueva colección se creará.

La forma más básica del método COPY toma dos argumentos, el *URL* de un *recurso fuente* y un *URL del destino*. La operación COPY quedará cumplida cuando el recurso fuente sea copiado y esté disponible en el destino. El método COPY puede aplicarse recursivamente a una colección. Las propiedades contenidas en el recurso fuente se copian en el destino. Las propiedades latentes (vivas) al copiarse tienen que pasar a ser propiedades dormidas, debido a que no son o no pueden ser soportadas como latentes en el destino. Un ejemplo de esto es cuando se intenta copiar una fuente con la propiedad latente "hora del día" y el destino se encuentra en un servidor diferente en otra parte del mundo, entonces deberá registrar la "hora del día" del servidor destino, con lo cual ya no coincide con el

fuente, por lo tanto se crea una propiedad dormida para ser copia en el destino y después se colocará el valor de dicha propiedad a la hora de la copia.

El método COPY puede fallar si todas las propiedades latentes no se copian como tales. Una restricción específica es que COPY no deberá fallar por alguna razón relacionada a la propiedad tal como la incapacidad para copiar algunas propiedades.

El método MOVE se define como el equivalente lógico de un COPY seguido por un DELETE, ejecutado como una operación atómica. Al igual que COPY, MOVE puede aplicarse recursivamente a una colección. Las propiedades se manejan de la misma forma y con las mismas opciones que el método COPY.

Consideraciones de diseño para las operaciones de espacio de nombre

El diseño de espacio de nombre WebDAV fue motivado por las necesidades de autoría de clientes y servidores quienes cuentan con una jerarquía estricta de espacio de nombre. Una jerarquía estricta de espacio de nombre es una en la cual si el recurso "a" y "a/b" ambos existen entonces "a" deberá ser una colección y deberá contener "a/b" como miembro. El carácter "/" especifica a su izquierda al nombre del contenedor. Muchos sistemas, especialmente ambientes de versiones más avanzadas, no requieren una estricta jerarquía de espacio de nombre. Dicho espacio de nombre es usualmente plano que emplea relaciones para expresar pertenencia. Sin embargo clientes y servidores quienes requieren una estricta jerarquía de espacio de nombre, incluyendo todo el sistema de archivos, no pueden operar propiamente si la jerarquía del espacio de nombre requeridos no se encuentra.

A lo largo de la presente sección se ha explicado las características distintivas del protocolo WebDAV, así como de las principales propiedades que satisfacen los requerimientos que un protocolo de autoría con colaboración en la *Web* debe soportar y brindar para autores dispersos geográficamente. Dicha actividad se logra mediante operaciones realizadas por métodos que permiten copiar, mover y recuperar información del recurso en construcción. En la siguiente sección describiremos un poco más estos métodos.

3.4 Métodos

WebDAV extiende el protocolo http/1.1 con nuevos métodos los cuales requieren parámetros. En HTTP/1.1 los parámetros para los métodos se codifican como encabezados. Sin embargo, a diferencia de HTTP/1.1 algunos parámetros de los métodos de WebDAV se codifican como cuerpos de solicitudes XML. Estos parámetros son proporcionados por el usuario y están generalmente relacionados con aplicaciones específicas. Por lo tanto, el uso de WebDAV requiere cualquier nuevo cliente *Web* o extensiones de navegadores existentes los cuales soportan la funcionalidad adicional.

Ya se ha mencionado que el protocolo WebDAV extiende el protocolo HTTP/1.1 haciendo posible la colaboración y esto es a través de la adición de nuevos métodos, tales como:

- PROPPATCH
- PROPFIND
- COPY
- MOVE
- MKCOL
- LOCK
- UNLOCK

También actualiza la semántica de algunos métodos ya existentes en el protocolo HTTP/1.1, que consiste principalmente en tomar en cuenta nuevos encabezados o condiciones que dan otro significado e influyen en el resultado de los métodos. Estos métodos son:

- GET
- PUT
- DELETE
- OPTIONS

Los nuevos encabezados añadidos al protocolo HTTP/1.1 son :

- DAV:
- If:
- Depth:
- Overwrite:
- Destination:
- Lock-Token:
- Timeout:

- Status-URI

A continuación se da una breve descripción de los nuevos métodos HTTP:

PROPPATCH:

Procesa las instrucciones especificadas en el cuerpo de la solicitud para establecer y/o remover propiedades definidas sobre el recurso identificado por la Petición-URI. La petición del método PROPPATCH se estructura de la siguiente forma:

```
PROPPATCH /inscripción.doc http/ 1.1.
Host: cs.cinvestav.mx
Content-type: text/xml
Content-Length: xxx

<?xml versión="1.0"?>
<?namespace href="http://www.cinvestav.mx/servescolares/dav" as="D"?>
<?namespace href="http://www.cic.mx/servescolares/z39.50/" as="Z"?>
<D:propertyupdate>
<D:set>
  <D:prop>
    <Z:authors>
      <Z:Author>Jorge Palacios</Z:Author>
      <Z:Author>Raúl Flores</Z:Author>
    </Z:authors>
  </D:prop>
</D:set>
<D:remove>
  <D:prop><Z:Copyright-Owner/></D:prop>
</D:remove>
</D:propertyupdate>
```

La respuesta a la petición anterior es la siguiente:

```
HTTP/1.1 207 Multi-Status
Content-Type: text/xml
Content-Length: xxxxx

<?xml version="1.0"?>
<?namespace href="http://www.cinvestav.mx/aservescolares/dav" as="D"?>
<?namespace href="http://www.cic.com/ aservescolares/z39.50/" as="Z"?>
<D:multistatus>
  <D:response>
    <D:href>http:// cs.cinvestav.mx /inscripcion</D:href>
    <D:propstat>
      <D:prop><Z:Authors/></D:prop>
      <D:status>HTTP/1.1 424 Method Failure</D:status>
    </D:propstat>
    <D:propstat>
      <D:prop><Z:Copyright-Owner/></D:prop>
      <D:status>HTTP/1.1 409 Conflict</D:status>
    </D:propstat>
    <D:responsedescription> Copyright Owner can not be deleted or
      altered.</D:responsedescription>
  </D:response>
</D:multistatus>
```

En el ejemplo, el cliente solicita establecer los valores de la propiedad <http://www.cic.mx/servescolares/z39.50/Authors> así como eliminar la propiedad <http://www.cic.mx/servescolares/z39.50/Copyright-Owner>. Sin embargo, la propiedad *Copyright-Owner* no se eliminó. El código de estatus *Method Failure*, para la propiedad *Authors* indica que no se llevó a cabo, lo cual crea un conflicto para la eliminación de la propiedad *Copyright-Owner*. Esto se especifica con el código de estatus *409 Conflict* que indica que "el cliente envió un valor cuya

semántica no es apropiada a la propiedad, incluye propiedades de solo lectura”

PROPFIND

Recupera las propiedades del recurso definido en la petición e indica si el recurso contiene o no miembros internos. La estructura de la petición para el método PROFIND, es la siguiente:

```
PROPFIND /files/ http/1.1
Host: cs.cinvestav.mx
Depth: 0
Content-type: text/xml
Content-Length: nnnn

<?xml versión="1.0"?>
<?namespace href="http://www.cinvestav.mx/servescolares/dav"as="D"?>
<D:propfind>
  <D:href>http://cs.www.cinvestav.mx/inscripciones/bigbox</D:href>
  <D:href>http://cs.www.cinvestav.mx/inscripciones/author</D:href>
  <D:href>http://cs.www.cinvestav.mx/inscripciones/DingAling</D:href>
  <D:href>http://cs.www.cinvestav.mx/inscripciones/Random</D:href>
</D:propfind>
```

La respuesta a dicha solicita se especifica a continuación:

```
HTTP/1.1 207 Multi-Status
Content-Type: text/xml
Content-Length: xxxxx

<?xml version="1.0"?>
<?namespace href="http://www.cinvestav.mx/servescolares/dav"as="D"?>
<?namespace href="http://cs.cinvestav.mx/inscripciones" as="R"?>
<D:multistatus>
  <D:response>
    <D:href>http://cs.cinvestav.mx/files/</D:href>
    <D:propstat>
      <D:prop>
        <R:bigbox>
          <R:BoxType>Box type A</R:BoxType>
        </R:bigbox>
        <R:author>
          <R:Name>J.J. Johnson</R:Name>
        </R:author>
      </D:prop>
      <D:status>HTTP/1.1 200 OK</D:status>
    </D:propstat>
    <D:propstat>
      <D:prop><R:DingAling/><R:Random/></D:prop>
      <D:status>HTTP/1.1 403 Forbidden</D:status>
      <D:responsedescription> The user does not have access to
        the DingAling property.
      </D:responsedescription>
    </D:propstat>
  </D:response>
  <D:responsedescription> There has been an access violation error.
</D:responsedescription>
</D:multistatus>
```

En este ejemplo, el método PROPFIND se ejecuta en la colección *http://cs.cinvestav.mx/files/*. El método se aplica a toda la colección y no solamente a alguno de sus miembros. En el método se especifica la solicitud de cuatro propiedades cuyos valores son enviados pero solamente las dos primeras propiedades son regresadas. No obstante, la tercer y cuarta propiedades no se obtienen, debido a que el usuario no tiene derecho de acceso a la propiedad causando una violación de acceso.

COPY

Copia el recurso especificado en la petición URI junto con los miembros internos del recurso en la localidad especificada en el encabezado "Destino". El recurso y los diferentes niveles existentes de la colección se copiarán en esa jerarquía.

La petición del método es sencilla y se estructura de la siguiente forma:

```
COPY /oolmedo/inscripción.doc http/1.1
Host: cs.cinvestav.mx
Destination: http://cs.cinvestav.mx/coordinacion/inscripcion.doc
```

La respuesta a dicha solicitud es:

```
http/1.1 204 No Content
```

En el código anterior se ejemplifica el copiado del recurso */oolmedo/inscripción.doc* al destino <http://cs.cinvestav.mx/coordinacion/inscripcion.doc>. El código de estatus *204 No content*, en la respuesta indica que el recurso existe en el destino y fue sobrescrito.

MOVE

En tanto no se aplique a colecciones, la operación de movimiento sobre un recurso es el equivalente lógico de una copia (COPY), seguido del borrado del origen y donde ambas operaciones se realizan atómicamente.

El método MOVE no garantiza el movimiento de un recurso a cualquier destino, es decir, tan solo puede controlar y garantizar la operación en el mismo servidor. Pero si lo que se quiere es mover un recurso a otro destino, el método no es capaz de hacerlo.

El siguiente ejemplo muestra la forma que tiene el método:

```
MOVE /oolmedo/inscripción.doc http/1.1
Host: cs.cinvestav.mx
Destination: http://www.cinvestav.mx/servescolares/inscripcion.doc
```

La respuesta a la petición es:

```
http /1.1 201 Created
Location: http://www.cinvestav.mx/servescolares/inscripcion.doc
```

Este ejemplo muestra que el recurso `/oolmedo/inscripción.doc` se moverá al destino <http://www.cinvestav.mx/servescolares/incipción.doc>. El contenido del recurso se sobrescribirá si el destino no está vacío. En este caso, en el destino el recurso no existe, por lo tanto se crea y la respuesta es el código `201 created`.

MKCOL

Se usa para crear una nueva colección en la localidad especificada en la petición URI. Si el recurso especificado en la petición no está vacío, la operación puede fallar. Si durante la operación no existen ancestros, es decir, que debe de existir la colección previa donde se creará la nueva, esto es, si se desea crear `/a/b/c`, deberá existir `/a/b`, para que tenga éxito la petición. La petición es la siguiente:

```
MKCOL /oolmedo/inscripciones/ http/1.1
Host: cs.cinvestav.mx
```

La respuesta es la siguiente:

```
http/1.1 201 Created
```

El anterior ejemplo, realiza la creación de la colección `/oolmedo/inscripciones/` en el servidor `cs.cinvestav.mx`.

LOCK

Se usa para realizar un bloqueo en algún tipo de acceso sobre el recurso identificado en la Petición-URI. La petición se envía de la siguiente forma:

```
LOCK /coordinación/inscripciones/inscripción.doc http/1.1
Host: cs.cinvestav.mx
Timeout: Infinite, Second-4100000000
Content-Type: text/xml
Content-Length: nnnn
Authorization: Digest username="abc"
                  Realm= abc@cs.cinvestav.mx, nonce="...",
                  Uri = "/coordinación/inscripciones/inscripción.doc",
                  Response="...", opaque="..."

<?xml versión="1.0"?>
<?namespace href="http://www.cinvestav.mx/servescolares/"as="D"?>
<D:lockinfo>
  <D:locktype><D:write/></D:locktype>
  <D:lockscope><D:exclusive/></D:lockscope>
  <D:owner>
    <d:href>http://www.cinvestav.mx/computacion/contact.html</d:href>
  </D:owner>
</D:lockinfo>
```

La respuesta a dicha solicitud es la siguiente:

```
HTTP/1.1 200 OK
Lock-Token: <opaquelocktoken:e71d4fae-5dec-22d6-fea5-00a0c91e6be4>
Timeout: Second-604800
Content-Type: text/xml
Content-Length: xxxxxx
```

```

<?xml version="1.0"?>
<?namespace href = "http://www.cinvestav.mx/servescolares/" as="D"?>
<D:prop>
  <D:lockdiscovery>
    <D:activelock>
      <D:locktype><D:write/></D:locktype>
      <D:lockscope><D:exclusive/></D:lockscope>
      <D:owner>
        <D:href>http://www.cinvestav.mx/computacion/contact.html </D:href>
      </D:owner>
      <D:timeout>Second-604800</D:timeout>
      <D:locktoken>
        <D:href>
          opaquelocktoken:e71d4fae-5dec-22d6-fea5-00a0c91e6be4
        </D:href>
      </D:locktoken>
    </D:activelock>
  </D:lockdiscovery>
</D:prop>

```

El código en la solicitud ejemplifica la forma en como se crea un bloqueo de escritura exclusivo sobre un recurso */coordinación/inscripciones/inscripción.doc*. El recurso <http://www.cinvestav.mx/computacion/contact.html> contiene información del propietario del bloqueo. El servidor tiene establecido un *timeout* en el recurso, el cual causa que el bloqueo automáticamente se desactive después de una semana (604800 segundos). En el código de la respuesta se especifica en el encabezado *Lock-Token* el URL que identifica de forma única al bloqueo para esta solicitud.

UNLOCK

Remueve el bloqueo identificado por el símbolo del bloqueo en el encabezado *Lock-Token* de la solicitud en la Petición-URI y de todos los otros recursos incluidos en el bloqueo.

```

UNLOCK /coordinación/inscripciones/inscripción.doc http/1.1
Host: cs.cinvestav.mx
Lock-Token: <opaquelocktoken:a515cfa4-5da4-22e1-f5b5-00a0451e6bf>
Authorization: Digest username="abc"
  Realm= abc@cs.cinvestav.mx, nonce="...",
  Uri = "/coordinación/inscripciones/inscripción.doc",
  Response="...", opaque="..."

```

Para la solicitud anterior, se recibe la siguiente respuesta:

```
HTTP/1.1 204 No Content
```

El ejemplo muestra la eliminación del bloqueo en el recurso *coordinación/inscripciones/inscripción.doc*. Si el bloqueo incluye más de un recurso, el bloqueo es eliminado de todos los recursos incluidos en el bloqueo. El código de status *204 No Content* incluido en la respuesta se usa en lugar del código *200 Ok*, ya que no existe un cuerpo para la respuesta.

3.5 Implementaciones

El trabajo que se ha desarrollado bajo un ambiente de un protocolo de publicación con colaboración incluye: código libre disponible en la red que compone servidores y clientes DAV y algunos productos comerciales. A continuación se enlistan:

3.5.1 Servidores

- PyDAV. Está escrito en Python corre en Unix y Windows. El almacenamiento persistente para recursos es el sistema de archivos, mientras el almacenamiento y recuperación de propiedades se maneja por un manejador de base de datos.

La arquitectura del servidor mantiene una clara separación entre el código que maneja el protocolo WebDAV y el código que maneja el repositorio de recursos y propiedades. Esta arquitectura permite manejar otros protocolos para acceder al contenido de otros repositorios y por el protocolo WebDAV para acceder el contenido de otros repositorios, tales como un sistema de manejo de documentos o una base de datos relacional.

- Microsoft IIS 5 (*Internet Information Server*). Es un servidor *Web* de aplicaciones y HTTP de Microsoft para Windows 2000.
- Mod_dav. Es un módulo Apache que provee soporte DAV para el servidor de *Web* Apache. Provee todas las facilidades básicas DAV para manipulación de recurso en el mismo servidor de *Web* y las propiedades mismas de estos. Además maneja el establecimiento y eliminación de bloqueos sobre recurso, así que los clientes pueden tener acceso exclusivo para modificar dichos recursos.
- Magi. Es un servidor personal *Web* que permite a los usuarios almacenar y recuperar información fácilmente. Un servidor Magi restringe quién tiene acceso a descargar y ver archivos usando una lista de "amigos". Los usuarios pueden compartir y sincronizar información en su propia PC, laptop o palmtop usando un servidor Magi desde donde sea como si estuviera en un directorio local.

3.5.2 Clientes

- **POSTIES.** Esta aplicación provee colaboración para compartir notas. Escrito en Java, Posties almacena cada nota como un recurso *Web* separado en una colección.
- **SITECOPY.** Escrito en C por Joe Orton, éste programa utiliza el WebDAV para actualizar un sitio *Web* de un directorio local. Sincroniza un sitio remoto de un directorio local, actualizando, borrando y moviendo archivos en un servidor como sea necesario. Los cambios en el servidor no se reflejan en el directorio local; la sincronización es solamente una forma.
- **Microsoft IE 5 (Internet Explorer).** Soporta WebDAV a través de una extensión de espacios de nombres los cuales permiten a un servidor WebDAV ser desplegado usando el mismo Explorador de Windows para navegar el sistema de archivos local. Esta interfaz permite copiar, mover, borrar, renombrar, crear nuevos archivos, crear una nueva colección y ofrecer funcionalidad *dragdrop*.
- **Microsoft Office 2000.** Se integra en Microsoft IE 5 y por lo tanto todas las características que soporta Microsoft IE 5. Microsoft Office 2000 maneja recursos directamente en su caja de diálogo File-Open y File-Save.
- **Cadáver.** Es un cliente DAV en línea de comandos. Soporta envío, descarga, despliegue en pantalla, operaciones de espacio de nombre (mover/copiar), creación y destrucción de colecciones, y operaciones de bloqueo.

3.6 Sinopsis

En este capítulo se ha descrito el protocolo WebDAV, sus contribuciones, sus características y métodos implementados que cumplen con los requisitos para ofrecer colaboración en grupos.

Se han descrito las características más destacadas de WebDAV que son (aparte de proporcionar acceso a la escritura del documento vía HTTP) el bloqueo o posibilidad de cerrar dicho acceso de escritura, mecanismo imprescindible en un entorno de trabajo con recursos compartidos, para evitar que el trabajo de un nuevo usuario sobrescriba al del anterior. La posibilidad de describir propiedades (meta datos) del documento en formato XML, como título, asunto, autor, fecha,

tamaño, que aunque no aparecen en el documento proporcionan información sobre el mismo y pueden ser fácilmente gestionados por el protocolo DAV (permitiendo funciones de listado de datos, búsqueda inteligente, etc.).

WebDAV también ofrece la posibilidad de renombrar, mover, copiar o borrar documentos en el propio servidor y la posibilidad de llevar un registro de las sucesivas versiones o revisiones del documento (pudiendo recuperar cualquiera de las anteriores). También permitirá a los usuarios crear enlaces virtuales entre recursos.

Uno de los aspectos más interesantes de WebDAV es que la posibilidad de trabajar o escribir no se limita a documentos de tipo texto. La posibilidad de trabajo conjunto se extiende a elementos HTML, GIF, JPEG, etc. De hecho ya existen aplicaciones gráficas que admiten este protocolo. WebDAV no pone ninguna restricción al tipo de documentos (elementos) sobre los que se puede trabajar.

Algunos de los objetivos del protocolo WebDAV han estado disponibles "desde siempre" empleando diversos medios. Por ejemplo la transferencia de archivos que WebDAV propone por vía HTTP se viene realizando de siempre vía acceso FTP. La posibilidad de controlar las sucesivas versiones de un documento o programa es el objetivo del conocido CVS. La principal razón por la que WebDAV propone "reinventar la rueda" es la de que estas funciones pasan a ser más flexibles y poderosas en el nuevo protocolo, que al basarse en HTTP, está mucho más desarrollado.

WebDAV permitirá a multitud de dispositivos escribir de forma segura en la red, Por ejemplo, usando una cámara digital que tenga un móvil incorporado al tomarse una fotografía, ésta podría publicarse en la *Web*. Otros colaboradores podrían, también de forma inmediata, realizar trabajos de retoque o postproducción sobre la imagen. Un buscador inteligente localizaría de forma inmediata la imagen para cualquier navegante (por ejemplo, localizando las imágenes de un mismo autor).

Con WebDAV desaparecen las diferencias entre el espacio de almacenamiento local (disco duro local) y el espacio *Web*. Será mucho más práctico para las empresas guardar sus documentos en servidores *Web* desde donde podrán asegurar su inmediata disponibilidad de forma segura. Cualquier usuario dotado del navegador adecuado podrá establecer vínculos, enlaces virtuales entre documentos.

Capítulo 4 Soluciones actuales

En este capítulo se presentarán dos de los sistemas más conocidos desarrollados para facilitar la colaboración en la *Web*: *BSCW system (Basic Support for Cooperative Work)* y *DreSS (Document Repository Service Station)*. Además se busca mostrar el estado del arte en la integración de herramientas de colaboración y su uso para gestionar documentos.

4.2 Sistemas de automatización de workflow con gestión documental

La *World Wide Web* posee un gran potencial como plataforma para sistema de *groupware*. Las personas crean y cambian documentos localmente y publican estos en un servidor *Web*, donde sus colaboradores pueden tener acceso a ellos. Sin embargo, los navegadores y servidores *Web* proveen principalmente acceso de lectura a documentos. Por esta razón, los sistemas *groupware* basados en *Web* fueron implementados usando *scripts* cliente/servidor de tal forma que implementen las operaciones básicas necesarias.

Los sistemas *groupware* basados en *Web* son aplicaciones que permiten a varios usuarios, ubicados en diferentes sitios, cooperar y producir documentos en forma estructurada. Por otra parte, el *groupware* está orientado a mejorar la productividad de grupos de trabajo, los cuales realizan sus labores bajo ciertas reglas de negocios y entre los cuales hay un intercambio bidireccional de información.

A continuación se analizarán sistemas *groupware*, que han proporcionado soluciones de automatización de *workflow*, a los procesos de negocios en diferentes áreas de la industria. Los sistemas presentados son:

- *BSCW system (Basic Support for Cooperative Work)*
- *DreSS (Document Repository Service Station)*

4.3 BSCW system (*Basic Support for Cooperative Work*)

BSCW es una aplicación *Web* que soporta la cooperación tanto síncrona como asíncrona entre los miembros de un grupo de trabajo. Las principales ventajas son:

- Creación de zonas de trabajo entre los miembros del equipo en donde pueden compartir documentos para el desarrollo de proyectos. La organización se hace mediante carpetas y documentos, ordenando la información de manera jerárquica. Una zona de trabajo contiene diferentes tipos de objetos: carpetas, documentos, enlaces, artículos y citas.
 - La carpeta se identifica cuando se accede por un único miembro, cuando éste invita a otros, se convierte en una zona de trabajo. A la vez una carpeta puede tener un número indeterminado de ellas. El aspecto de la zona es totalmente dependiente de quien la accede, y quien solo podrá acceder a los objetos a los cuales tiene derecho.
 - Para el manejo de documentos, soporta diferentes formatos: texto, imágenes, sonidos e incluso archivos comprimidos. Para publicarlos simplemente se crea una carpeta en BSCW.
 - Los artículos son como una lista de distribución dentro de las zonas o carpetas y es por medio del cual se entabla una discusión con los otros miembros sobre los documentos que se comparten. Por medio de este canal que es visto por el resto del grupo al conectarse, podría ser un sustituto del correo electrónico
 - Permite tener vínculos a otros lugares del *Web*, almacenando únicamente su URL.
 - Permite las citas y reuniones, al crear un espacio en donde compartir documentos en la *Web*, y de alguna forma promover el trabajo en grupo, debe de existir un mecanismo donde se puedan establecer reuniones o encuentros entre ellos para resolver problemas. Lo cual puede ser un encuentro presencial, dejando indicado mediante mensaje, el lugar elegido. Si se realiza en la red, habrá que decidir cual es el medio como videoconferencia, o Chat. En ambos casos hay que establecer la fecha y hora de inicio y fin. Por último se invita a los participantes notificándoles por correo electrónico. Una vez recibida, se puede aceptar o declinar la invitación y rechazarla. Incluso se puede dejar un mensaje a los demás invitados para comunicar eventos o cambio de planes. No es necesario ser parte de un grupo de trabajo para ser invitado.
- Diseño de políticas de derechos de acceso. Siendo un sistema en la *Web*, es

necesario contar con un modelo de seguridad que permita un acceso limitado. El primer paso es el registro de usuarios, de quienes se determinan los privilegios que dependen del papel que desempeñan. Se debe definir la política a seguir antes de establecer la estructura del área ya que un error podría provocar difusión no deseada de información. Por tal razón, es importante analizar que las carpetas y las zonas de trabajo con las que van a contar los miembros que formarán parte del proyecto, sean creadas en grupos homogéneos basados en un buen modelo de concesión de derechos de acceso.

- Se debe definir las funciones que podrán realizar los usuarios dentro del área de trabajo, así como los derechos sobre los objetos: acceso a la información, creación de nuevos objetos, copiar, recortar, borrar, invitar o expulsar a miembros del grupo.

Existen dos caminos para definir derechos: crear grupos de trabajo dentro de la zona y/o modificarlos en los objetos que contienen.

- Clases de usuarios. Dependiendo de los accesos, existen tres tipos de usuarios:
 - § Administradores, son los gestores de cada zona y tienen todos los derechos por defecto. Pueden invitar o expulsar a otros miembros e incluso decidir sobre su capacidad de trabajo. Al invitar a otro como tal, tiene la posibilidad de delegar en él sus funciones y dejar el cargo.
 - § Miembros comunes, sus capacidades son limitadas, entre ellos se encuentran los propietarios de los documentos que se encuentran en el escritorio de trabajo.
 - § Anónimo, no es miembro activo de la zona de trabajo, por lo cual es necesario indicarle el URL exacto en donde puede encontrar el documento que se le permite acceder. Este usuario es útil cuando se quiere divulgar el contenido de un documento en forma pública
- Búsqueda de información, se tiene la posibilidad de realizar consultas tanto en *Web* las cuales se pueden realizar a través de buscadores como Altavista, Excite, Infoseek o Lycos, así como búsqueda en la zona trabajo y carpetas de las cuales se es miembro. Se puede limitar la búsqueda por el tipo de objeto o por su contenido.

4.4 DreSS

DreSS es un conjunto de herramientas que permiten a autores depositar y/o actualizar documentos en un servidor *Web*, sin la necesidad de autenticarse en el servidor *Web*, y sin la necesidad de acceder al sistema de archivos de la computadores (ya sea por *nfs* o por otros medios).

DReSS se desarrolló para los siguientes propósitos:

- DreSS muestra que un servidor *Web* puede usarse como un repositorio de documentos para múltiples autores, sin la necesidad de autenticarse (login) con el servidor.
- DreSS muestra que un repositorio de documentos puede construirse usando un navegador y servidor *Web* estándar y el protocolo *http*.

DreSS permite a usuarios adicionar documentos a un servidor en forma controlada. El autor de un documento decide quien puede editar (o borrar) el documento y el administrador decide quien puede ser un autor. Los autores pueden crear subdirectorios y colocar documentos individuales así como directorios completos en el servidor.

DreSS fue desarrollado para aplicaciones de hipertexto en donde en muchos casos los autores generan diferentes páginas (documentos individuales) enlazados entre sí. Aunque DreSS implementa operaciones bloqueo para prever pérdida de actualizaciones, no se propone para aplicaciones donde muchos autores modifican el mismo documento frecuentemente. Tales aplicaciones requieren control de versiones, las cuales no se ofrecen porque DreSS se enfoca al hipertexto y por tal razón es importante manejar nombres relativos de documentos. Los usuarios pueden colocar un directorio completo una vez y usar enlaces entre los documentos contenidos en el directorio.

DreSS se ha empleado en sistemas para el manejo de sitios *Web* de los empleados de un departamento. También se ha utilizado DReSS para cursos en línea de hipermedia. Los estudiantes pueden colocar sus tareas en un servidor *Web* del departamento, en donde se evalúan sus aciertos por *scripts*. Previamente los estudiantes envían un e-mail o usan FTP para enviar las tareas al profesor, quien deberá entonces (manualmente) activar el software de evaluación. Puesto que es un curso de hipermedia, los estudiantes muchas veces regresan un hiperdocumento, el cual consiste de un número de archivos HTML separados pero enlazados. La actualización del directorio es una característica esencial para ésta aplicación.

4.5 Sinopsis

Los sistemas descritos han logrado automatizar algunos procedimientos de gestión documental. Como se puede apreciar, es importante establecer reglas para llevar a cabo esta automatización. Una de las características de estos sistemas es la creación de una estructura de directorios, donde se depositará la información a ser consultada y/o actualizada por los participantes, así como la forma en que éstas operaciones se realizarán a través de una interfaz de usuario y que además requiere de un servidor *Web*. Así mismo, podemos observar que los sistemas centralizan la información a ser consultada en un repositorio ubicándola en un solo sitio, a diferencia de la solución presentada en esta tesis en donde la información no se centraliza en una sola computadora.

Capítulo 5 Caso de estudio: inscripción escolar

En este capítulo se describe el caso de estudio utilizado para mostrar la gestión de documentos en la *Web*. El caso de estudio utilizado es el proceso de inscripción, proceso que permite mostrar un sistema de *workflow* continuo, así como la interrelación de los participantes al realizar sus tareas dentro del proceso.

Se identifican cada uno de los elementos del proceso de inscripción que son parte de un *workflow* con soporte para la colaboración y continuo, así como también como la gestión del documento de inscripción dentro del proceso.

5.1 Descripción del proceso

Un escenario para mostrar la interrelación de las diferentes entidades de un sistema de *workflow* con soporte para la colaboración y continuo es el *proceso de inscripción* de un alumno a un programa de estudios, por ejemplo de estudios de postgrado. Mismo que se implementa en un sistema interactivo que permite mostrar el objetivo a cumplir en la presente tesis.

El sistema lleva a cabo un proceso secuencial en el cual intervienen cinco participantes: *alumno, tutor de estudios, coordinador académico, secretaria y servicios escolares*. Al iniciar el proceso, el alumno solicita su inscripción al ciclo escolar. La solicitud de inscripción deberá ser evaluada por el resto de los participantes, finalizando el proceso al aceptar al alumno en el siguiente ciclo escolar. El proceso se esquematiza en la figura 5.1.

La interacción de los participantes se describe a continuación:

- a. El alumno inicia el proceso solicitando su inscripción al ciclo escolar del programa de postgrado. Acude con su tutor de estudios para discutir las materias que cursará y que deben indicarse en su inscripción.
- b. El tutor de estudios revisa la solicitud de inscripción.

- b.1. Si el tutor de estudios está de acuerdo, firma la solicitud de inscripción, la cual será enviada al coordinador académico.
- b.2. En caso en que el tutor de estudios no esté de acuerdo con lo especificado en la solicitud de inscripción, discutirá con el alumno las materias a cursar y tratará de llegar a un acuerdo.
- c. La solicitud de inscripción se evalúa por el coordinador académico, quien puede:
 - c.1. Aceptar la solicitud de inscripción, misma que el alumno entrega a la secretaria con las firmas de autorización.
 - c.2. Rechazar la solicitud la cual es discutida con el tutor de estudios y el alumno para proceder a corregirla. (punto b.2).
- d. La secretaria envía la solicitud de inscripción a servicios escolares.
- e. Servicios escolares lleva a cabo la revisión de la solicitud de inscripción. Emite la decisión tomada con respecto a la solicitud de inscripción.
 - e.1. Si la solicitud de inscripción se acepta por servicios escolares,
 - e.1.1. Archiva una copia de la solicitud de inscripción en los expedientes administrados por el área de servicios escolares.
 - e.1.2. Envía una copia a la secretaria, quien archiva esta inscripción en el expediente del Alumno.
 - e.2. Si la solicitud de inscripción es rechazada por servicios escolares, la secretaria informa y entrega la solicitud de inscripción al coordinador académico para su corrección.

5.2 Identificación de entidades de workflow en el proceso

Cada uno de los elementos que intervienen en el proceso de inscripción tiene una función específica que los identifica dentro de un *workflow*. Por ejemplo, los participantes dentro del proceso (alumno, tutor de estudios, coordinador, secretaria, servicios escolares), realizan actividades definidas y regidas por políticas (reglas del proceso). En esta sección se identifica a las entidades que conforman el *workflow*, de acuerdo a la definición del proceso.

Los elementos que conforman al *workflow* se relacionan de la forma en como se especifica en la figura 5.2., en la cual nos enfocaremos para identificar y describir las entidades dentro del proceso de inscripción.

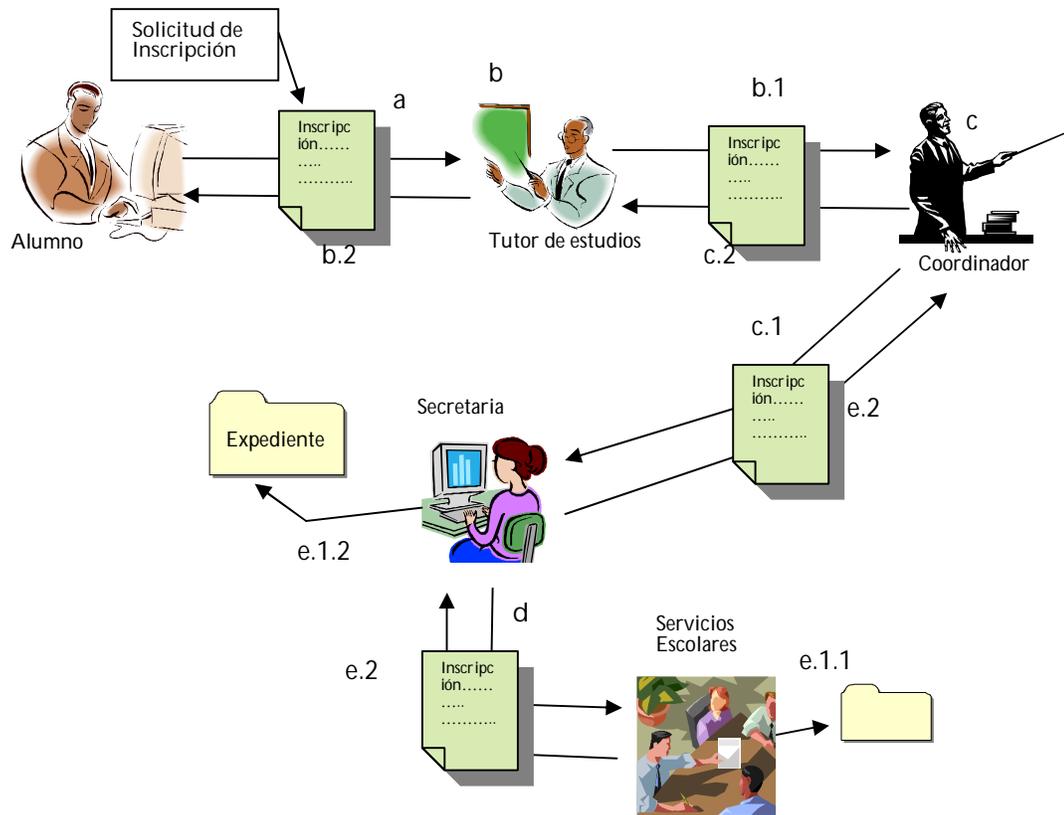


Figura 5. 1: Flujo de información en el proceso de inscripción

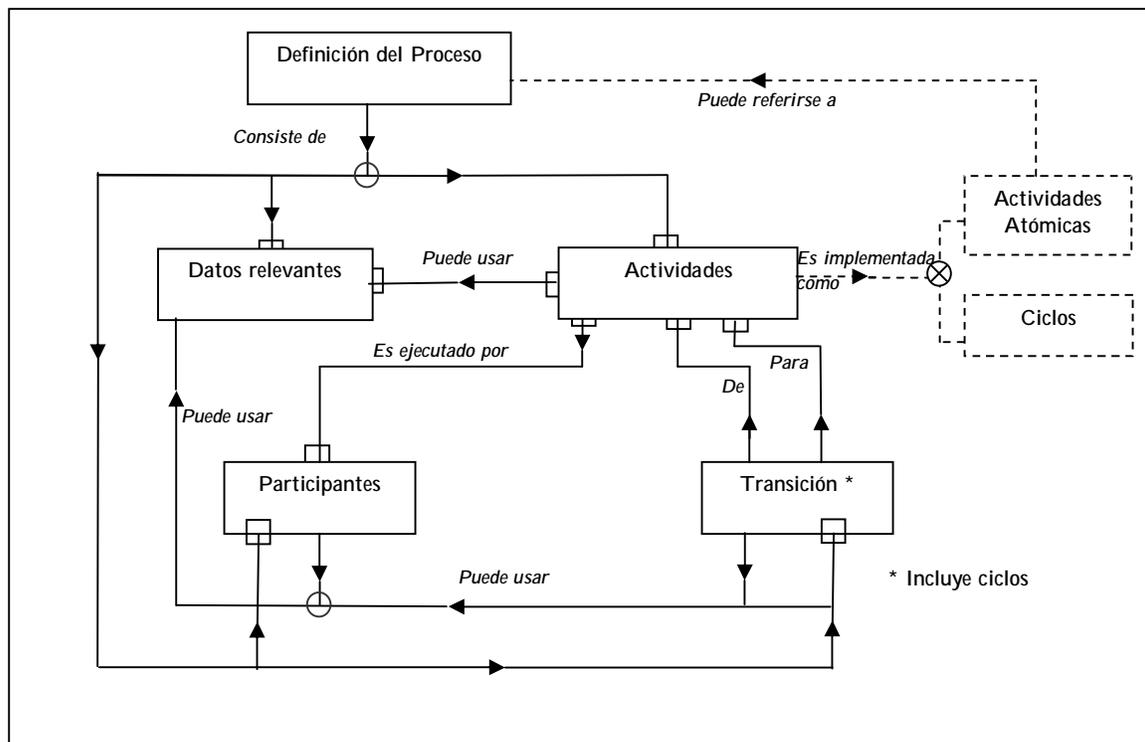


Figura 5. 2: Meta modelo de la definición del proceso

Siguiendo el diagrama del meta-modelo de la definición de proceso, identificamos y definimos los siguientes elementos.

Definición del proceso

Es el proceso de inscripción que realiza un alumno al ciclo escolar que comienza, y donde para la aceptación de dicha inscripción, se evalúa por otros participantes, quienes realizan actividades bien definidas para llevar a cabo dicho objetivo. El proceso de inscripción ya se explicó en la sección anterior.

Participantes

De acuerdo a la descripción del proceso de negocio, en este caso el proceso de inscripción, se identifican cinco participantes, los cuales son:

- *Alumno*. Persona quien solicita su inscripción al ciclo escolar.
- *Tutor de estudios*. El personal encargado de la autorización de la forma de inscripción de sus alumnos asignados.
- *Coordinador académico*. Persona quien es responsable de la autorización de todas las inscripciones al área dirigida por él, las cuales han sido previamente autorizadas por cada tutor de estudios a sus alumnos correspondientes.
- *Secretaria*. El personal de ayuda para la recopilación y captura en formato electrónico de las formas de inscripción de cada alumno, habiendo sido autorizadas por el coordinador académico. Así mismo es la persona quien mantiene la relación con servicios escolares para completar el proceso de inscripción.
- *Servicios escolares*. El departamento principal de la institución el cual rige y decide finalmente que sean aceptadas las inscripciones de los alumnos al ciclo escolar.

Actividades

Las actividades identificadas dentro del proceso de inscripción son:

- Solicitar inscripción
- Llenar forma de inscripción
- Revisar forma de inscripción
- Capturar inscripción
- Gestionar forma de inscripción
- Archivar documento de inscripción

Estas actividades como se especifica en el meta modelo del *workflow* pueden ser asignadas a uno o más participantes para su realización.

Datos relevantes

La información de mayor importancia en el proceso, debido a que es indispensable su existencia para poder realizar desde la solicitud de inscripción, hasta su aceptación es:

- Datos generales del alumno que se inscribe (nombre completo, matrícula, especialidad).
- Lista de asignaturas a impartirse en el ciclo escolar y nombre del catedrático asignado para impartirla.
- Reglamento de inscripción que ayuda a determinar la aceptación o rechazo de la inscripción.

Estos datos son utilizados en una o varias de las actividades que se llevan a cabo en el proceso de inscripción.

Transición

Ya hemos descrito las actividades, participantes e información relevante necesarios para realizarlas, el punto durante la ejecución de una instancia del proceso y una actividad se completa para pasar el control a otro participante es precisamente la transición.

La transición es el paso donde finaliza una actividad y se pasa el control para el inicio de una nueva que incluso puede ser el final del proceso. En este caso las

transiciones son:

- Autorizar la inscripción la cual está comprendida dentro de la actividad de revisar la inscripción.
- Rechazar la inscripción la cual se genera a partir de la resolución que se tome de la actividad de revisar la inscripción.
- Archivar la inscripción la cual se realiza al momento de aceptación, es decir, se le notifica al alumno su aceptación de inscripción al ciclo escolar.

Esta transición, está regida como ya se ha definido en el capítulo 2, por condiciones que se cumplen para dejar el control a la siguiente actividad. Las condiciones se dividen en pre-condiciones y post-condiciones, para el proceso de inscripción se mencionan las siguientes:

Pre-condiciones

- a. El alumno debió cumplir satisfactoriamente el ciclo escolar anterior.
- b. La solicitud se realice dentro del periodo de inscripción.
- c. Las materias elegidas por el alumno formen parte del programa de estudios de su especialidad.
- d. La forma de inscripción contenga las firmas de autorización (Tutor de estudios o coordinador).

Post-condición

- a. La forma de inscripción se encuentre debidamente hecha, es decir, contenga los datos relevantes tanto del alumno como de las materias que va a cursar.
- b. La forma de inscripción contenga la firma de autorización. En este caso estamos hablando de la firma del participante inmediato anterior, específicamente del tutor de estudios, coordinador académico, y servicios escolares.
- c. La forma de inscripción se encuentre capturada.
- d. La forma de inscripción se archive.

5.3 La gestión de documentos dentro del proceso

Se han descrito los elementos que intervienen dentro del proceso de inscripción, los participantes, las actividades a realizar, las interacciones que son regidas por condiciones, todo esto para poder llevar a cabo la gestión del documento de inscripción dentro del proceso.

Se describió anteriormente las operaciones básicas que se realizan con un documento: obtenerlos, analizarlos, gestionarlos y/o publicarlos. A continuación explicaremos como se llevan a cabo estas operaciones dentro del proceso de inscripción para ser integrables en el sistema de gestión documental.

Obtener el documento

Esta operación comienza cuando al solicitar la forma de inscripción el alumno, éste proporciona los datos que contendrá el documento (Datos generales del alumno, materias, ciclo escolar), datos que posteriormente formaran parte de la siguiente operación: *analizar*.

Recordemos que al obtener los documentos no sólo se enfoca en la generación de uno nuevo, si no también involucra la recepción de la información a partir de otros medios o de otras personas, esto es, cuando el resto de los participantes recibe el documento de inscripción lo obtiene de otro participante e incluso lo busca en un repositorio.

Analizar el documento

Una vez que el documento de inscripción se genera, es enviado a los participantes de uno en uno para que lo analicen. En esta operación el participante lee el documento, se entera del contenido del mismo, revisa que esta información cumpla con los reglamentos establecidos, puede hacer observaciones y anotaciones en el mismo y, finalmente, para que la actividad finalice deberá firmar de autorización y de ésta forma la actividad finaliza y se pasa el control al siguiente participante.

Gestionar el documento

Una vez revisado y analizado el documento, el participante (tutor de estudios, coordinador académico, servicios escolares), decide gestionarlo cuando toma una decisión, es decir, si decide autorizar o rechazar el documento lo remite al participante inmediato posterior o anterior respectivamente.

Durante la operación de gestión, el documento puede ser: (i) *remitido* cuando el momento que se envía a otro participante, (ii) *controlado* cuando el tutor de estudios analiza la forma de inscripción y acepta las materias elegidas por el alumno y, si en algún momento no lo esta, él puede decir al alumno que cambie las

materias por otras más afines a su especialidad.

Finalmente se puede decretar el documento, esto es cuando la solicitud de inscripción es autorizada por el último participante y ésta es remitida para archivar.

5.4 Discusión

El proceso de inscripción descrito en el presente capítulo, abarca cada paso que debe de realizar el alumno, desde saber cual es el período de inscripción, solicitar su forma de inscripción (en papel) hasta sostener entrevistas con su tutor de estudios o coordinador académico y esperar a que la solicitud sea aceptada por servicios escolares.

En este proceso el manejo de la información (inscripción) en un medio impreso como el papel resulta poco eficiente, ya que si tomamos en cuenta que el alumno al solicitar la inscripción se le entrega la forma en papel para que se inscriba, este documento impreso al ser autorizado es entregado a la secretaria para que ella genere otro documento a partir del original el cual será posteriormente enviado a servicios escolares quien archivará una copia del mismo. Vemos claramente la creación de diferentes copias de un documento que contiene la misma información para una sola persona, y si esto lo multiplicamos por el número de alumnos que se inscriben, estamos hablando de un número considerable de documentos impresos.

Teniendo diferentes copias archivadas del mismo documento, en diferentes lugares, la búsqueda del documento en un momento dado se vuelve ineficiente, sin mencionar aquellos casos de pérdida o de acceso no autorizado a información confidencial.

Sin analizamos esto podemos apreciar que el costo es mayor que los beneficios que podemos obtener. Ya que hay que considerar las hojas de papel que se utilizarán para imprimir cada documento y la copia del mismo, y si en un momento se necesitara volver a imprimir, aumentamos aún más el costo, ya que incluye recursos de impresión, tinta, electricidad, hojas de papel, etc. Los beneficios en realidad son pocos o ningunos, en donde tal vez, el único sea la inscripción del alumno.

El documento de inscripción es un elemento indispensable en este proceso, ya que por medio de éste, el alumno hace constar su aceptación al ciclo escolar. Siguiendo los puntos que indica la gestión de documentos, las operaciones que se realizan con

este documento son en primera instancia, la *recuperación*, es decir, como el alumno consigue el formato de inscripción al solicitárselo a la secretaria.

La siguiente operación a realizar con el documento es el *análisis*. Obviamente al buscar ser autorizada la inscripción el alumno entrega el documento a los demás participantes, los cuales proceden a leerlo, y hacerle los comentarios necesarios, lo cual implica una modificación. Dentro de los comentarios se pueden subrayar o marcar observaciones especiales para tomarse en cuenta.

Las operaciones en paralelo que pueden realizarse con el documento son: *gestión y publicación*. Al gestionar el documento, se puede decretar que sería la decisión tomada por el participante en cuestión, controlando de cierta forma la información o el estatus en que se encuentre el documento de inscripción y finalmente remitirlo al participante en cuestión según la decisión (aceptar o rechazar) del actual participante. Publicarlo implica dar a conocer al alumno la decisión tomada con respecto a su solicitud de inscripción y quedando constancia de la misma en un documento impreso archivando en un expediente, el cual se encuentra almacenado en un archivero.

El proceso de inscripción trae consigo problemas de comunicación y dependencia directa de la presencia física de los participantes, así como la disponibilidad de tiempo de los mismos. En un entorno corporativo los problemas aumentan ya que las personas al estar separadas geográficamente podrían no coordinar correctamente sus acciones.

Estos puntos plantean los inconvenientes del manejo de este proceso de forma manual, lo que lleva a la implementación de un esquema de gestión de documentos en donde se lleven a cabo las tareas a realizarse por los participantes dentro del proceso de inscripción pero dando solución a los problemas presentados, con lo cual hay que tomar en cuenta lo siguiente.

El sistema de gestión de documentos debe tener la capacidad para almacenar grandes cantidades de documentos, los cuales sean de fácil acceso para los participantes del proceso y con un tiempo mínimo de recuperación. De igual forma deberá permitir el control, la gestión, la distribución y la comparación de las diferentes versiones de los documentos.

5.5 Sinopsis

A lo largo del presente capítulo se definen cada uno de las entidades que actúan en el caso de estudio: *Proceso de inscripción*, tomado para mostrar la funcionalidad del objetivo de la tesis. Dichas entidades consisten desde la definición del proceso mismo así como cada uno de los participantes, las funciones que realiza dentro del proceso de inscripción, las reglas que definen la forma de cómo deben de llevar a cabo dichas funciones y los inconvenientes con que actualmente cuenta este proceso.

Capítulo 6 Diseño e implementación de SGD-Web

En el presente capítulo se describirá el diseño e implementación para la gestión de documentos, utilizando el proceso de inscripción explicado en el capítulo anterior. El diseño se basa en la estructura de directorios permitida por el protocolo WebDAV. El directorio se representa por una página *Web* la cual contiene subdirectorios representados por páginas estáticas o dinámicas.

Así mismo, en este capítulo se describirán a detalle la estructura que se utilizó para llevar a cabo la implementación, es decir, las características del lenguaje de marcado utilizado (HTML) para adicionar a la etiqueta `<a>` los parámetros necesarios para que junto con los métodos del protocolo WebDAV, se realice el movimiento de los documentos entre las computadoras de cada uno de los participantes del proceso. Así como también la tecnología Java que se usó para la programación, describiendo cada una de las clases utilizadas.

6.1 Diseño de la solución

La creación, almacenamiento, búsqueda, recuperación de documentos son algunas de las necesidades que un sistema de gestión documental deberá satisfacer y dentro del proceso de inscripción del caso de estudio, éstas operaciones son parte importante del mismo.

El flujo de información entre los participantes, implica el movimiento de archivos de un repositorio a otro e incluso de un equipo a otro, siendo este último caso, el utilizado en ésta implementación, es decir, cada participante cuenta con un equipo de cómputo desde el cual realiza el movimiento del documento hacia la máquina del otro participante.

Todos los equipos que participan en este proceso de inscripción deben estar interconectados a través de una red de computadoras, ya sea Intranet y/o Internet, pudiendo ser accesibles a través de su dirección IP o de su nombre. Para poder realizar esta comunicación entre equipos a través de la *Web* específicamente en este caso de estudio, se hace uso de protocolos de comunicación como es HTTP y

principalmente del protocolo WebDAV, el cual extiende HTTP para permitir la colaboración en la *Web*, como se explicó en el capítulo 4.

Una de las características del Protocolo WebDAV es el manejo de una estructura de directorios con características de lectura y/o escritura, según lo que deseamos realizar. Cada directorio creado permite contener documentos (archivos). Esta facilidad de estructurar un repositorio de documentos se puede visualizar como se muestra en la figura 6.1.

Organizar los documentos en una estructura jerárquica permite:

- Mejorar la organización de cada uno de los participantes del proceso, ya que se organiza la información en directorios, para llevar a cabo de forma óptima sus tareas.
- Se garantiza la seguridad de la información, debido a que se tienen organizados los documentos según el tipo de información que almacenan; dicha organización depende de la operación a realizarse con el documento, evitando la eliminación accidental de información
- Facilitar la administración de los archivos para realizar tareas como eliminación, renombramiento, copiado y movimiento.
- Es fácil la implementación, ya que de esta forma se puede especificar el origen y destino de los documentos, sin la necesidad de involucrarse con atributos del propio archivo, como es fecha de creación, autor y demás propiedades que podían ayudarnos a identificar de forma única al archivo.
- Los recursos son accesibles usando el protocolo http.

Con esta característica se diseñó una estructura de directorios para cada uno de los participantes que intervienen en el proceso de inscripción, donde se almacenan los documentos, y hacia donde son remitidos o depositados por cada colaborador. Esto se muestra en la figura 6.2.

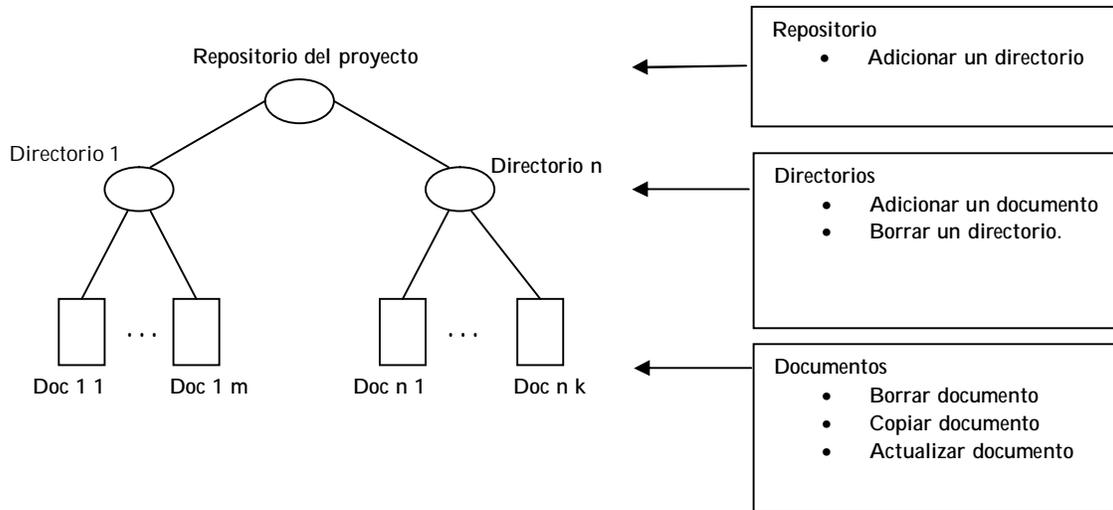


Figura 6. 1: Funcionalidad y estructura del repositorio de un proyecto basado en WebDAV

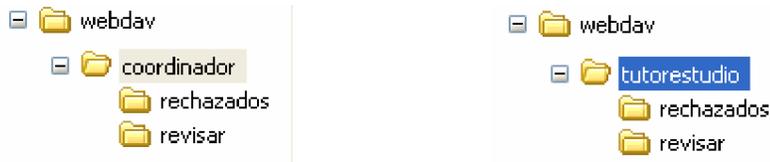


Figura 6. 2: Estructura de directorio creada para un participante.

El diseño de la interfaz usada para la representación del escritorio del participante en el proceso y en donde se realizan las operaciones para la gestión de documentos, se muestra en la figura 6.3. En ella se especifica un menú de actividades a realizar por el usuario *por revisar, rechazados, aprobados, expediente*, en el lado izquierdo de la pantalla, y que son mostradas según el perfil del usuario. En el lado derecho se despliega la información, es decir, la lista de documentos y las acciones (aceptar o rechazar) a tomar sobre dicho documento (Tabla 6.1).

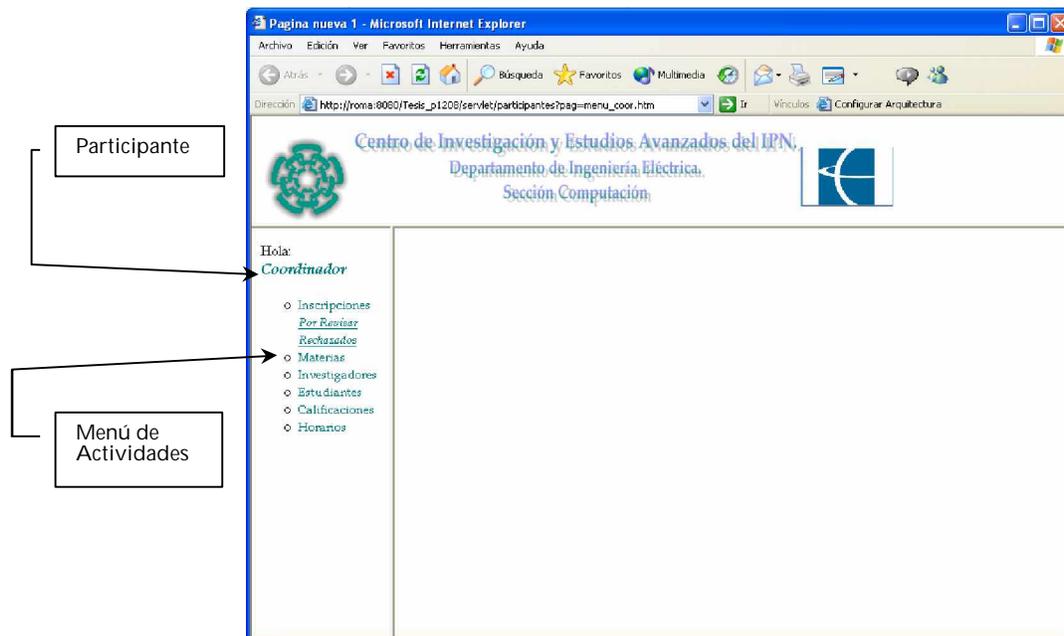


Figura 6. 3: Diseño de la simulación del escritorio del participante

El diseño implementado para la toma de decisión sobre el documento de inscripción: *autoriza o rechaza*, dependiendo de la opción del participante. Se muestra a continuación:

DOCUMENTO	ACCIONES	
Inscripcion	Autorizar documento a participante	Rechazar documento a participante

Tabla 6. 1: Acciones a tomar sobre un documento

Al tomar una decisión y al seleccionar la opción, se realiza el paso del documento, es decir, se envía el archivo de la computadora fuente a la computadora destino. Como se muestra en la figura 6.4.

Al diseñar de esta forma la aplicación, se muestra como están interactuando las diferentes entidades (participantes, documentos, actividades, reglas) dentro del proceso de inscripción. Como ya se ha explicado en una empresa existen procesos, los cuales son realizados por personal, quienes tienen definidas actividades a realizar para lograr el objetivo. En el proceso de inscripción cada participante se le ha definido actividades, explicadas en el capítulo 5, las cuales son representadas en el menú del participante (ver figura 6.3).

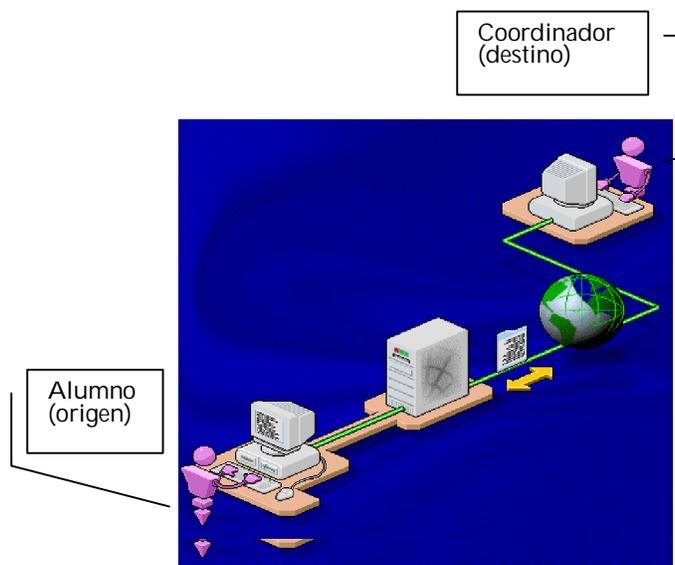


Figura 6. 4: Movimiento del documento de un participante a otro.

Dentro de este proceso, el paso de información es representado al momento de seleccionar una de las dos opciones *autoriza o rechaza*, que se especifican en la tabla 6.1., la cual podemos definir como una instancia de la actividad a realizar por el participante. Y la instancia del proceso es precisamente cuando el documento es depositado en el escritorio del participante, por ejemplo al coordinador.

La lista de actividades a realizar dentro del proceso, la cual se ejecuta por uno o varios participantes dentro del proceso, ayuda a mejorar la administración de documentos, pues se controla desde la creación del documento, modificación, análisis y gestión del mismo, y se previene el trabajo doble, controla la versión del documento, el almacenamiento y se ofrece facilidad de obtención de la información.

De acuerdo a los diferentes tipos de *workflow*, descritos en el capítulo 2, este proceso de inscripción y la forma en como se diseñó, se está hablando de un *workflow* colaborativo, esto es, un grupo de personas colabora para la creación de un documento, en este caso es la autorización que cada uno da, y se hace uso de la tecnología de Internet, diseñándose a través de interfaces *Web*.

Al desarrollar este sistema sobre la *Web*, nos permite la colaboración con personas que están ubicadas en diferentes lugares geográficamente, con lo cual se está cumpliendo con las características de un sistema de gestión documental, que debe ser corporativo. Otra ventaja es que las personas pueden tener acceso en el momento que lo deseen, la disponibilidad de la información y el manejo de la

misma independientemente del tipo de contenido, es decir, si es un documento de texto, imagen, etc.

6.2 Implementación de la solución

A continuación se detallará la forma como en conjunto se utilizaron el protocolo de transferencia de hipertexto y los métodos usados para este intercambio de información.

En base al diseño de la aplicación, cada participante tiene opciones que le permite realizar acciones (toma de decisiones) sobre el documento de inscripción: *crea*, *autoriza* o *rechaza*. De estas tres acciones, la creación del documento no realiza el movimiento del documento, pero realiza la creación de la estructura de directorios y el movimiento del documento a uno de los directorios. Dando inicio al proceso de inscripción, las acciones de autorizar o rechazar, llevan a cabo un movimiento del documento de una computadora a otra. Ver figura 6.5.

Para llevar a cabo esta manipulación de recursos, se utiliza los métodos que implementa WebDAV para la creación de directorios, copiar recursos de un origen a un destino o mover recursos.

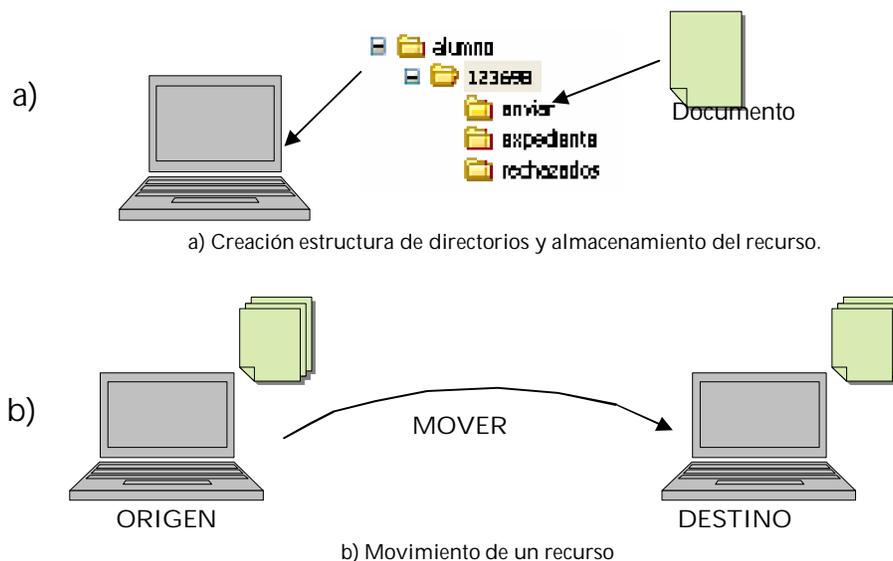


Figura 6. 5: Manipulación de los recursos.

6.2.1 Arquitectura de la aplicación

En el capítulo 3 se explicó el protocolo WebDAV y los métodos que implementa para permitir trabajo colaborativo en la *Web*. Así como las características principales de este protocolo.

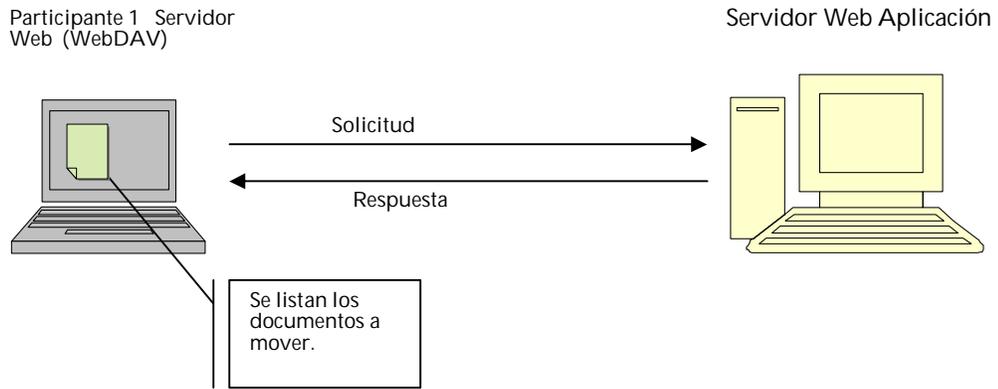
Para lograr la transferencia de documentos de un repositorio a otro, los cuales se pueden localizar en diferentes computadoras, es necesario que cada una cuente con un servidor *Web* que soporte el protocolo WebDAV. Dichos servidores ya se discutieron en el capítulo 3.

La arquitectura usada en la implementación consta de un servidor *Web*, donde se almacena la aplicación, es decir, en este servidor *Web* se encuentran las páginas html y los servlets utilizados para la operación "*move*".

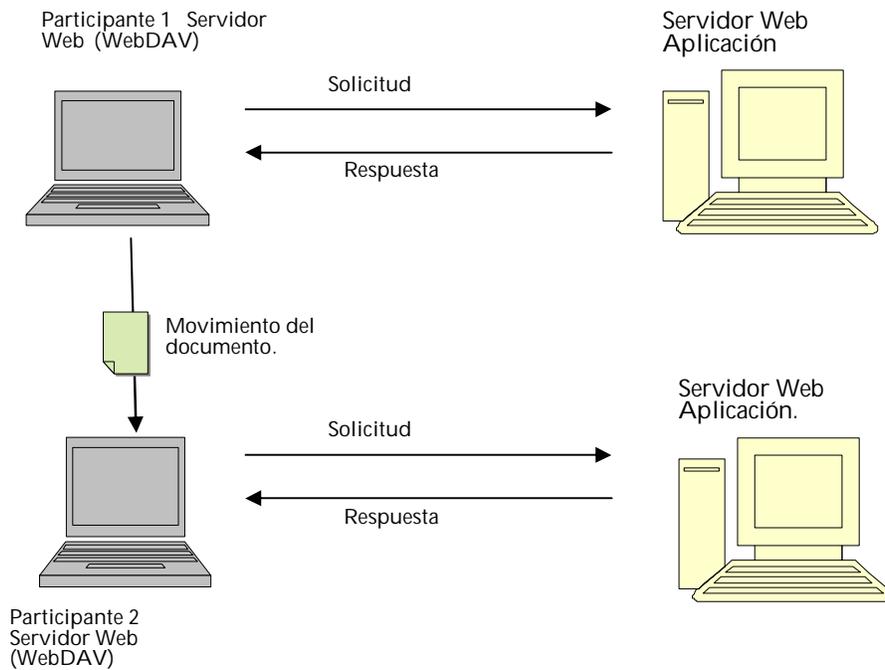
Cada una de las computadoras participantes accesan a la aplicación haciendo solicitudes al servidor *Web*. Dichas peticiones envían parámetros al Servidor *Web* para realizar las operaciones que recuperen el recurso a mover, es decir se listan los documentos disponibles, ver figura 6.6. a). Con los documentos se lleva a cabo la operación "*move*". El servidor *Web* responde al cliente realizando el movimiento del documento desde la computadora del participante 1 hacia la computadora del participante 2, como se ilustra en la figura 6.6. b). Por lo tanto los recursos que se transmiten de una computadora a otra, en realidad están ubicados en las máquinas participantes y es entre éstas máquinas donde se realiza la gestión del recurso, como se ilustra en la figura 6.6.

Las computadoras participantes son identificadas por medio de su nombre por el servidor *Web*. Cada uno de las computadoras participantes debe tener instalado un servidor *Web* que soporte WebDAV. Dentro de la estructura de directorios de cada uno de los servidores *Web* participantes, se debe configurar en el directorio de publicación el directorio WebDAV, permitiendo las operaciones de lectura y escritura para los demás participantes y es en este directorio donde se creará la estructura de directorio de acuerdo al participante, y donde se almacenaran los documentos que interactúan en el proceso.

Esta arquitectura permite tener diferente número de participantes que intervengan en el proceso, ya que la aplicación se encuentra almacenada en un sitio pero el recurso solo se almacena en la computadora de cada participante. El recurso jamás se almacena, ni siquiera de manera temporal en el servidor *Web*.



a) Solicitud que realiza el participante donde se listan los documentos (recurso a mover)



b) Movimiento del documento de un ordenador a otro.

Figura 6. 6: Arquitectura de la aplicación.

6.2.2 Creación de estructura de directorios y almacenamiento del documento

Una de las características de WebDAV es el soporte para crear y borrar una

colección (directorios), agregando y removiendo miembros a/desde una colección (directorio) y para listar los miembros de una colección (directorio). Ver figura. 6.7.

El método de WebDAV utilizado para la creación de ésta colección es *MKCOL*. La sintaxis del método es:

```
MKCOL directorio http/1.1  
Host: servidor
```



Figura 6. 7: Miembros de una colección

El proceso por medio del cual se realiza la creación de la estructura de directorios, para almacenar los documentos del alumno, es el siguiente:

- a. Para la creación del documento de inscripción se accesa a la página "*inscribir.htm*", donde se introduce el número de matrícula del alumno, parámetro que es enviado a la clase "*Buscar_datos*", la cual recupera los datos del alumno, que serán incluidos en el documento. Los datos que se obtienen se almacenan en variables *session* y a su vez son pasados como parámetros a la clase "*crear_dir*".
- b. La clase "*crear_dir*" crea la estructura de directorios, ilustrada en la figura 6.7, utilizando el método MKCOL. Estos directorios se crean en el siguiente orden:
 - i. Primero se crea el directorio del alumno, el cual se referencia por el numero de matricula. Para poder crear este directorio, es importante que exista el directorio padre *alumnos*.

```
MKCOL /Webdav/alumnos/12369 HTTP/1.1  
Host: roma  
Content-Type: text/plain
```

- ii. Una vez creado el directorio del alumno "*12369*", se crean dentro de ésta colección, los directorios: *enviar*, *rechazados*, y *expediente*:

```
MKCOL /Webdav/alumnos/12369/enviar HTTP/1.1
Host: roma
Content-Type: text/plain
```

```
MKCOL /Webdav/alumnos/12369/rechazados HTTP/1.1
Host: roma
Content-Type: text/plain
```

```
MKCOL /Webdav/alumnos/12369/expediente HTTP/1.1
Host: roma
Content-Type: text/plain
```

- c. Una vez creada la estructura del directorio, se crea el documento de inscripción. Los parámetros del número de matrícula así como los demás datos del alumno obtenidos a través de la clase "Buscar_datos" (punto a) son enviados a la clase "crear_inscripcion", la cual almacena el documento en el directorio "enviar", utilizando el método *PUT*, de la siguiente forma:

```
PUT /Webdav/alumnos/12369/enviar/12369.txt HTTP/1.1
Host: servidor
Content-Length: longitud
Content-Type: text/plain
"contenido"
```

Finalmente se llega a la construcción de la estructura de directorio para el alumno, ilustrado en la figura 6.7. Para el resto de los participantes se puede llevar a cabo la secuencia de operaciones descritas antes de creación de la estructura, pero en el desarrollo hecho en esta tesis, la estructura de directorio del resto de participantes es configurado en el momento de la configuración del servidor *Web*.

Se puede revisar el código de cada una de las clases que intervienen en esta creación de directorios en el apéndice A.

6.2.3 Movimiento del documento

El protocolo WebDAV implementa el método "move" que lleva acabo el movimiento de un recurso, desde un directorio origen a un directorio destino. La sintaxis de este método es el siguiente:

```
Move /Webdav/alumno/123698/enviar/123698.txt
http/1.1.
Host: roma
Destination: http://roma/Webdav/tutorestudio/
```

Podemos observar en esta sintaxis, y por las mismas propiedades del método, que esta operación (*move*) sólo es posible de un directorio a otro, contenidos en la misma computadora (servidor), por lo tanto para fines de nuestro objetivo, el método "*move*" no es posible utilizarse. Sin embargo aprovechando las características del protocolo, que también implementa los métodos: *GET*, *PUT*, Y *DELETE*, que en conjunto y en ese orden se utiliza para realizar la misma operación "*move*" logrando mover un documento de una computadoras origen a una computadoras destino.

El primer paso es reconocer que opción es seleccionada por el participante, de esta manera se sabe a que computadoras y a que directorio se consultara para la obtención de los documentos a visualizar por el participante. En esta operación se utiliza la etiqueta `<a>`, donde se le pasan dos parámetros: *user* y *host*, como se muestra a continuación:

```
<a href= "http://url/op?user=participante&host=maquinaparticipante">
```

Donde:

- *url/op*, es la dirección donde se encuentra el archivo a utilizar para la acción y *op*, es la opción a realizar por el participante: *Enviar*, *Rechazados*, *Por revisar*, *Revisados*, *Aprobados*, según sea el caso
- *user*, es el usuario quien realiza la acción.
- *host*, es la dirección (nombre) de la máquina del usuario que realiza la acción.

Con estos valores, lo primero que hace es llamar a la página especificada en el url, y según el *user* y *host*, el sistema sabe a que máquina y en que repositorio tiene que ir a revisar. Si existen archivos se muestran al usuario con las posibles acciones que puede realizar según la opción elegida por el usuario.

Por ejemplo:

Si *user* = *Tutor de estudios*, *host*: *atlantidas*, y la opción elegida fue: "Por revisar", esto es:

```
<a href= "http://computacion/revisar_doc?user=tutorestudio&host=atlantidas">Por  
Revisar</a>
```

Estos dos parámetros se envían a la clase "revisar_doc" y utilizando el método GET, de la forma siguiente:

```
GET http://atlantidas:8080/Webdav/tutorestudio/revisar/
```

Se obtienen los documentos que se encuentren en el directorio de "revisar" y que esperan ser autorizados o rechazados.

DOCUMENTO	ACCIONES	
Inscripcion	Enviar documento al Coordinador Académico (Autorizar)	Rechazar documento al alumno

Cualquiera de estas dos acciones mostrada en la tabla anterior, es un hipervínculo, que va a realizar el movimiento del documento de una computadora origen a una computadora destino. La forma de la etiqueta <a> es la siguiente:

```
<a href="url?RSource=repositorio\archivo&HSource=servidor&PSource=puerto&RTarget=repositorio\archivo&HTarget=servidor&PTarget=puerto">
```

Donde:

- *url*, es la dirección del *servlet* utilizado para realizar el movimiento del archivo: "Mover_archivo"
- *RSource*, es el recurso origen, es decir, el archivo a mover al repositorio destino. Está constituido de la ruta y el nombre de directorio.
- *HSource*, es el servidor origen, a donde va a ir a buscar el RSource.
- *PSource*, es el puerto utilizado por el servidor origen.
- *RTarget*, recurso destino, es decir, el directorio a donde se moverá el archivo.
- *HTarget*, servidor destino, es decir, la máquina donde se depositará el documento.
- *PTarget*, puerto utilizado por el servidor destino

Todos estos datos se envían al *servlet* (Mover_archivo) el cual realiza el

movimiento del documento, dicho movimiento se lleva a cabo bajo las siguientes operaciones:

- a. Primero obtiene el archivo, de la computadora origen a través del método *GET*, utilizado de la siguiente forma:

```
GET RSource http/1.1
```

Se almacena el contenido del archivo, en una "cadena".

- b. Una vez obtenido el documento se coloca en la computadora destino, utilizando el método *PUT*, como se muestra a continuación:

```
PUT RTarget http/1.1
Host: HTarget
Content-Length: longitud del contenido
Content-Type: text/plain
"cadena"
```

- c. Por último se procede a eliminar el archivo del origen, por medio del método *DELETE*:

```
DELETE RSource HTTP/1.1
Host: HSource
```

Éste es el procedimiento que lleva a cabo el movimiento de un documento, desde un repositorio origen a un repositorio destino, situados en diferentes computadoras. Independientemente de la operación elegida (*revisar, rechazados, aprobados, enviar*) y acción tomada (*autorizar, rechazar*), el punto esencial es saber en donde se encuentra el documento y en donde se va a depositar. A nivel código la forma en como se envían los parámetros y se envían las solicitudes al servidor *Web* para ejecución de los métodos utilizados se puede revisar en el apéndice A.

6.2.4 Descripción de clases

La aplicación se desarrollo en su mayoría por páginas *Web* dinámicas, con las cuales los participantes realizan sus tareas dentro del proceso, como es la creación y envío del documento de inscripción por parte del alumno y la revisión del documento para su autorización o rechazo por parte del resto de los participantes. En ésta implementación se uso la tecnología Java Servlets. En la figura 6.8., se ilustran las diferentes clases utilizadas en la implementación.

6.3.4.1. Creación de estructura de directorios y almacenamiento de archivo

En esta sección se describen las clases utilizadas para la creación de la estructura de directorios del alumno, así como también la clase que realiza el almacenamiento del documento (archivo).

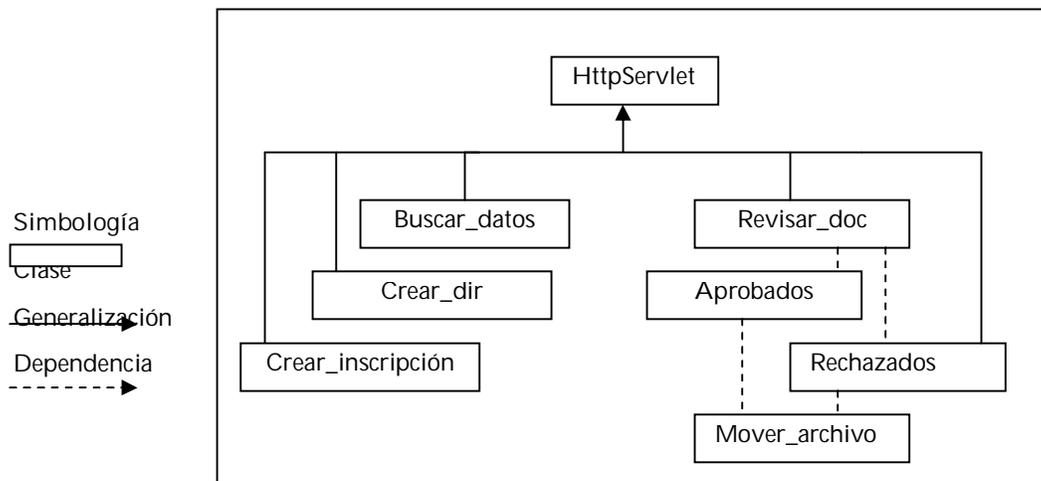


Figura 6. 8: Diagrama de clases

Clase Buscar_datos

Esta clase recibe como parámetro el número de matrícula del alumno, por medio de la cual realiza una búsqueda en la base de datos de *alumnos* obteniendo los datos generales del alumno que serán incluidos en el documento de inscripción.

Buscar_datos	
<p>Responsabilidades</p> <ul style="list-style-type: none"> • Recibe el número de matrícula introducido en la página html "Inscribir" • Mostrar al usuario los datos obtenidos del alumno • Indicar al usuario si se genera el documento. 	<p>Colaboradores</p> <ul style="list-style-type: none"> • Crear_dir • Crear_inscripcion

Clase Crear_dir

Recibe la información de la clase *Buscar_datos*, de donde se toma el número de matrícula para crear la estructura de directorios para almacenar la documentación correspondiente al alumno.

Crear_dir	
<i>Responsabilidades</i>	<i>Colaboradores</i>
<ul style="list-style-type: none"> • Recibe la información del alumno (matrícula) obtenida de la clase <i>Obtener_datos</i>. • Crea la estructura de directorios para el alumno. 	<ul style="list-style-type: none"> • <i>Buscar_datos</i> • <i>Crear_inscripcion</i>

Clase *Crear_inscripcion*

Recibe la información del alumno como, matrícula, datos generales y necesarios para la creación del documento de inscripción.

Crear_inscripcion	
<i>Responsabilidades</i>	<i>Colaboradores</i>
<ul style="list-style-type: none"> • Recibe la información del alumno • Crea y almacena el documento de inscripción en el directorio "enviar". • Notifica al alumno la creación del documento. 	<ul style="list-style-type: none"> • <i>Crear_dir</i>

6.3.4.2. Movimiento del documento

A continuación se explican las clases utilizadas para acceder a los directorios *revisar*, *rechazados*, *aprobados*, de cada uno de los participantes de acuerdo su función dentro del proceso así como la clase mueve el documento de inscripción de un participante a otro.

Clase *revisar_doc*

Esta clase recibe como parámetros el usuario y el servidor para desplegar los archivos (documentos) a ser revisados por el participante. Mostrándose con las dos acciones a llevar a cabo con el documento: autorizar o rechazar.

Revisar_doc	
<i>Responsabilidades</i>	<i>Colaboradores</i>

<ul style="list-style-type: none"> • Recibe la información: usuario, servidor, a través del url, de la etiqueta <a> • Accesa al directorio "revisar" del participante • Despliega los documentos a ser revisados por el participante. • Muestra las acciones a tomarse con respecto al documento: autorizar o rechazar. 	<ul style="list-style-type: none"> • Mover_archivo
---	---

Clase rechazados

Accesa al directorio "rechazados" del participante, utilizando los parámetros de usuario, servidor, para reconocer que participante es y que máquina debe acceder, para mostrar todos aquellos documentos que le fueron rechazados al participante junto con las acciones a realizar.

Rechazados	
<p><i>Responsabilidades</i></p> <ul style="list-style-type: none"> • Recibe la información: usuario, servidor, a través del url, de la etiqueta <a> • Accesa al directorio "rechazados" del participante • Despliega los documentos que le fueron rechazados al participante. • Muestra las acciones a tomarse con respecto al documento. 	<p><i>Colaboradores</i></p> <ul style="list-style-type: none"> • Mover_archivo

Clase aprobados

Recibe los parámetros usuario y servidor, con los cuales sabe que participante es el que está realizando el proceso y accesar al servidor y directorio de "aprobados" desplegando todos los documentos que ya fueron aprobados por todos los participantes del proceso.

Aprobados	
<p><i>Responsabilidades</i></p> <ul style="list-style-type: none"> • Recibe la información: usuario, servidor, a través del url, de la etiqueta <a> • Accesa al directorio "aprobados" del participante • Despliega los documentos que ya fueron aprobados 	<p><i>Colaboradores</i></p> <ul style="list-style-type: none"> • Mover_archivo

- | |
|---|
| <p>por los participantes.</p> <ul style="list-style-type: none"> • Envía el documento al expediente del alumno |
|---|

Clase Mover_archivo

Esta clase es la principal debido a que lleva a cabo la operación "move". Éstarecibe los parámetros que se necesitan para saber el origen y destino del documento, dichos parámetros se envían por las tres clases anterior descritas: *revisar_doc*, *rechazados*, *aprobados*.

Mover_archivo	
<p>Responsabilidades</p> <ul style="list-style-type: none"> • Recibe la información: origen y destino en la etiqueta <a> , de donde y hacia donde se moverá el documento • Ejecuta los métodos <i>GET, PUT, DELETE</i>, en ese orden para el movimiento del documento • Envía el documento al siguiente participante, independientemente de la acción tomada: <i>autorizar o rechazar</i> 	<p>Colaboradores</p> <ul style="list-style-type: none"> • Mover_archivo

6.4. Sinopsis

En este capítulo se muestra el diseño de la aplicación, es decir, las interfaces que se crean para representar el escritorio de cada uno de los participantes. Se explica la arquitectura de la aplicación, es decir, donde se encuentra almacenada las clases y paginas html que componen la aplicación. Finalmente se explica como se lleva a cabo la implementación de esta aplicación, es decir, cual fue el mecanismo con el cual se realiza la operación "move".

Capítulo 7 Uso del sistema: SGD-Web

Este capítulo describe el manual de usuario del sistema SGD-Web (Sistema de Gestión de documentos en el *Web*) desarrollado en esta tesis. Se describe como inicializar el servidor *Web*, la ejecución de la aplicación. Se explica como cada usuario accesa a la aplicación, se ilustran las opciones que le son mostradas al usuario para la realización de sus actividades dentro del proceso de inscripción.

7.1. Requerimientos del sistema

Para el uso del sistema se deben de cumplir los siguientes requerimientos

- Sistema Operativo Windows 2000 o XP
- Microsoft Access
- Servidor *Web* Apache Tomcat 4.0
- Internet Explorer 5.0 o superior

7.2. El Servidor *Web*

El servidor *Web* utilizado es Apache Tomcat 4.0, el cual se encuentra disponible en el sitio de Jakarta Project. La utilización de éste servidor, se debe al soporte que tiene del protocolo WebDAV, así como también el manejo de la tecnología Java Servlets.

Una vez que se instala el servidor *Web* Apache Tomcat 4.0, se crea un directorio de "webapps", en donde se realiza la publicación de los sitios *Web*. En este directorio se instala la aplicación, es decir, el directorio que contiene tanto las páginas html, como los servlets. El directorio de nuestra aplicación es "SGD-Web", en donde en el directorio raíz se encuentran todas las páginas html y dentro de éste el directorio WEB-INF donde se almacenan todas las clases de los servlets. Dentro del directorio raíz, se almacena la base de datos de nombre "alumnos". Todo esto se configura en el Servidor *Web*. Esta estructura se muestra en la figura 7.1.

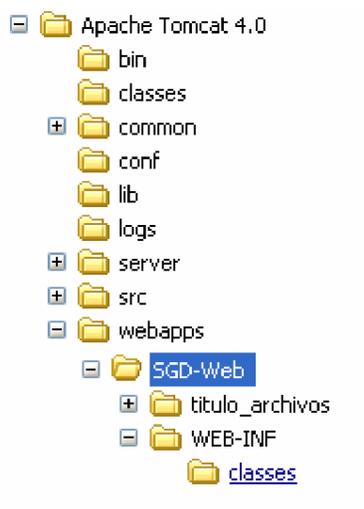


Figura 7. 1. Directorio de Publicación de Tomcat

Para iniciar el servidor *Web Tomcat 4.0*, se debe de ejecutar la opción de Start Tomcat 4.0, la cual se encuentra en el grupo de programas Apache Tomcat 4.0, tal y como se muestra en la figura 7.2.

Al ejecutar el programa se obtiene una ventana de línea de comandos, donde se indica la ejecución del servidor *Web*, la cual se muestra en la figura 7.3.

Para cada uno de las computadoras de los participantes se debe de igual manera instalar el servidor Apache Tomcat, el cual soporta WebDAV. Dentro del directorio de WebDAV se debe de crear la estructura de directorios del participante, la cual ya se explico en el capítulo 4.

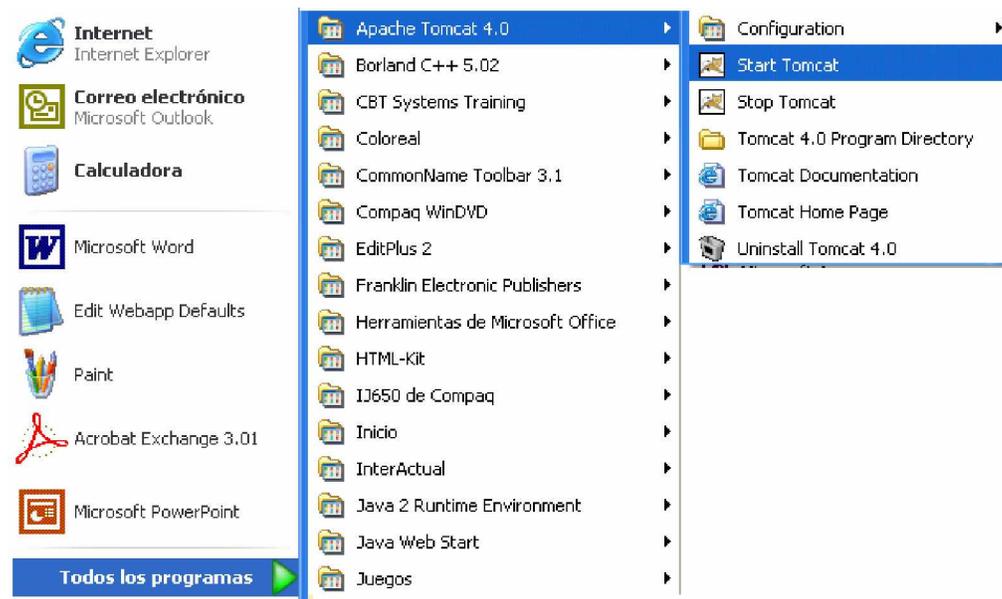


Figura 7. 2. Opción que inicializa el servicio de Tomcat.

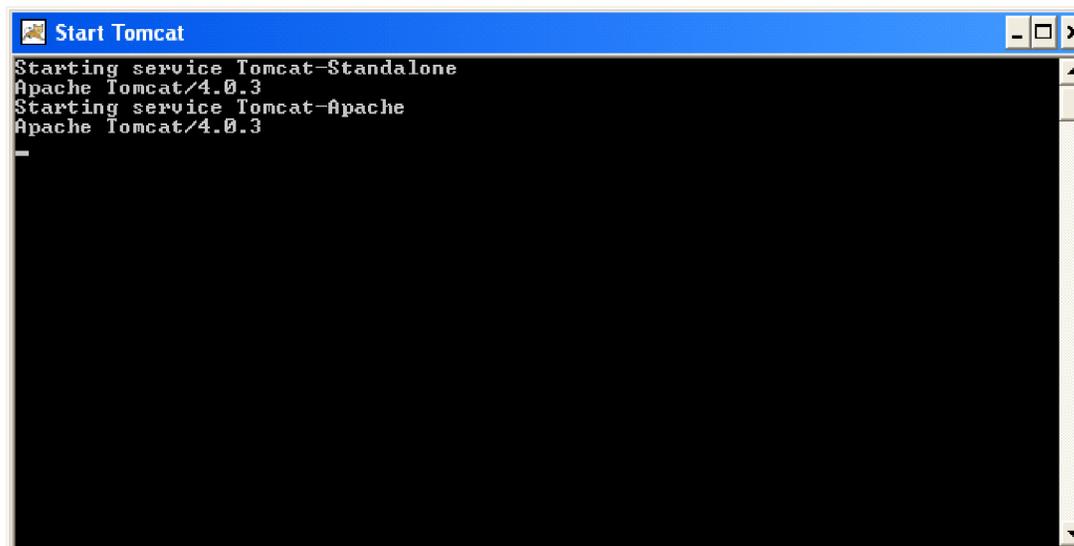


Figura 7. 3. Ventana que indica el servicio inicializado de Tomcat.

7.3. Uso del sistema

A continuación se describe el uso de la aplicación. Para inicializar la aplicación se tiene que abrir el explorador de Internet Explorer 5.0 o superior, en donde en la barra de direcciones se debe de colocar la siguiente dirección.

<http://roma:8080/SGD-Web/index.htm>

La página de inicio se muestra figura 7.4. En ésta página se introduce el nombre de usuario y la contraseña para entrar al sistema, en donde se reconoce que usuario es y mostrar el menú de opciones correspondiente a las actividades que realiza dentro del proceso de inscripción.



Figura 7. 4. Página de inicio del sistema.

7.4.1. Menú del alumno

Para el caso del alumno se muestra el menú de la figura 7.5., en donde puede realizar cualquiera de las 3 opciones que tiene para la opción de inscripción: *Crear inscripción, enviar, rechazados*. Así como también puede acceder su "expediente" donde se encuentran sus documentos ya autorizados.

Crear Inscripción

Al seleccionar esta opción se despliega la página donde el alumno introducirá su número de matrícula, ver figura 7.6., y al pulsar el botón de buscar, se le presentarán en pantalla todos los datos que corresponde a su información (Figura 7.7.) para ser incluidos en su documento de inscripción creándose éste al momento de pulsar el botón de generar documento, y se le notifica al alumno que el documento ha sido creado y que revise su escritorio (ver figura 7.8).



Figura 7. 5. Menú del alumno.

Enviar

Esta opción muestra al alumno los documentos que tiene que enviar a su tutor de estudios para que le sean autorizados o rechazados, ver figura 7.9. , al seleccionar el hipervínculo de "Documento a enviar al tutor de estudios", se envía el documento al escritorio del tutor, y se despliega nuevamente la pagina del escritorio del alumno, pero mostrado los documentos restantes a enviar a su Tutor de estudios, como se muestra en la figura 7.10., para este caso ya no existe documentos disponibles para enviar al tutor de estudios.



Figura 7. 6. Pantalla donde introducirá el número de matricula el alumno



Figura 7. 7. Datos del documento de inscripción.



Figura 7. 8. Notificación de creación de documento.

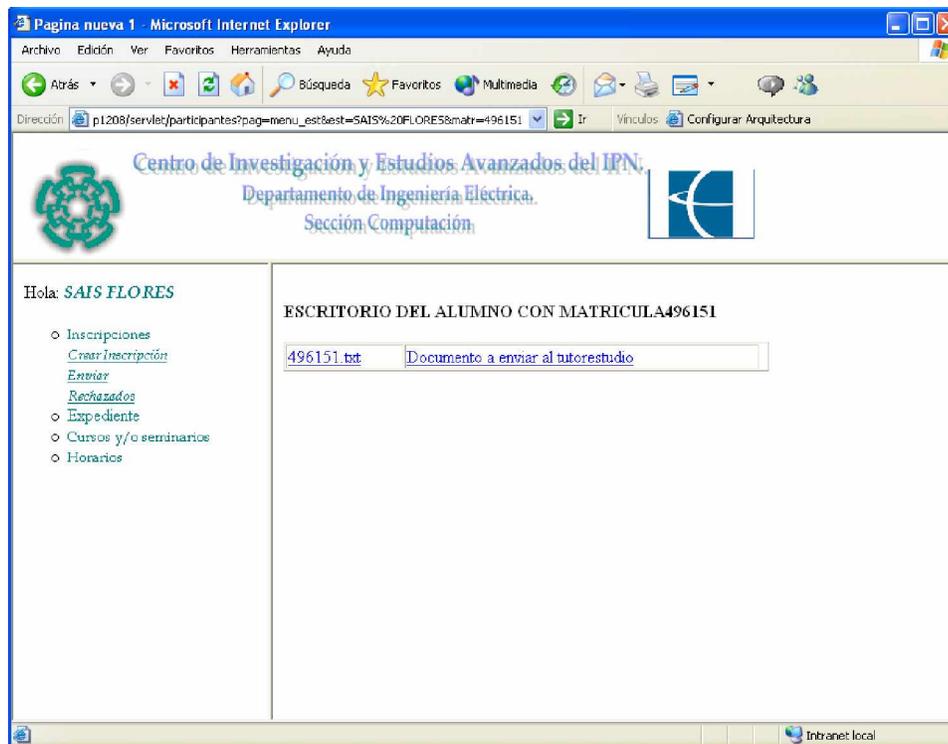


Figura 7. 9. Documentos a enviar por el alumno a su Tutor de estudios.

Rechazados

Al seleccionar esta opción el alumno podrá visualizar los documentos que su tutor de estudios le ha rechazado, figura 7.11., y que tendrá que corregir para después volver a ser enviado a su tutor de estudios y cuyo documento corregido estará disponible en la opción de "enviar".

Expediente

En esta opción el alumno tiene disponible los documentos que ya han sido autorizados por los participantes del proceso. El proceso ha llegado a su fin de manera satisfactoria. Ver Figura 7.12.



Figura 7. 10. Escritorio del alumno, sin documentos disponibles.



Figura 7. 11. Documentos rechazados al alumno.

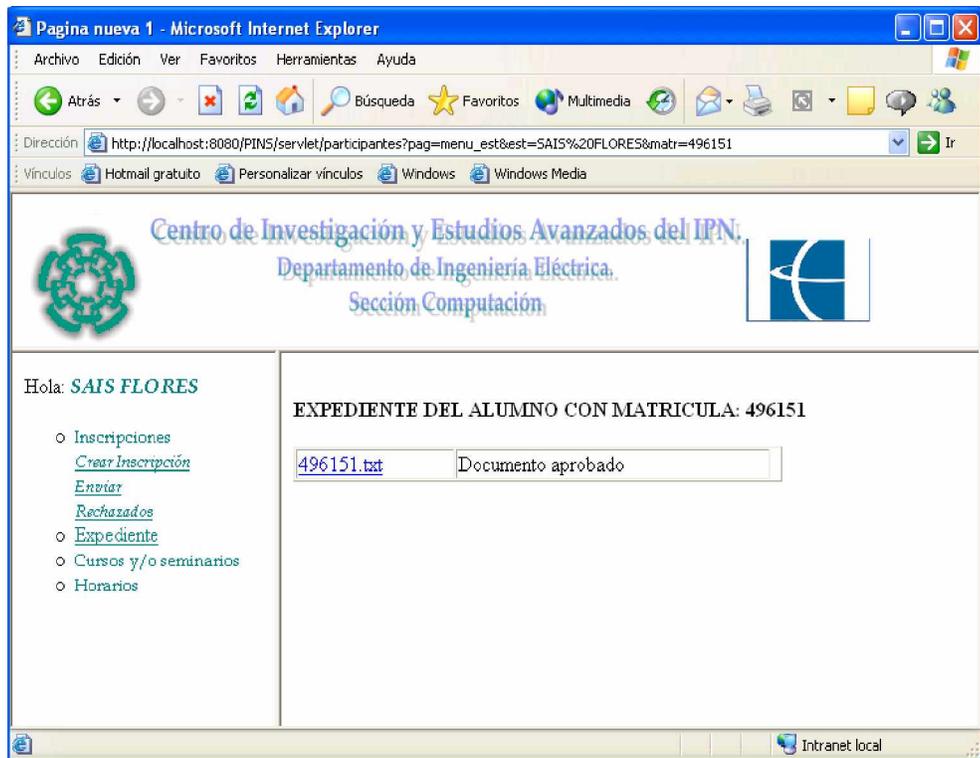


Figura 7. 12. Muestra el documento en el expediente del alumno.

7.4.2. Tutor de estudios

Este participante básicamente solo tiene dos opciones en el proceso de inscripción: Por revisar y rechazados, como se muestra en la figura 7.13.

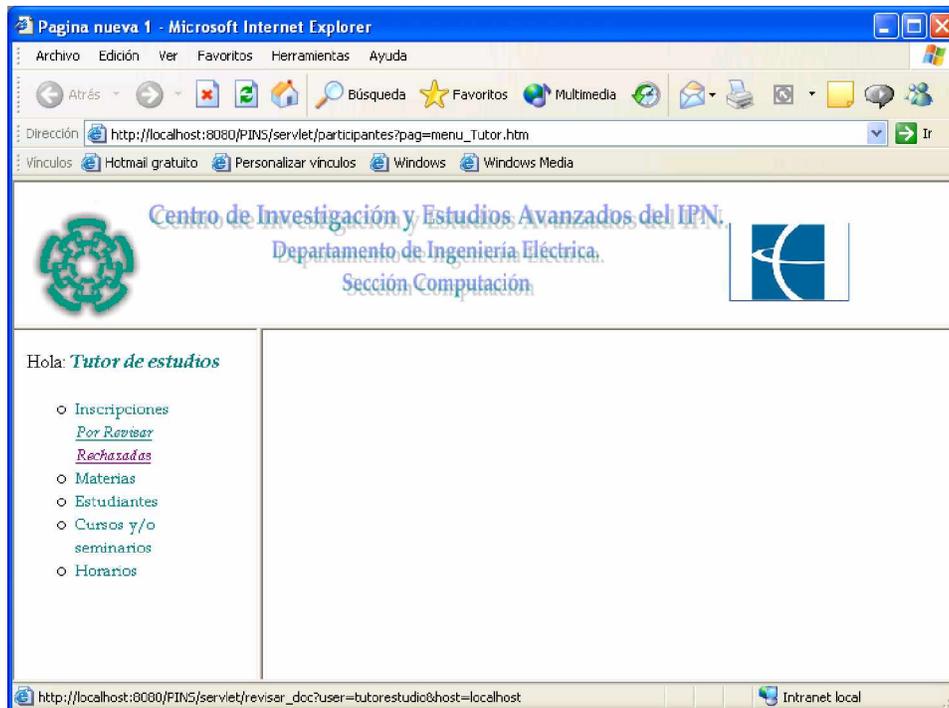


Figura 7. 13. Menú del Tutor de estudios.

Por revisar

Con esta opción el Tutor de estudios puede visualizar todos los documentos que los estudiantes a su cargo le envían para que él los autorice o los rechace, ver figura 6.14. Si el tutor de estudios autoriza la inscripción tendrá que seleccionar el hipervínculo "Documento autorizar (enviar al coordinador académico)", en caso contrario selecciona al hipervínculo "Rechazar documento al alumno", cualquiera que sea su decisión. Una vez realizado el movimiento del documento, el sistema volverá a mostrar la página de la opción "Por revisar", desplegando los documentos aún disponibles para su revisión. Ver figura 7.15.



Figura 7. 14. Escritorio del Tutor de Estudios. Autoriza o rechaza documentos

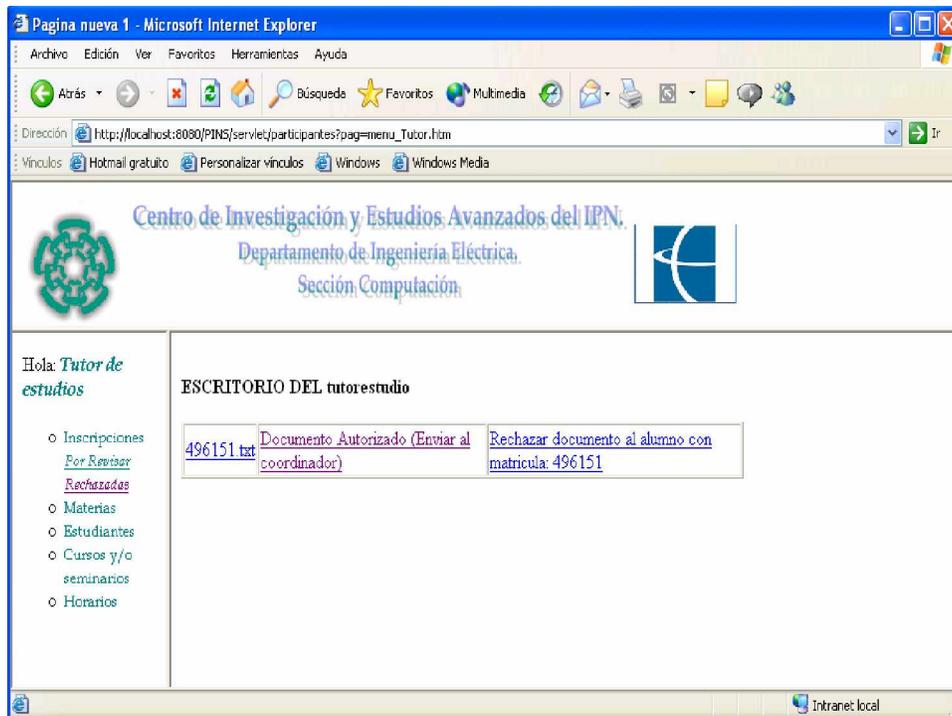


Figura 7. 15. Escritorio del Tutor de Estudios, después de autorizar un documento.

Rechazados

Seleccionar en esta opción se visualizan los documentos que fueron rechazados por el coordinador académico y que el tutor de estudios tendrá que revisar bien enviar al alumno para su corrección. Ver figura 7.16.

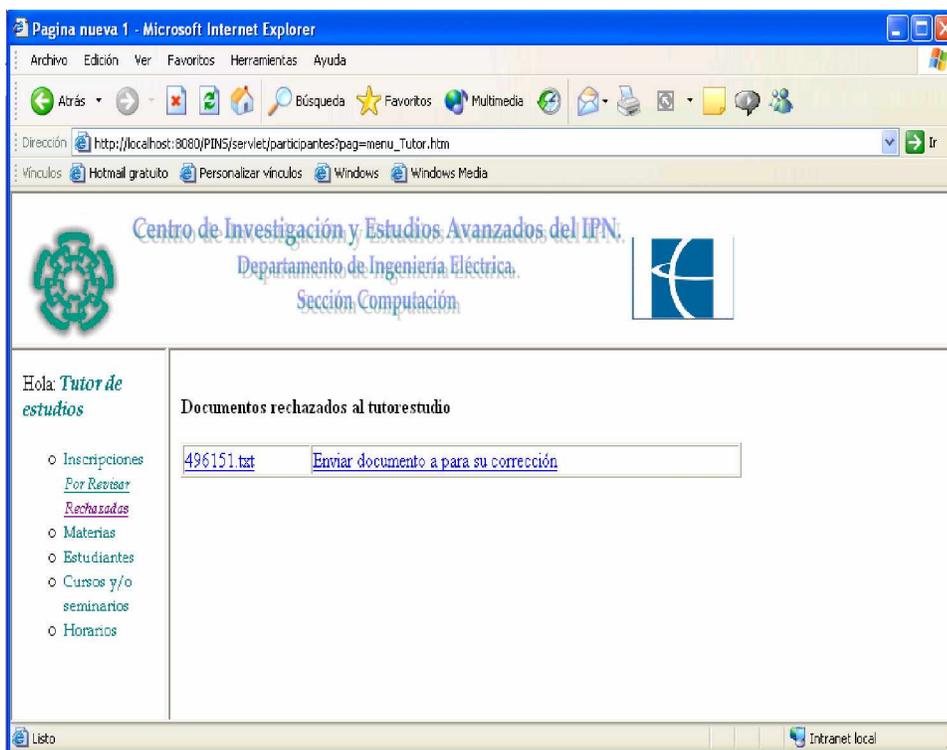


Figura 7. 16. Documentos rechazados al Tutor de Estudios.

7.4.3. Coordinador académico

Este participante básicamente tiene las mismas opciones en el proceso, que el tutor de estudios: Por revisar, rechazados. Ver figura 7.17.



Figura 7. 17. Menú del Coordinador Académico.

Por revisar

Al seleccionar esta opción el coordinador académico podrá visualizar los documentos que tiene que revisar, para que los autorice o rechace. Ver figura 7.18. En caso de seleccionar el hipervínculo "Autorizar documento (Enviar a la secretaria)" o bien el hipervínculo "Rechazar documento al tutor de estudios", se realizará la operación de mover el documento, visualizándose después la misma página pero con los documentos restantes a revisar por el coordinador académico, como se muestra en la figura 7.19.

Rechazados

Esta opción contiene los documentos que no fueron autorizados por servicios escolares y que la secretaria se encarga de enviárselos nuevamente al coordinador académico, para que éste proceda a su revisión o lo regrese al tutor de estudios. Ver figura 7.20.



Figura 7. 18. Documentos a ser revisados por el Coordinador Académico.



Figura 7. 19. Escritorio del Coordinador Académico, sin documentos a revisar.

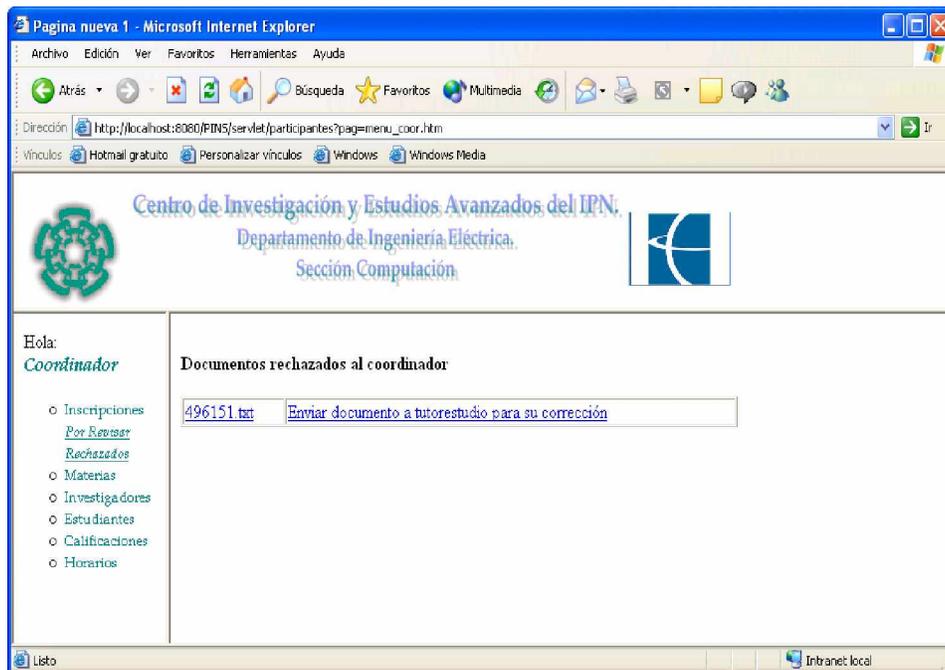


Figura 7. 20. Documentos rechazados al Coordinador Académico.

7.4.4. Secretaria

El papel de la secretaria dentro del proceso es una especie de recolector de documentos, por lo tanto las opciones para este proceso son revisados, rechazados, y aprobados. Ver figura 7.21.

Revisados

Esta primera opción despliega todos los documentos que ya fueron autorizados por tutor de estudios y coordinador académico y que la secretaria enviará a servicios escolares para su revisión y pueda determinar si queda autorizado o rechazado, como se muestra en la figura 7.22.

Rechazados

Al seleccionar al directorio de rechazados y muestra a la secretaria los documentos que fueron rechazados por servicios escolares y que debe de devolver al coordinador académico para que sean corregidos. Ver figura 7.23



Figura 7. 21. Menú de la Secretaria.

Aprobados

Al seleccionar esta opción se muestran los documentos que servicios escolares ha revisado y aprobado. Dichos documentos serán enviados al expediente del alumno. Ver figura 7.24

7.4.5. Servicios escolares

Las opciones que servicios escolares lleva a cabo en el proceso de inscripción es únicamente la revisión de los documentos. Ver figura 7.25.



Figura 7. 22. Escritorio de la Secretaria.



Figura 7. 23. Documentos rechazados al Coordinador Académico.



Figura 7. 24. Documentos aprobados. La secretaria envía al expediente del alumno.

Por revisar

Al seleccionar esta opción se muestran los documentos que le fueron enviados a servicios escolares por parte de la secretaria y los cuales han sido revisados y autorizados por el tutor de estudios y coordinador académico.

Servicios escolares realiza la revisión de estos documentos para autorizarlos o rechazarlos, seleccionado en el hipervínculo correspondiente. Ver figura 7.26



Figura 7. 25. Menú de Servicios Escolares.

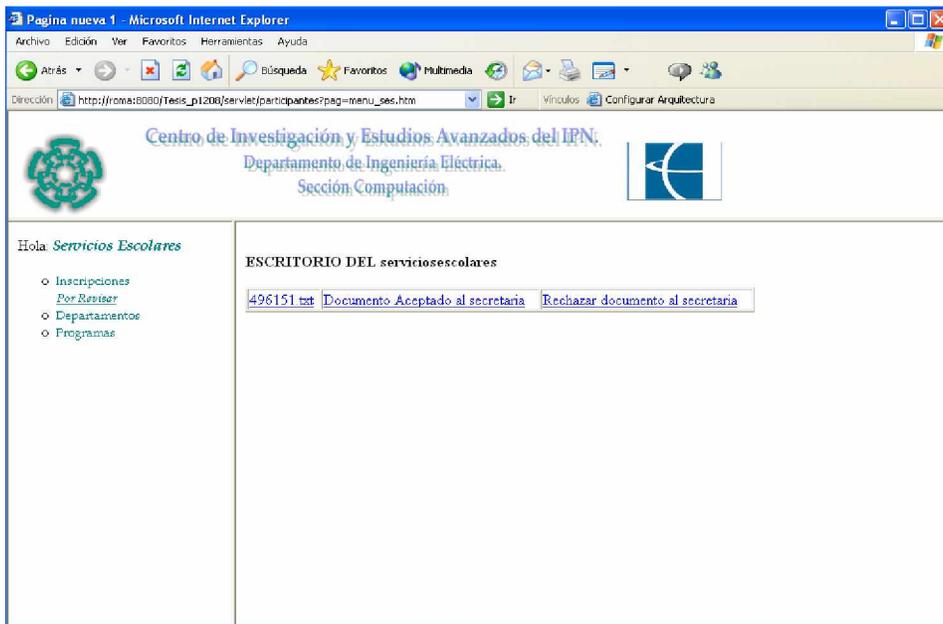


Figura 7. 26. Escritorio de Servicios Escolares.

7.5. Sinopsis

Todas las opciones con las que cuenta el sistema para intercambiar el documento de inscripción de un participante a otro, es decir, como el archivo se mueve de una computadora a otra, se presentaron en este capítulo. Se explico la interacción del participante con el sistema dependiendo de las actividades que realiza dentro del proceso, donde algunos participantes llevan a cabo la misma tarea regidas bajo diferentes reglas o en ocasiones bajo reglas semejantes.

Al implementar el sistema de gestión de documentos, para el caso de estudio: el proceso de inscripción, sobre *Web*. Donde se crean interfaces personalizadas para cada participante, proporciona:

Seguridad para el usuario a través de la autenticación inicial, con nombre de usuario, y proporcionar solo la información correspondiente a dicho usuario para llevar a cabo sus actividades dentro del proceso.

Especificación de sus actividades, al crear un menú con opciones de la lista de actividades, le proporciona al usuario una clara descripción de la actividad a realizar sobre los documentos que se muestran al seleccionar la opción, es decir, para el caso de revisar, es explicito que deberá leer el documento y analizarlo para tomar una decisión sobre el mismo. Para la opción de rechazados volverá a revisar los documentos para saber la causa de rechazo y poder remitir al participante anterior siguiente.

Fácil manejo de los documentos. Al agrupar en un menú de lista de actividades a realizar por el usuario, donde los nombres de las opciones son explicitas de las actividades, le facilita al usuario el poder acceder aquellos documentos a revisar, rechazados, para que pueda concluir con su actividad sobre la entidad.

Conocimiento del flujo de trabajo, al presentar la lista de documentos con las opciones de autorizar o rechazar, especificando en cada una el nombre del participante a quien se remitirá según la decisión tomada.

Al implementar el proceso de inscripción en un sistema interactivo sobre *Web*, y con una autenticación inicial del participante del proceso, permite al usuario tener

un mejor control de las actividades que debe realizar e identificar los documentos según el estado en el que actualmente se encuentran, es decir, si están por ser revisados, o han sido rechazados o finalmente aprobados.

Capítulo 8 Conclusiones

En este trabajo se han presentado y discutido los problemas relacionados con la gestión de documentos. Para resolver algunos de éstos problemas, principalmente aquellos relacionados con la interoperabilidad y facilidad de uso, se definieron en esta tesis las características que debe cumplir un sistema de gestión documental, tanto a nivel departamental como corporativo. El sistema SGD-Web propuesto es un sistema de gestión documental desarrollado como aplicación *Web*, cuya interoperabilidad descansa en el uso de estándares ampliamente aceptados como Web-DAV y cuya facilidad de uso se deriva del entorno de escritorio que ofrece al usuario para organizar la revisión, edición y transferencia de documentos. El sistema SGD-Web ha sido construido sobre WebDAV, una extensión del protocolo de transporte HTTP que ofrece, entre otros, mecanismos para leer, escribir y eliminar documentos entre servidores. Por otra parte, las interfaces de usuario HTML que ofrece SGD-Web siguen la metáfora del escritorio de oficina organizado por bandejas las cuales implementan la funcionalidad básica de mensajería de documentos.

De acuerdo al perfil del participante y a su función dentro del proceso, la definición e implementación del proceso de inscripción se presenta al usuario a través de un sistema interactivo en la *Web* que permite mostrar las diferentes actividades a realizar. De esta forma, SGD-Web provee al usuario del control de las acciones a realizar sobre los documentos puestos en su escritorio. El usuario puede leer entonces los documentos, analizarlos y evaluarlos para tomar decisiones sobre su contenido, lo cual garantiza que la acción tomada sobre el documento sea consistente y confiable. La disponibilidad del documento es el evento que inicia las acciones, no se podrá proceder con la siguiente fase del proceso en tanto no se concluya la que le precede en la secuencia de acciones.

Al desarrollar la tesis se han cumplido los objetivos planteados:

1. Implementar las operaciones básicas para la gestión de documentos. Las operaciones básicas de gestión de documentos como recuperar, almacenar, eliminar y mover, se implementaron directa o indirectamente por combinación de los métodos *GET*, *PUT* y *DELETE* del protocolo WebDAV.
2. Facilitar la colaboración entre los participantes del sistema de gestión documental. La gestión de documentos se simplifica enormemente al usuario al

quedar ocultos los parámetros de las operaciones en el elemento <a>, los cuales envían los métodos *GET*, *PUT* y *DELETE* al servidor WebDAV para realizar las operaciones correspondientes.

A continuación discutiremos algunos aspectos importantes relacionados con estos logros.

Con relación a la implementación de las operaciones básicas de un sistema de *workflow*, al realizar la transferencia de documentos de una computadora a otra por la aplicación de los métodos *GET*, *PUT* y *DELETE*, en ese orden, podemos encontrar el problema que al estarse llevando a cabo la petición de estos métodos se pierda la conexión, lo cual daría lugar a las inconsistencias que surgen por el manejo de copias múltiples. Por ejemplo, si la conexión se pierde después de aplicar la operación *PUT* y antes de aplicar la operación *DELETE*, provocaría que se tenga una copia del documento en dos sitios, lo cual puede dar lugar a problemas de identificación y actualización del documento. Una posible solución requiere el uso de mecanismos similares aquellos usados en los sistemas de colección de basura. Considerando aquellas copias no eliminadas de los documentos como elementos que deben eliminarse tan pronto como sea posible, SGD-Web requiere de mecanismos que permitan tanto identificar a aquellos documentos que han participado en una operación como determinar si la operación fue concluida exitosamente. De esta forma, para aquellas operaciones no concluidas satisfactoriamente, se podría recuperar la lista de los documentos involucrados con el fin de buscar y eliminar las copias inválidas. Este mecanismo haría que operaciones como *MOVE* se convirtieran en operaciones atómicas, lo cual podría hacer de SGD-Web un sistema transaccional.

Con relación a facilitar la colaboración, la simplificación de un proceso de gestión documental que involucra a varios participantes es un problema importante porque afecta directamente tanto a la productividad como a la disminución de errores. SGD-Web aumenta la productividad porque simplifica la conducción del *workflow* mediante interfaces *Web* sencillas, desligando a los participantes de tareas tediosas que demandan tiempo. Aunque en esta tesis se ha desarrollado un sistema estático con participantes y rutas de documentos predefinidos, *workflows* más flexibles y capaces se pueden obtener si el sistema se combina con tecnologías de páginas dinámicas que permitan modificar dinámicamente ya sea a los participantes como a las rutas de los documentos. Podría calificarse a esta forma de *workflow* como dinámico o reconfigurable. No es difícil reconocer la importancia que puede tener este tipo de *workflow*, ya que la gran mayoría de los escenarios de automatización están sujetos a los continuos cambios en la disponibilidad de recursos (p.ej., la falta

temporal o definitiva de personal lo que resulta en la modificación de las rutas de documentos) que se tienen que realizar para ajustar la conducción de los procedimientos administrativos tanto departamentales como corporativos. De ésta forma este trabajo contribuye con un mecanismo simple que puede ayudar a la automatización parcial de procedimientos administrativos orientados a documentos, en donde los documentos se consideran elementos básicos para la representación, el intercambio de información y la coordinación de actividades.

El caso de estudio abordado en este trabajo es un *proceso de inscripción a un programa de estudio*, el cual se puede clasificar como un *workflow* de colaboración continuo. El proceso de inscripción se puede considerar naturalmente con un sistema de gestión documental, ya que se remite el documento de inscripción para su autorización de un participante a otro.

El sistema SGD-Web demuestra las enormes posibilidades de aplicación que puede ofrecer un sistema de gestión documental con base en la *Web*. A continuación presentaremos algunas extensiones posibles de este trabajo.

8.1. Trabajo futuro

El trabajo como ya se mencionó es útil dependiendo de las condiciones de uso. Existen mejoras que se pueden emprender como trabajo futuro o con las que se pueden complementar con otros trabajos, entre los cuales se puede mencionar:

- La incorporación de métodos transaccionales para definir operaciones atómicas.
- La integración de un sistema de notificación a cada participante de la existencia de recurso a ser revisados.
- La implementación de firmas electrónicas para cada uno de los participantes, de tal forma que se llegue a una automatización completa del documento.
- La integración de un sistema gráfico para la definición del flujo de información que permita seleccionar a los participantes o modificar las rutas del flujo de documentos.

Al igual que como se hizo en el caso de estudio, se puede implementar este trabajo en los procesos de inscripción de otras instituciones, e incluso hablando de

diferentes áreas de la industria, entre las cuales podemos mencionar, la industria de las telecomunicaciones y la industria bancaria, entre otras. Hoy en día la industria busca automatizar en gran medida los procesos de negocios para aumentar la productividad a la vez que se busca reducir los costos de operación. Por eso consideramos que el aporte de este trabajo de tesis puede ser importante para lograr estos objetivos.

Referencias

- 1 Malone T., Crowston K, The Interdisciplinary Study of Coordination, *ACM Computing Surveys*, Vol. 26, No. 1, pp. 87-119, 1994.
- 2 Carriero N., Gelenter D., Linda in context, *Communications of the ACM*, Vol.32, No. 4, 1989.
- 3 <http://www.lcc.uma.es/~gisum/coordinacion/presentac-es.html>, junio 2002.
- 4 Jablonski S., Bussler C. *Workflow Management: Modeling concepts, Architecture and Implementation*, International Thomson Computer Press, 1996.
- 5 David Hollingsworth, The Workflow Management Coalition Specification: the Workflow Reference Model, *Technical Report TC00-1003*, Workflow Management Coalition, Hampshire, UK, January 1995.
- 6 Ralf Steinmetz and Klara Nahrstedt, *Multimedia: Computing, Communications and Applications*, Prentice Hall, 1995.
- 7 Jeff Conklin, Hypertext: An introduction and Survey, *Survey & tutorial series*, September 1987.
- 8 Luc Moreau and Wendy Hall, On the Expressiveness of Links in Hypertext Systems, *The computer journal*, Vol. 41, No.7, 1998.
- 9 José Oscar Olmedo, A Blackboard Architecture for Hypermedia Systems, In *Proceedings of 8th International Congreso on Computer Science Research*. Colima, México, 2001.
- 10 Hughes, Shoffner and Hammer, *Java Network Programming*, Manning Publications Co., May 1999.
- 11 R. Orfali, D. Harkey, J. Edwards, *The Essential Client/Server Survival Guide*, John Wiley & Sons, 1995.
- 12 Juan Manuel Correa Hdez, *SAWF-HW: Sistema de Administración de*

- Workflow basado en el modelo de Coordinación de Hyperworlds*, Tesis de maestría, Departamento de Ingeniería eléctrica, Cinvestav, México, D.F., octubre 2002.
- 13 Victor Hugo Hernández Hdez, *Protocolo para el manejo colaborativo de documentos Web con soporte para desconexión*. Tesis de maestría, Departamento de ciencias Computacionales, Cenidet, Cuernavaca, Mor., junio 2002.
 - 14 J. Slein, F. Vitali, E. Whitehead, and D. Durand, Requirements for a Distributed Authoring and Versioning Protocol for the World Wide Web, RFC 2291, WEBDAV Working Group, February 1998, <http://www.faqs.org/rfcs/rfc2291.html>.
 - 15 Y. Goland, Jr. E.J. Whitehead, Asad Faizi, S.R. Carter, and Del Jensen, Extension for Distributed Authoring and Versioning on the World Wide Web – WEBDAV, *Technical report*, UCI, March 1997.
 - 16 Fredj Dridi, Gustaf Neumann, How to implement Web-based Groupware Systems based on WebDAV, *Proc. of WETICE 99, IEEE 8th Intl. Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise*, Stanford, CA, June 1999.
 - 17 Whitehead, E. J. Jr. & Goland, Y. Y., WebDAV: A network protocol for remote collaborative authoring on the Web, *Proceedings of the Sixth European Conference of Computer Supported Cooperative Work*, Copenhagen, Denmark, pp. 291-310, 1999.
 - 18 Siminiani Mariano, *Intranets, empresa y gestión documental*, McGraw-Hill, pp 3-130, 1997.
 - 19 Fischer Layna & Workflow Management Coalition, *Workflow handbook 2001*. Future Strategies Inc., 2001.
 - 20 W. Appelt, U. Busbach, The BSCW System: A WWW based Application to Support Cooperation of Distributed Groups, GMD-German National Research Centre for Information Technology, Sankt Augustin, Germany, 1996.
 - 21 World Wide Web Consortium, <http://www.w3.org/>

22 Apache Tomcat, <http://jakarta.apache.org/tomcat/index.html>

23 WebDAV, <http://www.webdav.org/>