

**DISEÑO E IMPLEMENTACIÓN DE UNA BIBLIOTECA VIRTUAL BASADA EN
EL BUSCADOR PARA UNIVERSIDADES**

JESSICA LEOTTAU ANDRADE

MONICA PATRICIA PINZON VIVES

JOSE MIGUEL TORRES SCAFF

CORPORACIÓN UNIVERSITARIA TECNOLÓGICA DE BOLIVAR

FACULTAD DE INGENIERIA DE SISTEMAS

CARTAGENA DE INDIAS, D.T. H. Y C.

2001

**DISEÑO E IMPLEMENTACIÓN DE UNA BIBLIOTECA VIRTUAL BASADA EN
EL BUSCADOR PARA UNIVERSIDADES**

JESSICA LEOTTAU ANDRADE

MONICA PATRICIA PINZON VIVES

JOSE MIGUEL TORRES SCAFF

**Trabajo de grado presentado como requisito para optar el título de Ingeniero de
Sistemas**

DIRECTOR

GONZALO GARZON

Ingeniero de Sistemas

CORPORACIÓN UNIVERSITARIA TECNOLÓGICA DE BOLIVAR

FACULTAD DE INGENIERIA DE SISTEMAS

CARTAGENA DE INDIAS, D.T. H. Y C.

2001

Nota de Aceptación

Presidente del Jurado

Jurado

Jurado

Fecha: _____

Cartagena de Indias, D. T. H. Y C., 27 de abril 2001.

Señores:
Departamento de Investigaciones
Corporación Universitaria Tecnológica De Bolívar
Cartagena D. T. H. y C.

Respetado Señores:

Presentamos para su consideración el Proyecto de Grado titulado: **"DISEÑO E IMPLEMENTACIÓN DE UNA BIBLIOTECA VIRTUAL BASADA EN EL BUSCADOR PARA UNIVERSIDADES"** como requisito para optar el título de Ingeniero de Sistemas.

Atentamente,

JESSICA LEOTTAU ANDRADE

MONICA P. PINZON VIVES

JOSE M. TORRES SCAFF

Cartagena de Indias, D. T. H. Y C, 27 de abril 2001.

Señores
Departamento De Investigaciones
Corporación Universitaria Tecnológica de Bolívar
Cartagena D. T. y C.

Respetado Señores:

Por solicitud de los estudiantes MONICA PATRICIA PINZON VIVES, JESSICA LEOTTAU ANDRADE y JOSE MIGUEL TORRES SCAFF, dirigí a satisfacción el proyecto de grado titulado: "**DISEÑO E IMPLEMENTACIÓN DE UNA BIBLIOTECA VIRTUAL BASADA EN EL BUSCADOR PARA UNIVERSIDADES**", como requisito para optar al título de Ingeniero de Sistemas.

Espero que el contenido y las normas aplicadas cumplan con los requisitos exigidos por esta dirección.

Atentamente,

GONZALO GARZON
Ingeniero de Sistemas
Decano Facultad de Sistemas
C.U.T.B.

Cartagena de Indias D. T. H. Y C., 27 de abril de 2001.

ARTICULO 105

La Corporación Universitaria Tecnológica de Bolívar se reserva el derecho de propiedad intelectual de todos los trabajos de grados aprobados y no pueden ser explotados comercialmente sin autorización.

DEDICATORIA

A mi mamá por su esfuerzo y perseverancia.

A mi abuela Eva por todo su cariño.

A mi hermana por su compañía.

A mi padre que de donde este se que esta orgulloso de mi.

A todos que de una u otra forma me dieron su apoyo y amor. Gracias.

Jessica Leottau Andrade

La dedico a mis padres y hermanos por su amor y apoyo incondicional, . Gracias por todo.

Mónica Patricia Pinzón Vives

La dedico a mis padres que con mucho esfuerzo hicieron posible mi realización como Ingeniero de Sistemas.

Jose Miguel Torres Scaff

AGRADECIMIENTOS

Los autores expresamos nuestros mas sentidos agradecimientos a todas las personas que de una u otra forma hicieron posible este proyecto, en especial al Ingeniero Gonzalo Garzon por su dirección en todo el proyecto y al Matemático Moisés Quintana por su asesoría. A la Corporación Universitaria Tecnológica de Bolívar por brindarnos esta oportunidad para trabajar en su desarrollo como Universidad.

LISTA DE TABLAS

Tabla 1. Bases de datos.....	35
Tabla 2. Ventajas de Linux.....	44
Tabla 3. Ventajas de LyX.....	47
Tabla 4. Ventajas de LyX.....	51
Tabla 5. Índice general de áreas.....	57
Tabla 6. Funciones del menú Archivo.....	82
Tabla 7. Clases de documentos.....	90
Tabla 8. Variables del Archivo CONFIG.....	137

LISTA DE FIGURAS

Figura 1. Funcionamiento XSL.....	32
Figura 2. Diferencia entre el SGML, XML y HTML.....	33
Figura 3. Diagrama de la estructura de Sets.....	62
Figura 4. Diagrama de la estructura Book.....	64
Figura 5. Diagrama de la estructura de Bookbiblio.....	69
Figura 6. Diagrama de la estructura de Part.....	70
Figura 7. Diagrama de la estructura de Reference.....	71
Figura 8. Diagrama de la estructura de Preface	72
Figura 9. Diagrama de la estructura Chapter.....	72
Figura 10. Diagrama de la estructura de Appendix.....	73
Figura 11. Diagrama de la estructura de DocInfo.....	73
Figura 12. Diagrama de la estructura de Sections.....	76
Figura 13. Diagrama de la estructura de BridgeHeads.....	77
Figura 14. Diagrama de la estructura de Bibliography.....	79
Figura 15. Diagrama de la estructura de Glossary.....	80
Figura 16. Menú Archivo.....	81
Figura 17. Esquema de la Biblioteca Virtual.....	101
Figura 18. Sistema de directorios.....	105
Figura 19. Interface pagina principal.....	108
Figura 20. Interface de búsqueda.....	109
Figura 21. Formato para crear un nuevo documento.....	118
Figura 22. Formato para cambio de clase de un documento.....	119
Figura 23. Formato para la elección de entornos.....	120

Figura 24. Formato para insertar tablas y listas.....	122
Figura 25. Formato para insertar imágenes.....	125
Figura 26. Formato de tablas.....	126
Figura 27. Formato para el cambio de tipo de letras.....	127
Figura 28. Formato de carta.....	131
Figura 29. Formato para la realizacion de articulos.....	133

LISTA DE ANEXOS

ANEXO A. COMO HACER TUS DOCUMENTOS EN LYX.....	117
ANEXO B. INSTALACION Y CONFIGURACION DE HT://DIG.....	136
ANEXO C. CONFIGURACION E INSTALACION DEL SERVIDOR WEB APACHE	142

TABLA DE CONTENIDO

INTRODUCCION.....	24
1.MARCO TEORICO.....	26
1.1.BIBLIOTECA VIRTUAL.....	26
1.2.EL LIBRO ELECTRÓNICO	28
1.3.HTML Y XML.....	29
1.3.1.HTML, XML, versus SGML.....	32
1.4.DTD.....	34
1.5.PHP.....	34
1.6.LA WEB.....	37
1.7.BASES DE DATOS.....	38
1.7.1.Base de datos y el WWW.....	40
1.7.1.1.¿Porqué BDs a través del WWW?	40
1.7.1.2.Acceso a la Base de Datos:.....	40
1.7.1.3.Comunicación con el servidor de BD:	41
1.8.METADATOS.....	41
1.9.LINUX	42
1.9.1.Historia.....	42
1.9.2.¿Porque Linux?.....	43
1.9.3.LyX.....	44
1.9.3.1.Diferencias entre LyX y los procesadores de texto usuales.....	46
1.9.4.LaTeX.....	48
1.9.5.HT://DIG.....	51

2.PERFIL DE LA BIBLIOTECA DE LA CORPORACIÓN UNIVERSITARIA TECNOLÓGICA DE BOLIVAR.....	52
2.1.INVESTIGACIÓN PRELIMINAR.....	56
2.2.PROCEDIMIENTO DE CLASIFICACION DE LOS DOCUMENTOS.....	57
2.3.PROCEDIMIENTO DE CONSULTA Y PRESTAMO DE UN LIBRO.....	58
2.4.RESULTADOS DE LA INVESTIGACION.....	59
3.EDICION DE DOCUMENTOS.....	61
3.1.DTD DOCBOOK.....	61
3.1.1.Sets y Books.....	62
3.1.1.1.Metadatos de Set y Book	68
3.1.1.2.Atributos de Set y Book.....	69
3.1.1.3.Componentes de Book.....	70
3.1.2.Propósito general de los elementos Sect	74
3.1.3.Tablas de contenido y listas de títulos.....	77
3.1.4.Bibliografía.....	79
3.1.5.Glosary.....	80
3.2.LA INTERFACE DE LYX.....	81
3.2.1.Operaciones básicas sobre archivos.....	81
3.2.2.WYSIWYM: El espacio en blanco en LyX.....	84
3.2.3.Entornos.....	85
3.2.4.Secciones y Subsecciones.....	86
3.2.5.Listas y sublistas.....	87
3.2.6.Estrofas, Citas y otros.....	88
3.2.7.Clases de texto y modelos.....	89

3.2.8.Título del documento.....	90
3.2.9.Etiquetas y referencias cruzadas.....	91
3.2.10.Notas a pie de página y notas al margen.....	91
3.2.11.Bibliografía.....	92
3.2.12.Índice general.....	93
3.2.13.Soporte para Ecuaciones.....	94
3.2.14. Importar documentos LaTeX: reLyX.....	94
3.2.15.BibTeX.....	95
3.2.16.Errores.....	96
4.HT://DIG.....	97
4.1.INDEXACION.....	97
4.1.1.Estructura de la Base de Datos.....	97
4.1.2.Sistema Utilizado.....	97
4.1.3.Resumen de funcionamiento.....	97
4.1.3.1.Digging	98
4.1.3.2.Merging	98
4.1.3.3.Búsqueda.....	98
4.1.4.Programas Requeridos.....	99
4.1.4.1.libg++.....	99
4.1.4.2.GNU 'make'	100
4.1.4.3.Servidor Web Apache.....	100
5.BIBLIOTECA VIRTUAL.....	101
5.1.TRANSFORMACIÓN DE LOS DOCUMENTOS AL FORMATO REQUERIDO.....	102

5.2. CLASIFICACIÓN Y ORGANIZACIÓN DE LOS DOCUMENTOS.....	104
5.3.INDEXACION DE LOS DOCUMENTOS.....	105
5.4.BUSQUEDA.....	107
6.CONCLUSIONES.....	110
RECOMENDACIONES.....	112
BIBLIOGRAFIA.....	114
ANEXOS.....	116

GLOSARIO

ABSTRACTS: Resumen.

APPENDICES o APPENDIX: Apéndice.

ARTHEADER: Cabecera.

ARTICLES: Artículos.

AUTORGROUPS: Grupos de autores.

BIBLIODIV: División de la bibliografía.

BIBLIOENTRY: Entrada de la Bibliografía.

BIBLIOGRAPHY: Bibliografía.

BIBLIOMISC: Miscelánea.

BIBTEX: Permite construir bases de datos de referencias bibliográficas que pueden usarse en múltiples documentos.

BOOK: Libro. Es el nivel en el cual todos los documentos DocBook comienzan

BOOKBIBLIO: Contiene información acerca de Book de Article.

BOOKINFO: Permite ingresar información del documento en Book.

BOOKTITLE: Título de Book.

BRIDGEHEAD: Simula la estructura de una sección.

CHAPTER: Capítulo.

DTD: Declaración del tipo de documento (Document Type Declaration) proporciona la gramática para una clase de documentos SGML

DOCBOOK: Es una DTD que hace posible que podamos tener documentos cada uno con diferentes jerarquías y una gran cantidad de elementos para el contenido de los mismos.

DOCINFO: Permite ingresar información del documento

GLOSSARY: Glosario.

GRAPHICS: Gráficas.

HT://DIG: Herramienta diseñada para indexar sitios web o para indexar directorios y realizar búsqueda dentro de los mismos

INDEX: Indice.

LABEL: Etiqueta.

LATEX: Paquete de macros construido sobre TeX

LEGALNOTICE: Licencia.

LINK: Ligar, tener conexión.

LOT: Lista de títulos de tablas o figuras.

LOTENTRY: Entrada de la lista de títulos.

LYX: Procesador de texto visual, WYSIWYM, que usa toda la potencia de LaTeX para realizar el trabajo de composición.

METADATOS: Registro de información acerca de documentos que no hace parte de ellos mismos. Datos acerca de los datos.

MODESPECS: Modo de especificación.

PAGENUM: Numero de Pagina.

PART: Parte.

PARTINTRO: Introducción de Part

PREFACE: Prologo, Prefacio.

REFENTRY: Entrada del Reference.

REFERENCE: Referencia. Agrupa colecciones.

RENDERAS: Indica la apariencia que se quiere para el título de Sect..

REVVHISTORY: Historial.

SECT: Secciones.

SERIESINFO: Permite ingresar la información en las series.

SET: Grupo. Elemento más alto en la jerarquía de Docbook

SETINDEX: Grupo de índices.

SETINFO: Permiten ingresar información del documento es Set.

SIMPLESECT: Secciones atómicas.

SRCREDIT: Provee un crédito para la fuente de una imagen.

SUBTITLE: Subtítulo.

TEX: Lenguaje de composición con capacidad para uso y creación de macros.

TITLE: Título.

TITLEABBREV: Título abreviado.

TOC: Tabla de contenido.

TOCBACKS: Entradas para la parte final.

TOCENTRIES: Apariencia para la entrada de la actual tabla de contenido.

TOCFRONTS: Parte frontal de la tabla de contenido.

TOCLEVEL: Entradas para sect.

TOCPARTS: Entradas para Part.

TOCCHAPS: Entradas para chapters y appendices..

RESUMEN

El objetivo de este trabajo es diseñar e implementar una biblioteca virtual para la Corporación universitaria Tecnológica de Bolívar, basada en un motor de búsqueda que le permita a sus usuarios buscar, encontrar y consultar la información que estará organizada por categorías tales como tipos de documentos, materia, título, autor, etc. Después de realizar un análisis detallado de las deficiencias que se presentan en la biblioteca en todo lo relacionando con la consulta y préstamo de los libros o documentos que en esta podemos encontrar, decidimos utilizar el formato electrónico (Pdf, HTML, Postcrip, Word) de dichos documentos y crear una aplicación administradora de estos que agilizará la búsqueda de un documento específico y además permitirá la consulta de varios usuarios al mismo tiempo. La plataforma que utilizamos por ser una variante muy potente de Unix, gratuito, por permitir aprovechar las maquina obsoletas (386, 486, etc.), y por tener en la web todo tipo de herramientas y aplicaciones es la distribución SUSE 7.0 de LINUX. Proponemos también el uso de su procesador de texto LyX para transcripción de los documentos y el cual, nos permitirá transformar los documentos de forma inmediata a formatos como HTML, Poscript y PDF. De estos formatos el utilizado en este proyecto es el HTML, formato en el cual también fueron realizadas las interfaces de la aplicación, que realiza la búsqueda con PHP utilizando el indexador de Linux HT://DIg.

INTRODUCCION

En estos tiempos en donde la tecnología de la informática a evolucionado hasta el punto de automatizar muchos procesos a través de la WEB, es importante tener claro el concepto de BIBLIOTECA VIRTUAL.

Muchos son los conceptos emitidos y muchos los proyectos realizados entre los cuales encontramos bibliotecas virtuales tales como la Cervantes.com que provee a los usuarios de la Web una gran cantidad de importantes obras literarias con la facilidad de leerlas. Otros proyectos son las bibliotecas de universidades, muchas de las cuales solo pueden ser utilizadas por sus estudiantes, a través de la validación de un usuario y un password. Muchos de estos conceptos no incluyen la posibilidad de mostrar la información completa de un documento, si no, que por el contrario, se limitan a mostrar una reseña bibliográfica. Es precisamente uno de los objetivos de este proyecto, facilitar a los estudiantes de la Corporación Universitaria Tecnológica de Bolívar una herramienta que le permita visualizar el texto completo de cualquier documento que incluye información a cerca de un tema determinado.

Nuestro esquema supone que una Biblioteca Virtual, está constituida por un conjunto de aplicaciones que facilitan la búsqueda y administración de libros o documentos en formato

electrónico, que serán de uso exclusivo de los estudiantes, los cuales necesitarán hacer uso de su nombre de usuario y password activos.

Teniendo en cuenta que este trabajo hace parte de un proyecto de la institución, que tiene entre sus objetivos fomentar la investigación y el desarrollo de aplicaciones en Linux, investigamos a fondo las herramientas que esta plataforma provee, para adaptarlas a nuestros fines. De este modo, este trabajo le presenta al lector un editor de texto llamado LyX posiblemente nuevo para él, y un manual que le enseñará como usarlo para escribir sus propios documentos, como Trabajos de Grado, Artículos, entre otros, y transformarlos al formato requerido por la Biblioteca Virtual, que es el HTML. También le presenta los resultados de la investigación realizada sobre el HT://Dig, Indexador de Linux, para realizar la búsqueda, al igual que el funcionamiento de las interfaces que para estos fines fueron construidas.

Este trabajo le servirá de guía a todo aquel que desee saber como funciona la Biblioteca Virtual, como realizar consultas en esta, y como contribuir con sus documentos.

1.MARCO TEORICO

1.1.BIBLIOTECA VIRTUAL

El concepto de Biblioteca Virtual no es algo nuevo. Vannevar Bush, ingeniero electricista norteamericano que en 1939 fue nombrado director del NACA y mas tarde en 1941 director de Investigación Científica, desempeñó un importante papel en las investigaciones que dieron como resultado el movimiento de ideas que hicieron nacer la cibernética. Entre los avances tecnológicos que predijo se encuentra el proyecto Memex, el cual hacía posible el intercambio de información entre los científicos y que estos tuvieran acceso total a toda la información registrada.

El término Biblioteca Virtual es usado para referirse a un sistema de aplicaciones cuya función es extender la funcionalidad de una biblioteca convencional, por ejemplo, colocando al alcance de sus usuarios toda su colección en forma electrónica. Una definición más precisa, es cualquier aplicación o sistema de aplicaciones dirigido a proveer acceso y servicio a volúmenes grandes de información, la cual debe estar en forma organizada y digitalizada.

Una biblioteca digital ofrece una gran ventaja a quienes deben manejar gran cantidad de volúmenes, ya que permite que el lector acceda mediante una pantalla, a lo que busca o

busca, sin necesidad de deteriorar el material original, pudiendo, si lo desea imprimir el texto que necesita. La biblioteca en línea, ubicada en Internet, la gran telaraña, tiene la ventaja adicional, con respecto a la biblioteca tradicional (ya sea con texto en papel o digitalizado) de ofrecer la información, almacenada en muchos lugares intangibles, a la mayor cantidad posible de lectores. Y además, es accesible con un equipamiento mínimo: una computadora, un módem, una línea telefónica, luz eléctrica y un acceso a correo electrónico o (completo) a Internet. Internet abre todo su potencial para que la interconexión entre diferentes puntos, sirva para alcanzar ese material que de otra manera sería casi inaccesible.

Hay muchas formas en las que se puede organizar una Biblioteca Digital, pero entre todas las diferencias y particularidades hay unas características generales que tienen todas en común. Las Bibliotecas Digitales usualmente poseen grandes volúmenes de información de alta calidad. El material es catalogado y se asegura de que su origen, fecha y otras descripciones externas sean muy precisas. Muchas de las Bibliotecas Virtuales se preocupan por proveer acceso digital a material existente dentro de colecciones de bibliotecas tradicionales, y se concentra en material que fue originalmente hecho para medios análogos: bibliotecas de imágenes o texto escaneado, segmentos de vídeos y audiovisuales digitalizados. Otros proyectos extienden la biblioteca a un conjunto de datos científicos, bibliotecas de software o trabajo en multimedia. La gran parte de trabajos en esta área se concentran en proveer contenidos o métodos de acceso mejorados, de tal manera que se satisfaga las necesidades de información de un individuo

en particular, por ejemplo, un estudiante de química buscando información para un proyecto, un estudiante de secundaria bajando multimedia sobre química.

1.2.EL LIBRO ELECTRÓNICO

El "Libro Electrónico", así llamado, es un dispositivo similar a un visor o pantalla o monitor de computadora, pero más liviano, por-table y con funciones, lógicamente, acotadas a la lectura y al trabajo con textos.

Hay que tener en cuenta algunas ventajas y algunas dificultades que no lo hacen demasiado favorable para este mercado, por lo menos en el muy corto plazo.

1. Recién están asomando a la plaza editorial, los primeros cuatro intentos de cuatro marcas diferentes.
2. El precio del soporte resulta aún demasiado alto.
3. El material de lectura que se ofrece no está en español.
4. Las editoriales cuentan con la ventaja de no tener que pensar en el tema de almacenamiento y cantidad de libros para reposición.
5. Al eliminar el uso de papel, tintas y otros componentes, se evitan algunos de los problemas ecológicos que afectan al planeta.
6. El editor minimiza el riesgo, si la publicación no se vende, permitiendo de esa manera que muchos autores noveles cuenten con una oportunidad que ahora no existe en cuanto a la impresión en papel.

7. Estos libros electrónicos se manejan con baterías recargables, lo que permite movilizar el material de lectura.
8. En general, como siempre ocurre con las novedades, el precio de los libros es demasiado alto, comparándose en algunos casos con el valor del libro en papel, cuando no debiera ser así. En este punto, tener en cuenta que nos estamos refiriendo a un tipo de libro preparados para una marca determinada de visor de lectura.
9. Los Libros Electrónicos tienen una alta capacidad de almacenamiento, pero no ilimitada, por ello, algunas empresas ofrecen la posibilidad de guardar los libros que ya no se leen, en línea.
10. La desactualización de los equipos. En este momento de acelerada evolución tecnológica, hay una rotación muy alta de equipamiento, y los precios de los libros electrónicos y la calidad de diseño que ofrecen, hacen dudar ante una compra acelerada. No hay problemas en cuanto al cambio de software, que sería recargable como cualquier e-book adicional, pero el problema mayor se presenta con la "arquitectura" de ese libro (mejor pantalla, definición de colores, menor peso, etc.).

1.3.HTML Y XML

El lenguaje HTML, a pesar de su sencillez, es sin duda un invento prodigioso. Es el más exitoso sistema de presentación de documentos de la historia. Desde que apareció el WWW, gracias al HTML hemos podido publicar y acceder a más información de la que jamás hemos podido imaginar.

El HTML ha sido víctima de su propio éxito. El gran crecimiento de Internet, los intereses comerciales y la necesidad de poder realizar páginas Web vistosas, ha dado lugar a que en poco tiempo este lenguaje haya evolucionado muy rápidamente y, por desgracia, no siempre por el camino más adecuado. Actualmente sigue siendo igual de rígido e inflexible como era en un principio. Y es que es un lenguaje limitado en cuanto que no nos permite realizar sobre Internet todas las aplicaciones o cosas que nos gustaría.

Estas razones han obligado a los miembros del W3 Consortium a, en lugar de desarrollar nuevas versiones de HTML, desarrollar un nuevo lenguaje (llamado metalenguaje) que han denominado XML (Extensible Markup Language) que aproveche las innegables ventajas del HTML pero que a su vez permita realizar muchas cosas más. Esto no significa, al menos por el momento, el fin del HTML.

La idea que subyace bajo el XML es la de crear un lenguaje muy general que sirva para muchas cosas. El HTML está diseñado para presentar información directamente a los humanos, y esto sin duda es algo bueno, pero es un lenguaje complicado de procesar para los programas informáticos. El HTML no es bueno porque no indica lo que está representando, se preocupa principalmente de que eso tiene que ir en azul, o con un tipo de letra determinada, pero no te dice que lo que está mostrando es el título de un libro o el precio de un artículo. El XML hace precisamente esto: describe el contenido de lo que etiqueta.

Es evidente que no hay que ser ningún programador experto para entender que cualquier programa informático podrá trabajar de forma más eficiente sobre el segundo ejemplo que sobre el primero.

Esto permitirá, por ejemplo, realizar motores de búsqueda mucho más eficaces, lo que nos permitirá un acceso más rápido y eficiente a la información. Nos permitirá acceder a nuestras páginas favoritas desde nuestro teléfono móvil, o desde la radio de nuestro coche, en el momento en el que los programas de reconocimiento de voz trabajen con XML. Facilitará el intercambio de información y la cooperación entre las empresas facilitando el comercio electrónico, etc. Y es que el XML busca precisamente crear la capacidad de hacerlo todo en la Web.

La potencia de esta forma de trabajar radica en que estamos etiquetando e identificando el contenido, olvidándonos en un principio por la forma de presentarlo. Posteriormente se le dará un formato de salida. El W3c está trabajando actualmente en el desarrollo de un lenguaje de hojas de estilo que nos lo permitan, denominado XSL (Extensible Style Language). Mediante una XSL podremos transformar un documento XML en otro XML (por ejemplo en HTML) o convertirlo a un formato de impresión: RTF, PDF, etc.

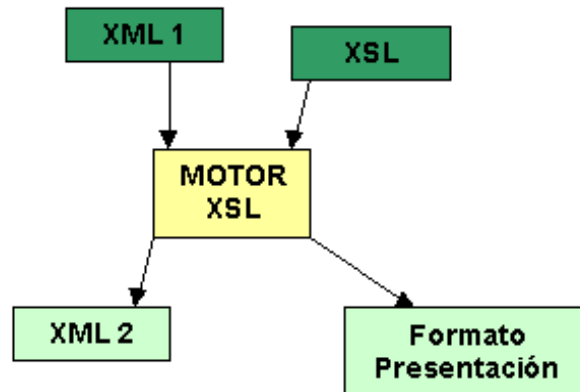


Figura 1. Funcionamiento XSL

Si el HTML supuso una revolución porque permite la comunicación entre las personas, el XML supondrá una revolución porque va a permitir la comunicación entre las máquinas.

1.3.1.HTML, XML, versus SGML

No se debe pensar que el XML es un HTML++. . Tanto el XML como el HTML tienen su base en el SGML. El SGML (Standard Generalized Markup Language, ISO 8879) es el estándar internacional para la definición de la estructura y el contenido de diferentes tipos de documentos electrónicos. Es decir, es un metalenguaje que nos permite definir lenguajes para definir la estructura y el contenido de nuestros documentos. La definición de la estructura y el contenido de un tipo de documento se realiza en una DTD. En ella definimos los elementos que conformarán ese tipo de documentos y como tienen que estar organizados para que sea correcto.

Un ejemplo de DTD es por ejemplo la que define cómo tendrán que ser los documentos HTML. Por tanto, el HTML no es más que un tipo de documento SGML que se utiliza en la Web, y esto es importante, ya que aquí radica su principal diferencia con el XML.

El XML no es ningún tipo de documento SGML, sino que es una versión abreviada de SGML optimizada para su utilización en Internet. Esto significa que con él vamos a poder definir nuestros propios tipos de documentos (podremos definir nuestras propias etiquetas) y, por tanto, ya no dependeremos de un único e inflexible tipo de documento HTML.

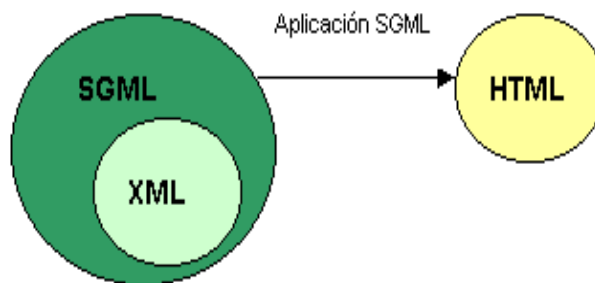


Figura 2. Diferencia entre el SGML, XML y HTML

El XML más que un HTML++ hay que considerarlo como un SGML-- optimizado para su utilización en Internet. Como escribió Richard Ligth en su libro Presenting XML, "XML ofrece el 80% de las ventajas del SGML con un 20% de su complejidad". Y es que los diseñadores de XML intentaron dejar fuera sólo aquellas partes que raramente se utilizan. Esta reducción resultó ser muy importante: la especificación XML ocupa aproximadamente 30 páginas, frente a las 500 del SGML.

1.4.DTD

El DTD o declaración del tipo de documento (Document Type Declaration) proporciona la gramática para una clase de documentos SGML. Esta gramática contiene la definición del conjunto de etiquetas que puede contener esa clase de documentos SGML. Un DTD es generalmente un fichero (o varios usados conjuntamente) que contiene la definición formal de un tipo de documento particular. Este define los nombres que pueden utilizarse en los elementos, dónde pueden aparecer y cómo se interrelaciona entre ellos.

1.5.PHP

PHP (acrónimo de "PHP: Hypertext Preprocessor") es un lenguaje interpretado de alto nivel embebido en páginas HTML y ejecutado en el servidor.

Una respuesta corta y concisa, pero que significa realmente? Un ejemplo nos aclarará las cosas: Podemos ver que no es lo mismo que un script CGI escrito en otro lenguaje de programación como Perl o C++. En vez de escribir un programa con muchos comandos para crear una salida en HTML, escribimos el código HTML con cierto código PHP embebido (introducido) en el mismo, que producirá cierta salida (en nuestro ejemplo, producir un texto). El código PHP se incluye entre etiquetas especiales de comienzo y final que nos permitirán entrar y salir del modo PHP.

Lo que distingue a PHP de la tecnología Javascript, la cual se ejecuta en la máquina cliente, es que el código PHP es ejecutado en el servidor. Si tuviésemos un script similar al de nuestro ejemplo en nuestro servidor, el cliente solamente recibiría el resultado de su ejecución en el servidor, sin ninguna posibilidad de determinar que código ha producido el resultado recibido. El servidor Web puede ser incluso configurado para que procese todos los ficheros HTML con PHP.

Al nivel más básico, PHP puede hacer cualquier cosa que se pueda hacer con un script CGI, como procesar la información de formularios, generar páginas con contenidos dinámicos, o mandar y recibir cookies.

Quizás la característica más potente y destacable de PHP es su soporte para una gran cantidad de bases de datos. Escribir un interfaz vía Web para una base de datos es una tarea simple con PHP. Las siguientes bases de datos están soportadas actualmente:

Tabla 1. Bases de datos

<i>Adabas D</i>	<i>Ingres</i>	<i>Oracle (OCI7 and OCI8)</i>
<i>dBase</i>	<i>InterBase</i>	<i>PostgreSQL</i>
<i>Empress</i>	<i>FrontBase</i>	<i>Solid</i>
<i>FilePro</i>	<i>mSQL</i>	<i>Sybase</i>
<i>IBM DB2</i>	<i>MySQL</i>	<i>Velocis</i>
<i>Informix</i>	<i>ODBC</i>	<i>Unix dbm</i>

PHP también soporta el uso de otros servicios que usen protocolos como IMAP, SNMP, NNTP, POP3, HTTP y derivados. También se pueden abrir sockets de red directos (raw sockets) e interactuar con otros protocolos.

PHP fue concebido en otoño de 1994 por Rasmus Lerdorf. Las primeras versiones no distribuidas al público fueron usadas en un sus páginas Web para mantener un control sobre quien consultaba su currículum. La primera versión disponible para el público a principios de 1995 fue conocida como "Herramientas para paginas Web personales" (Personal Home Page Tools). Consistían en un analizador sintáctico muy simple que solo entendía unas cuantas macros y una serie de utilidades comunes en las páginas Web de entonces, un libro de visitas, un contador y otras pequeñas cosas. El analizador sintáctico fue reescrito a mediados de 1995 y fue nombrado PHP/FI versión 2. FI viene de otro programa que Rasmus había escrito y que procesaba los datos de formularios. Así que combinó las "Herramientas para paginas Web personales", el "intérprete de formularios", añadió soporte para mSQL y PHP/FI vio la luz. PHP/FI creció a gran velocidad y la gente empezó a contribuir en el código.

Es difícil dar estadísticas exactas, pero se estima que a finales de 1996 PHP/FI se estaba usando al menos en 15.000 páginas Web alrededor del mundo. A mediados de 1997 este número había crecido a mas de 50.000. A mediados de 1997 el desarrollo del proyecto sufrió un profundo cambio, dejó de ser un proyecto personal de Rasmus, al cual habían ayudado un grupo de usuarios y se convirtió en un proyecto de grupo mucho más organizado. El analizador sintáctico se rescribió desde el principio por Zeev Suraski y

Andi Gutmans y este nuevo analizador estableció las bases para PHP versión 3. Gran cantidad de código de PHP/FI fue portado a PHP3 y otra gran cantidad fue escrito completamente de nuevo.

Hoy en día (finales 1999), tanto PHP/FI como PHP3 se distribuyen en un gran número de productos comerciales tales como el servidor Web "C2's StrongHold" y Redhat Linux. Una estimación conservativa basada en estadísticas de NetCraft (ver también Estudio de NetCraft sobre servidores Web), es que más de 1.000.000 de servidores alrededor del mundo usan PHP. Para hacernos una idea, este número es mayor que el número de servidores que utilizan el "Netscape's Enterprise server" en Internet.

A la vez que todo esto está pasando, el trabajo de desarrollo de la próxima generación de PHP está en marcha. Esta versión utiliza el potente motor de Scripts Zend para proporcionar altas prestaciones, así como soporta otros servidores Web, además de apache, que corren PHP como módulo nativo.

1.6.LA WEB

Por la WEB entendemos información en Internet, como es accesada por individuos usando la Red Mundial de Redes o World Wide Web, o alguna otra herramienta para el acceso a la información en redes de trabajo. La Web es accesada usando uno de los muchos navegadores ahora disponibles. Esta provee una interface con la información. Eso es, un usuario es presentado con un documento el cual incluye links a seguir y la

forma de salir. Por interactuar con el documento el usuario genera un nuevo documento a ser presentado. La Web, como un servicio de Internet, es ante todo pública. Un sitio Web puede proveer acceso a un gran número de usuarios a través del mundo.

1.7.BASES DE DATOS

En un principio no había computadores. Cuando los hubo, la primera forma de almacenamiento era plana, en grandes cintas magnéticas o perforadas. Después los informáticos empezaron a organizar sus datos en dos tipos de modelos: el modelo jerárquico y el modelo de redes. En el modelo jerárquico, los tipos de información se clasificaban en forma de árbol. En determinada base de datos, por ejemplo, la raíz de este árbol eran los registros de empresas. Cada empresa almacenaría los datos de un conjunto de departamentos, estos últimos serían responsables de guardar los datos de sus empleados, y así sucesivamente. Además del conjunto de departamentos, de una empresa podría ser propietaria de otro conjunto de registros, como bienes inmuebles o algo así.

El problema, como es fácil de imaginar, es que el mundo real no se adapta fácilmente a este tipo de organización. Por lo tanto, a este modelo de datos se le añaden "registros virtuales" que son, en el fondo, una forma primitiva de punteros entre registros.

El modelo de redes es un poco más flexible. Un registro puede contener un conjunto de otros registros. Además a cada registro de estos puede pertenecer al conjunto de otros registros. Este se lograba por medio de punteros.

Estos modelos desaparecieron, ya que eran sistemas navegacionales, ya que para poder obtener un dato había que tener una idea clara de cómo estaban organizados los datos. Pero lo que más afectaba estos modelos era que carecían de herramientas que facilitaran la consultas a la bases de datos. Fue donde apareció un matemático de la IBM, el cual no inventó el concepto de registro ya que esta ya existía, sino que se dio cuenta que si organizaban cada registro como un campo simple, podía diseñarse un sistema matemático que permitiera la descomposición de la información acerca de estos registros planos, con la seguridad de poder restaurar estos datos más adelante, con la ayuda de operaciones algebraica. A la teoría matemática que hay detrás de este modelo se le conoce como álgebra relacional, y es la base de nuestro conocido lenguaje SQL y del QBE. De esta forma cualquier persona puede interactuar con una base de datos.

En el mercado existen mucho manejadores de bases de datos. La pregunta es cuales son los criterios para evaluar un manejador de base de datos?

- Plataformas Soportadas.
- Soporte de tipos de datos y restricciones.
- Lenguaje de triggers y procedimientos almacenados.
- Implementación de transacciones: recuperación y aislamiento.
- Segmentación.
- Replicación.
- El Precio

1.7.1.Base de datos y el WWW

1.7.1.1.¿Porqué BDs a través del WWW?

- **Interface gráfica:** provista por el *browser*
- **Estandarización:** hojas en HTML
- **Soporte multiplataformas:** existen *browsers* para casi cualquier plataforma
- **Acceso a redes:** el servidor HTTP se encarga de las actividades de red, liberando al programador y al usuario

1.7.1.2.Acceso a la Base de Datos:

- La forma mas común es usando CGI como el *front-end* que accesa la base de datos (*back-end*)
- los principales fabricantes de BDs proveen diferentes métodos para accesar sus productos.
- El acceso puede ser:
 - leyendo directamente de los archivos que tienen los datos
 - comunicándose con el manejador de la BD

1.7.1.3. Comunicación con el servidor de BD:

- Se puede lograr con servidores que tienen soporte para estándares tales como ODBC o SQL
- De esta forma se hace interface a un cliente de ODBC o SQL que puede comunicarse con el servidor de BD
- Muchos sistemas populares brindan estas facilidades: Informix, SyBase, Oracle, Paradox, Interbase, Access, Approach, etc.

1.8.METADATOS

En las bibliotecas digitales es común querer el registro de información acerca de documentos que no hace parte de ellos mismos. Esto es algunas veces llamado "**Atributos del Documento**"; algunas veces "**Metadatos**" para indicar que ellos son datos acerca de los datos. Los metadatos ayudan en la descripción, organización y acceso de la información en la red.

Los metadatos consisten en información que caracteriza datos. Los metadatos son utilizados para suministrar información sobre datos producidos. En esencia, los metadatos intentan responder a las preguntas quién, que, cuando, donde, porqué y cómo, sobre cada una de las facetas relativas a los datos que se documentan. Los sistemas en línea para manejar metadatos necesitan manipular objetos predecibles en forma y contenido.

Las bibliotecas han sido tradicionalmente muy buenas con respecto a indexar, proceso mediante el cual se asocian metadatos con material bibliográfico. El catálogo de acceso para un dato en la biblioteca suministra metadatos acerca de ese dato. Hay una variedad de estándares usados para indexar en línea.

1.9.LINUX

1.9.1.Historia

El autor primario de Linux es Linus Torvalds. A partir de la versión original, ha sido mejorado por incontables personas de todo el mundo. Se trata de un clon del sistema operativo Unix, escrito desde cero. Ni USL, ni la Universidad de California, Berkeley, tuvieron participación en la escritura de Linux. Una de las cosas más interesantes de Linux es que el desarrollo ocurre simultáneamente alrededor del mundo. La gente ha contribuido desde Australia a Finlandia y esperamos que lo siga haciendo.

Linux comenzó como un proyecto para explorar el chip 386. Uno de los primeros proyectos de Linus fue un programa que pueda alternar entre la impresión de AAAA y BBBB. Esto luego evolucionó hacia Linux.

Linux ha sido registrado bajo los términos de la Licencia Pública General GNU (GNU General Public License) o GPL. Esta licencia, escrita por la Free Software Foundation (FSF), está diseñada para evitar que alguna persona restrinja la distribución de software.

En pocas palabras, dice que aunque Ud. cobre a alguien por entregarle una copia, no podrá impedir que ese alguien la regale. También significa que debe estar disponible el código fuente. Esto es útil para los programadores.

Cualquiera puede modificar Linux y aún distribuir sus modificaciones, siempre que mantenga el código bajo la misma licencia.

En Linux puede correr la mayoría del software popular en Unix, incluyendo el Sistema de Ventanas X. X se desarrolló en el Instituto Tecnológico Massachusetts, para permitir a los sistemas Unix la creación de ventanas gráficas, y la cómoda interacción con las mismas. En la actualidad, el sistema X se usa en todas las versiones disponibles de Unix.

Además de las dos variaciones de Unix, System V y BSD, existe un conjunto de documentos de estandarización publicados por la IEEE denominados POSIX. Linux antes que nada satisface los documentos POSIX-1 y POSIX-2. Su apariencia se asemeja mucho a la de BSD en ciertas partes, mientras que es parecido a System V en otras. Es una combinación (y para la mayoría de las personas, una buena) de los tres estándares.

1.9.2.¿Porque Linux?

Linux es un sistema operativo de gran aceptación mundial. Sus principales ventajas son:

Tabla 2. Ventajas de Linux

<i>Gratuidad.</i>	<i>Linux es Gratuito.</i>
Potencia.	Permite aprovechar al máximo las máquinas obsoletas (386, 486, etc..).
Compatibilidad con Unix.	Linux es una variante muy potente de Unix.
Aplicaciones de todo tipo disponibles con fuentes incluidas.	Circulan por la Web todo tipo de herramientas y aplicaciones.

1.9.3.LyX

Parte del reto inicial de usar LyX surge del cambio en la manera de pensar que debe hacer el usuario. En su momento, todo lo que teníamos para crear documentos eran máquinas de escribir, así que aprendimos verdaderas artimañas para evitar sus limitaciones. Subrayar, que es poco más que sobrescribir con el carácter “_”, se convirtió en un forma de resaltar texto. Para crear una tabla, se establecía a mano el ancho de cada columna y se ponían las tabulaciones necesarias. Lo mismo se aplicaba para cartas y otros textos sangrados a la derecha. Además, la ruptura de palabras al final de línea requería ser muy cuidadoso y previsor.

En otras palabras, todos hemos sido entrenados para preocuparnos por los pequeños detalles de “qué caracter va en qué lugar”.

Como consecuencia, casi todos los procesadores de texto se basan en esta mentalidad. Todavía usan tabuladores para añadir espacios en blanco. Todavía te tienes que preocupar de en qué parte exacta de la página saldrá cada cosa. Resaltar texto significa cambiar el tipo de letra, similar a cambiar la rueda de una máquina de escribir.

Aquí es donde LyX difiere de un procesador de texto corriente. El usuario no tiene que preocuparse de que una letra vaya en un sitio determinado. Le dice a LyX lo que está haciendo y él se preocupa de todo lo demás, siguiendo un conjunto de reglas llamado estilo. Veamos un pequeño ejemplo.

Supongamos que estamos realizando un informe. Queremos que comience con una sección llamada “Introducción”. Así pues, nos dirigimos a cualquiera que sea el menú de el procesador de texto que cambia el tamaño de fuente y elegimos un nuevo tamaño. Después cambiamos también a negrita. Seguidamente escribimos: “1._Introducción”. Por supuesto, si más tarde decidimos que esta sección pertenece a alguna otra parte del documento, o bien insertamos una nueva sección anterior a ésta, tenemos que cambiarle la numeración a ella y a todas las posteriores, además de las correspondientes entradas en el índice.

En LyX, nos dirigimos a la lista situada a la derecha de todos los botones y eliges Sección, y escribes “Introducción”.

Si cortamos y pegamos la sección en otra parte, todo es renumerado automáticamente. Se puede hacer incluso que LyX actualice cualquier referencia a la sección que esté dentro del fichero.

Con el procesador de texto tradicional hay problemas de consistencia. Cinco días más tarde, abrimos el informe y comenzamos la sección 4. Sin embargo, hemos olvidado que estábamos usando la letra en negrita de 18 puntos, y usamos la de 16, así que acabamos escribiendo el encabezado de la sección 4 con un tipo de letra distinto al que usamos para la sección 1. Este problema ni siquiera existe en LyX. El ordenador se encarga de todo el tedioso trabajo de llevar la cuenta de tamaños y fuentes. Al fin y al cabo, para eso está hecho.

Así pues, la idea esencial detrás de LyX es especificar lo que se está haciendo, no cómo hacerlo. En lugar de un procesador “lo que ves es lo que obtienes” (WYSIWIG, What You See Is What You Get), el modelo de LyX es “lo que ves es lo que quieres decir” (WYSIWIM, What You See Is What You Mean).

1.9.3.1. Diferencias entre LyX y los procesadores de texto usuales

Es importante describir las características de LyX, y una de las principales, WYSIWIM, es un concepto radicalmente diferente del que el 99% de la gente tiene del proceso de texto. Sus principales ventajas son:

Tabla 3. Ventajas de LyX

<i>Cosas que no encontraras en LyX</i>	<i>Cosas que presenta LyX</i>
Regla	Controles de sangrado
Tabuladores	Saltos de pagina
Espacios en blancos adicionales(usando enter)	Espacios entre lineas
	Espacio en blanco horizontal y vertical
	Tipos de letra y tamaño
	Estilo de letras

Los tabuladores, así como la regla (que nos muestra la posición de cada elemento en la página), son inútiles en LyX. El programa se preocupa de dónde tiene que ir cada cosa, no el usuario. Con los espacios en blanco adicionales ocurre lo mismo; LyX los añade conforme son necesarios, según el contexto. Al principio puede resultar molesto no poder escribir dos líneas en blanco seguidas, pero cobra mucho más sentido una vez que empezamos a pensar en términos WYSIWYM.

Aunque aparecen en LyX, no se necesitan normalmente. El programa se preocupa de estas cosas por el usuario, actuando en consecuencia según lo que se esté haciendo. Diferentes partes del documento son automáticamente puestas en diferente tamaño y estilo. El sangrado de cada párrafo es dependiente del contexto; cada tipo de párrafo se sangra de manera diferente. Los saltos de página se manejan también de forma automática.

En general, el espacio entre líneas, entre palabras y entre párrafos es variable, elegido por LyX.

Por último, éstas son las áreas en las que LyX (y LaTeX) sobrepasa a muchos procesadores de texto:

- Separación de palabras a final de línea
- Listas de cualquier tipo
- Matemáticas
- Tablas
- Referencias cruzadas

Por supuesto, muchos procesadores de texto modernos manejan símbolos matemáticos, tablas, separación de palabras a final de línea, e incluso comienzan a aproximarse a las definiciones de estilo y el concepto WYSIWYM. Sin embargo, acaban de empezar a incluir estas características, mientras que LyX está construido sobre el sistema de proceso de documentos LaTeX.

1.9.4. LaTeX

LaTeX es un sistema de preparación de documentos diseñado por Leslie Lamport en 1985. Fue construido gradualmente sobre un lenguaje de composición de documentos llamado TeX, creado por Donald Knuth en 1984. “TeX” se pronuncia como “blech” en

inglés . Sin embargo, muchos no comprenden qué es exactamente. TeX toma una secuencia de órdenes de composición escritos en un fichero ASCII, y los ejecuta. Es más complicado que una máquina de escribir, pero sin llegar a la especialización y complejidad de una auténtica imprenta. En cualquier caso, produce como salida un fichero de formato llamado “independiente de dispositivo”, o dvi para abreviar. El fichero dvi puede ser leído después por otro programa que acepte este formato, o convertido a otros formatos, como PostScript.

Si no tuviera más característica que ésta, sería simplemente un motor de composición. Sin embargo, TeX también permite definir macros. La mayoría de la gente que usa TeX está usando realmente un conjunto de macros para ocultar muchos de los detalles de composición. Esto es en lo que piensa la gente cuando habla de TeX. Los usuarios normales no trabajan con TeX puro, un esqueleto desnudo formado únicamente por comandos de composición. Sólo aquellos que crean conjuntos de macros lo hacen. En realidad se querían macros que fueran más orientadas al usuario y menos a la composición, un conjunto de comandos que sirvieran para componer secciones, tablas o fórmulas matemáticas de una forma consistente y uniforme, sin demasiadas complicaciones.

Ahora, de manera simultánea al desarrollo y crecimiento de LaTeX, otras personas están creando sus propios paquetes personalizados de macros para TeX, algunos para realizar trabajos en publicaciones matemáticas y cosas así. Unos usaron TeX directamente, otros comenzaron a modificar LaTeX. Para tratar de unificar este lío, un equipo de expertos en

LaTeX, empezaron a trabajar en LaTeX2e. Esta versión posee comandos que proporcionan una interfaz más fácil para la creación de macros, ayuda para usar las nuevas fuentes, y más mejoras. De hecho, LaTeX es en sí mismo un vasto lenguaje por derecho propio. Usuarios de todo el mundo han estado creando sus propios añadidos para LaTeX, además de los estándar.

Existen dos formas de extender LaTeX: las clases y los estilos. Una clase es un conjunto de macros de LaTeX (y de TeX) que describen un nuevo tipo de documento, como un libro o un artículo. Hay clases para transparencias, para publicaciones de física y matemáticas\ldots{}. ¡Algunas universidades tienen incluso una clase para su propio formato de tesis! Un estilo, a diferencia de una clase, no define un nuevo tipo de documento, sino un nuevo tipo de comportamiento, que puede ser utilizado por cualquier documento. Por ejemplo, LyX controla los márgenes de página y el espaciado entre líneas usando dos ficheros de estilo diferentes de LaTeX, diseñados para este fin. Hay ficheros de estilo para gran cantidad de cosas: imprimir etiquetas o sobres, cambiar el sangrado normal del texto, añadir nuevos tipos de letra, manipular gráficos, diseñar elaborados encabezados de página, personalizar bibliografías, alterar la posición y apariencia de notas a pie de página, tablas y figuras, personalizar listas, etc.

En resumen:

Tabla 4. Ventajas de LyX

<i>TeX</i>	Lenguaje de composición con capacidad para uso y creación de macros.
<i>Latex</i>	Paquete de macros construido sobre TeX.
<i>Clases</i>	Descripciones de un tipo de documento, usando LaTeX.
<i>Estilos</i>	Alteran algún aspecto del comportamiento normal de LaTeX.
<i>LyX</i>	Procesador de texto visual, WYSIWYM, que usa toda la potencia de LaTeX para realizar el trabajo de composición.

LyX funciona de manera diferente a otros procesadores de texto, usa LaTeX como motor de composición. Como este último, se centra en el contexto de tu escritura (lo que estás escribiendo). El ordenador se encarga entonces de cómo debe aparecer.

1.9.5.HT://DIG

Herramienta diseñada para indexar sitios web o para indexar directorios y realizar búsqueda dentro de los mismos y usar los resultados para ser mostrados en una página en HTML, PHP, etc.

2.PERFIL DE LA BIBLIOTECA DE LA CORPORACIÓN UNIVERSITARIA TECNOLÓGICA DE BOLIVAR

La Biblioteca de la CUTB, fue creada en 1971, iniciando con una colección de 500 volúmenes, para los programas que existían, Economía Industrial, Ingeniería Industrial, Ingeniería mecánica y Eléctrica. Contaba con una sala de lectura que albergaba unos 60 estudiantes, ubicada en el segundo piso de la casa del bloque A en la sede de Manga. Prestaba sus servicios de lunes a viernes, de 2:00 de la tarde a 9:00 de la noche y los sábados de 9:00 a 2:00 de la tarde.

En los años 1980, se crean otras salas de lectura, adquiriéndose un gran número de volúmenes de libros, creándose un espacio para la hemeroteca con algunos cuantos títulos de revistas en las distintas áreas académicas, con servicios de fotocopidora, ampliándose el horario de servicios de lunes a viernes de 9:00 a 12:00 del medio día y los sábados de 9:00 de la mañana a 3:00 de la tarde.

En el año 1984 se crea una base de datos bajo el formato SCIB (Servicio Colombiano de Información Bibliográfica), creada por el ICFES. Donde se empezó a introducir la información existente de aproximadamente 2000 registros, bajo el sistema operativo DOS y ambiente Informix. Pero en el año 1991 sufre un daño borrándose toda la información.

En los años 1992 se crea una base de datos bajo los mismo estándares del formato SCIB, en el sistema operativo UNIX, en ambiente Informix. Se procedió a realizar un nuevo inventario de toda la colección y una actualización del vocabulario bajo los estándares de las normas Angloamericanas para bibliotecas universitarias. En ese mismo año se empieza la construcción de otra biblioteca satélite en el campus de Ternera, para los programas de ingenierías de los niveles de 1 a 4 semestres, dotada con algunos textos y revistas, ubicada en el segundo piso del edificio de laboratorios, con una sala de lectura múltiple y servicios de fotocopidora, con un horario de 8:00 de la mañana a 12:00 del medio día y de 2:00 de la tarde a 8:00 de la noche de lunes a viernes, y los sábados, de 9:00 de a mañana a 1:00 de la tarde. Más tarde en los años 1995 se trasladan todos los programas de pregrado a Ternera, creándose otra biblioteca en el bloque de tecnologías, solo para los programas de psicología y ciencias económicas. Ampliándose la compra de libros y revistas. En el año 2001 se traslada nuevamente la biblioteca de los programas antes mencionados al campus de manga y se creó una sala de lectura en grupo en la biblioteca del campus de Ternera.

Hasta la fecha cuenta con 10.603 títulos de libros y 1.612 tesis de grado, con 11.276 volúmenes de libros y 2.322 volúmenes de tesis de grados. Se está reformando y ampliando el software de consulta bibliográfica.

El personal que trabaja en la biblioteca está integrado por 6 personas de la siguiente manera:

Un ***Jefe de Biblioteca*** que lleva el control de la biblioteca, ya que establece todo los servicios y tiene responsabilidad directa con el control del material que en esta se manipula.

Una persona encargada de los ***Procesos Técnicos***, encargada de clasificar los libros.

Dos ***Auxiliares de Circulación y Préstamo***, encargados del préstamo de los libros.

Un ***Auxiliar De Hemeroteca*** encargado de vigilar y controlar los movimientos relacionados con revistas, folletos, periódicos, etc..

Un ***Portero*** encargado de prestar el servicio de vigilancia y recibir los útiles que lleven consigo las personas que ingresen a la biblioteca.

Los usuarios de la biblioteca pueden ser estudiantes activos de la universidad, profesores, egresados, directivos o estudian provenientes de otras universidades o personas que soliciten nuestros servicios. Para usar los servicios de la biblioteca se requieren la presentación del carnet vigente.

El tipo de colecciones que se pueden encontrar en la biblioteca es:

Obras de referencia: guías, enciclopedias, diccionarios, anuarios, estadísticas y otros.

Colección de reserva: Aquellas obras que por circunstancias especiales son tales como el número de ejemplares existentes o la demanda, son exclusivamente de consulta interna con prestamos a domicilios los fines de semana.

Colección general: Compuesta por materiales que, por su naturaleza o por poseer un mayor número de ejemplares, pueden ser prestados a domicilio.

La Mapoteca: Compuesta por mapas de Colombia que nos ubican en su geografía y nos recuerdan su historia.

La Hemeroteca: Conformada por revistas, folletos, documentos, periodos y archivo vertical.

Ayudas Audiovisuales: La conforman vídeos, diapositivas, acetatos, edición y película.

Multimedia– CD Rom: Posee información sobre cultura general en este medio para la investigación de los usuarios.

La biblioteca cuenta con los siguientes servicios: de fotocopia de libros artículos de revista que se encuentren en la biblioteca, y de un servicio SIDES (sistema de información y documentación para Educación Superior), encargado de la conmutación bibliográfica por medio de fotocopias.

2.1.INVESTIGACIÓN PRELIMINAR

Los requerimientos de La biblioteca de la Universidad Tecnológica de Bolívar, establecidos por la instrucción y de los que ya se ha obtenido una cantidad de detalles, radican en la consulta y prestamos de libros por parte del cuerpo estudiantil y el profesorado, de los cuales muchos están en estado de deterioro, de otros solo existen pocos ejemplares y se ven situaciones como las siguientes:

- Gran cantidad de peticiones de libros por parte de los usuarios no son satisfechas, porque no quedan libros en estanterías para prestar, porque se están usando o están en mal estado físico.
- Deterioro físico de los libros, por uso normal, por rallado, mutilación de hojas o capítulos, uso excesivo o mal uso, fotocopiado, algunas pérdidas por robo, etc.
- Deficiencia en los equipos que proporcionan el sistema de búsqueda bibliográfica.
- Deficiencia en la búsqueda de los libros la cual es realizada por el bibliotecario con la referencia bibliográfica que el usuario toma del sistema de consulta bibliográfica.

Ademas la naturaleza de los procesos, la solicitudes e información detallada son conocidas ampliamente por nosotros los estudiantes que dirigimos la investigación, todo esto nos ha ahorrado el uso de técnicas para encontrar hechos.

2.2.PROCEDIMIENTO DE CLASIFICACION DE LOS DOCUMENTOS

Cuando se adquiere un libro, antes de ser ingresado a la biblioteca debe ser llevado al almacén donde se le coloca un numero de inventario después de ser revisado. Posteriormente es llevado a la biblioteca donde el auxiliar lo sella, lo clasifica y le agrega una ficha de control para futuras estadísticas.

Para la clasificación de los libros se usa el sistema SCDD20 (Sistema de Clasificación Decimal DEWEY), que se realiza de la siguiente manera:

Primero se ubica el libro en el índice general de áreas, seguido de un numero que se obtiene del índice específico por materia que depende del contenido del libro. Adicionalmente llevara otros códigos tales como el de autor, volumen, ejemplar y una banda de color que indica el tipo de libro.

Tabla 5. Indice general de áreas

000	GENERALIDADES
100	FILOSOFIA Y SICOLOGÍA
200	RELIGIÓN
300	CIENCIAS SOCIALES
400	Lenguas
500	TECNOLOGÍAS (CIENCIAS APLICADAS)
600	LAS ARTES
700	LITERATURA Y RETORICA
800	GEOGRAFIA E HISTORIA

El código de autor se obtiene de la tabla de CULTER. SAN. BORN., Escogiendo las tres primeras letras del apellido del autor. En caso de que sean dos autores se escogen las tres primeras letras del apellido del primer autor, y si son mas de tres autores se escogen las tres primeras letras del titulo del libro.

El código del volumen depende del numero de volúmenes que tenga un titulo.

El código de ejemplar depende de la cantidad de ejemplares que existan en la biblioteca del mismo libro

La banda de color será roja si es un libro de reserva, amarilla si es un libro general y se puede prestar y azul si es el único ejemplar y no se puede prestar.

2.3.PROCEDIMIENTO DE CONSULTA Y PRESTAMO DE UN LIBRO

El procedimiento de consulta y préstamo de un libro se lleva a cabo de la siguiente manera: el usuario al ingresar a la biblioteca hará uso de los computadores ubicados en la instalaciones de la biblioteca para hacer uso del sistema SIDES el cual presta el servicio de búsqueda de documentos en la biblioteca ya sea por Autor, Materia o Titulo de este. Este le arrojará la lista de los documentos encontrados al realizar la consulta los cuales tienen un código de clasificación que deberá ser entregado al auxiliar de circulación y préstamo, quien se encargará de buscar el libro en los estantes de la biblioteca. Si el usuario va a hacer uso del documentó en las áreas de lectura de la

biblioteca, hará entrega de su carnet al auxiliar y firmara la ficha de control para que se le haga entrega del documento. Si el usuario desea llevar el libro a su casa se verificara el tipo de usuario y el tipo de documentó, dependiendo de esto se sabrá que día puede llevarse el documento, y cuando debe regresarlo.

El auxiliar maneja adicionalmente unas tablas de control de préstamo por área según su clasificación general y por carrera, para efectos estadísticos al final del mes.

2.4.RESULTADOS DE LA INVESTIGACION

Como resultado de la investigación, se determinó óptimo la creación de un sistema que permitiera a los usuarios de la biblioteca visualizar los textos completos de los libros que se encuentran en la biblioteca en su formato electrónico, y ademas permitiera solucionar el problema de concurrencia.

En el esquema que proponemos tomamos como formato electrónico el HTML, y organizamos estos documentos de tal manera que puedan ser manipulados como las paginas de un sitio web, y de la misma manera puedan ser buscados.

Estos serían los procesos básicos del sistema a diseñar:

1. Transformación de los documentos al formato requerido.
2. Indexación de los documentos.

3. Búsqueda.

3.EDICION DE DOCUMENTOS

3.1.DTD DOCBOOK

La DTD Docbook hace posible que podamos tener documentos cada uno con diferentes jerarquías y una gran cantidad de elementos para el contenido de los mismos. Además nos ofrece la posibilidad de tener marcas para introducir a los documentos información no lineal como Hipervínculos.

DocBook soporta varios tipos de documentos a saber:

- Libros y colecciones de libros.
- Referencias
- Artículos

Los documentos no están obligados a comenzar por la raíz de la estructura de Docbook. Un documento puede empezar con cualquiera de los elementos de cualquier nivel de la estructura de Docbook, siempre y cuando los elementos del documentos se vean reflejados en el nombre del tipo del documentos en la declaración del DOCTYPE. Por ejemplo:

```
<!DOCTYPE para PUBLIC “ -//Davenport//DTD Docbook V3.0//EN”>
```

<Para>

Este es un pequeño ejemplo de documento que consiste en un simple párrafo Docbook

</Para>

3.1.1.Sets y Books

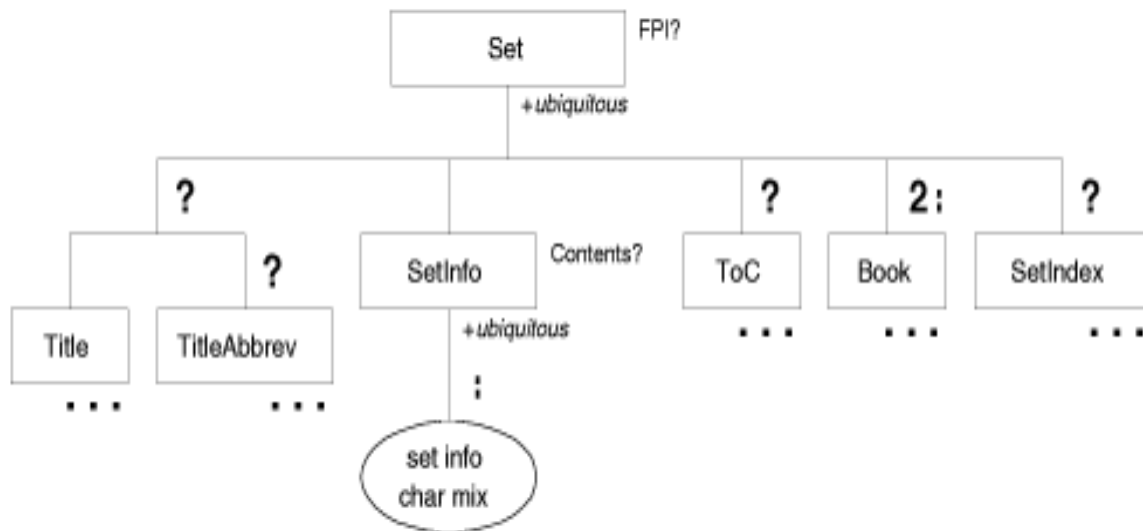


Figura 3. Diagrama de la estructura de Sets

El elemento Set es el más alto elemento en la jerarquía de Docbook. Contiene unos elementos Title y TitleAbbrev opcionales, un SetInfor opcional (que puede contener Title), un ToC opcional, dos o mas Books, y un SetIndex que es también opcional. Cada elemento Set, Book, RefEntry, y Article tienen sus propios elementos para metadatos.

Un sencillo ejemplo puede estar estructurado de la siguiente manera:

Set

Title

SetInfo

Book

Book

Book

SetIndex

El elemento Book es el nivel en el cual todos los documentos DocBook comienzan. Sin embargo, si su documento es iniciado con un Set, no es necesario usar Set a menos de que se quiera hacer algún tipo de procesamiento con la marca Set.

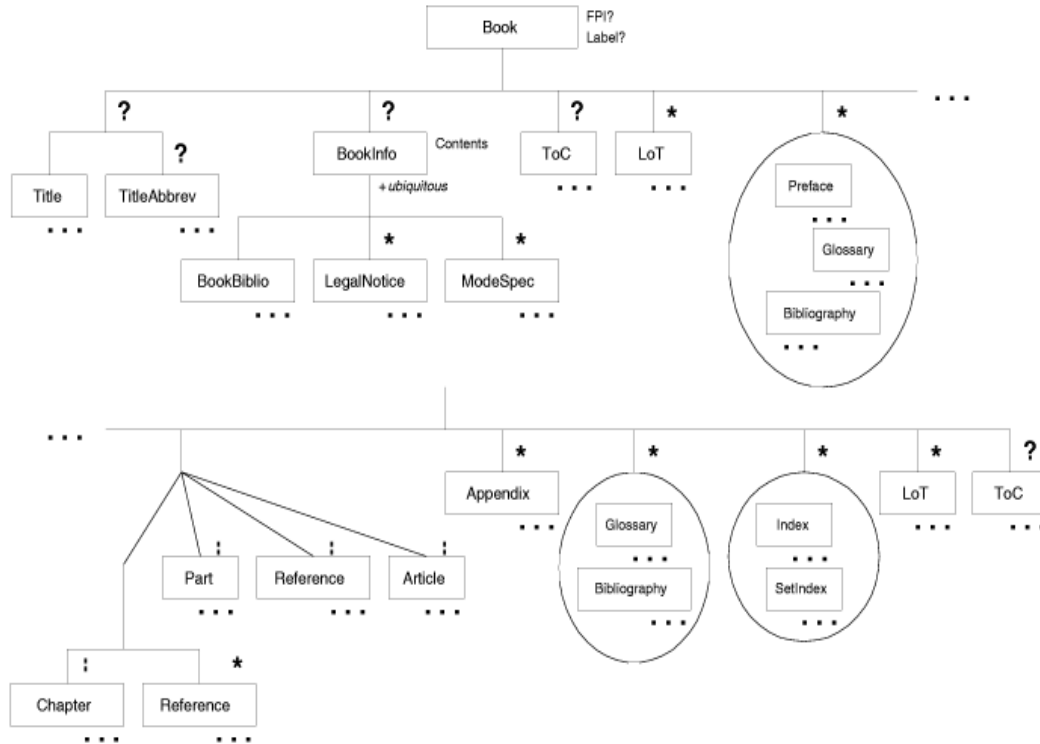


Figura 4. Diagrama de la estructura Book

La estructura de Book es relativamente compleja. Permite para una variedad de configuraciones de componentes principales Book. La estructura supone que se ubicarán tantos elementos ToC y Glossary donde se quiera que aparezcan en la salida del documento. El modelo del contenido de Book tiene tres grandes segmentos, correspondientes al tema de encabezado, cuerpo y cierre del documento.

- En el principio, Book puede contener un Title y un TitleAbbrev opcionales, un Book Info opcional (que también puede tener Title y TitleAbbrev), un ToC opcional, cero o mas LoTs (listas de títulos), y una mezcla de cero o más elementos Bibliography, Glossary, y Preface en cualquier orden. Podríamos tomar un lugar para guardar la

información del título de Book: directamente en Book o en un elemento BookBiblio de BookInfo.

- El cuerpo de Book puede tener elementos Chapter (capítulos) seguidos por cualquier número de elementos Reference, solo Reference, Solo Parts (que pueden contener elementos como capítulos), o sólo Articles.
- En el final, Book puede contener cero o más Appendices, una mezcla de cero o mas elementos Bibliography y Glossary en cualquier orden, una mezcla de cero o mas elementos Index y SetIndex en cualquier orden, cero o mas elementos LoTs, y un elemento ToC que es opcional. Ejemplo:

Book

BookInfo

ToC

LoT para figuras

LoT para tablas

Preface

Chapter

Chapter

Chapter

Chapter

Reference

Appendix

Appendix

Glossary

Bibliography

Index

Books puede tener elementos Chapter y Appendices agrupados en elementos Parts. Puede tener grupos de Appendices en un elemento Part colocando al final del elemento Book un elemento Part que contenga solo elementos Appendices. No se pueden mezclar Chapters y Appendices en una parte específica, aunque la estructura así lo permita. Ejemplo:

Book

BookInfo

Preface

Part

PartIntro

Chapter

Chapter

Part

PartIntro

Chapter

Chapter

Chapter

Part

PartIntro

Appendix

Appendix

Una referencia manual puede estructurarse así:

Book

BookInfo

ToC

LoT

Preface

Reference

Reference

Reference

Reference

Appendix

Appendix

Appendix

Appendix

Index

3.1.1.1. Metadatos de Set y Book

SetInfo y BookInfo permiten ingresar información del documento (también llamada Metadatos) en Sets y Books, respectivamente. SetInfo permite tomar la información de los elementos del documento y colocarlos en cualquier orden. BookInfo, por el contrario, impone un orden a sus subelementos. Contiene cero o más elementos Graphics que ilustran o representan los extremos de Book, un elemento BookBiblio requerido (que también impone un estricto orden a sus elementos), cero o más LegalNotices, y cero o más ModeSpecs para el almacenaje de la información relacionada con los elementos de enlace.

BookBiblio es utilizado en BookInfo y ArtHeader para ingresar información acerca de Book o Article y es también utilizado en BiblioEntry para ingresar información acerca de citas.

3.1.1.3. Componentes de Book

Appendix, Chapter, Index, Part, Preface, SetIndex y Reference son componentes permitidos dentro de Book. Bibliography y Glossary pueden estar entre los componentes de Book o pueden hacer parte de componentes de Book.

El cuerpo de un Book puede estar organizado comenzando con uno o dos niveles lógicos. Puede contener componentes de Book tales como Chapter, Appendices, y References directamente, o puede contener grupos de estos componentes organizados dentro de elementos Part.

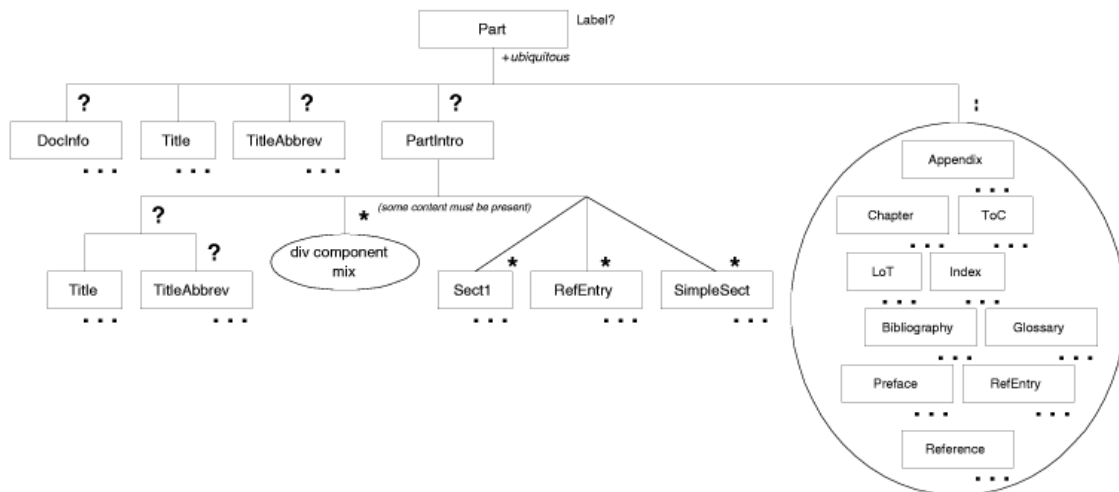


Figura 6. Diagrama de la estructura de Part

El modelo del contenido de Part, es muy amplio: contiene un elemento DocInfo opcional, un Title, un Title Abbrev opcional, un PartInfo opcional, y una mezcla de uno mas elementos Appendices, Chapter, Bibliography, Glossary, Prefaces, RefEntry y Reference en cualquier orden.



Figura 7. Diagrama de la estructura de Reference

Reference tiene una utilidad especial para agrupar colecciones de elementos RefEntry. Reference pueden ser usados, por ejemplo, para organizar páginas de manuales en secciones numeradas. Reference contiene un elemento DocInfo opcional, un Title, un TitleAbbrev opcional, un PartInfo opcional (que es también usado en Part), y uno o mas elementos RefEntry. El atributo Label en Reference puede tener un valor no numérico como palabra clave, mejor que un número, para la construcción formal del objeto Prefijo, prefijo de número de página, y otros.

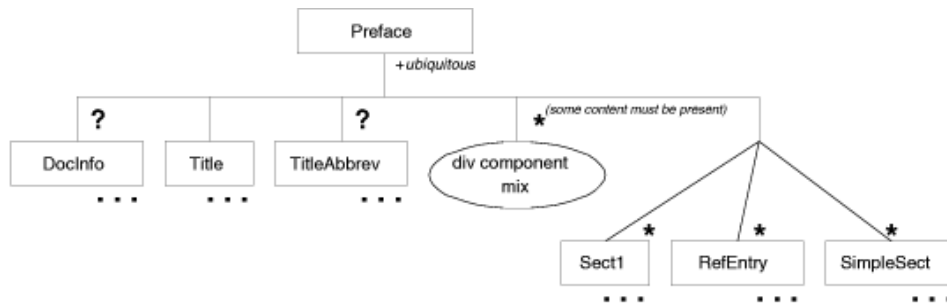


Figura 8. Diagrama de la estructura de Preface

El modelo del contenido de Preface, Chapter, y Appendix, son similares. Preface contiene un elemento DocInfo opcional, un Title, un TitleAbbrev opcional, cualquier número de objetos en el nivel de un elemento %divcomponent.mix, y opcionalmente un grupo de Sect o SimpleSect, o un grupo de RefEntry. Un Book puede tener múltiples Prefaces, por ejemplo, una sección de agradecimientos y otra un prefacio; estos se distinguen por sus títulos.

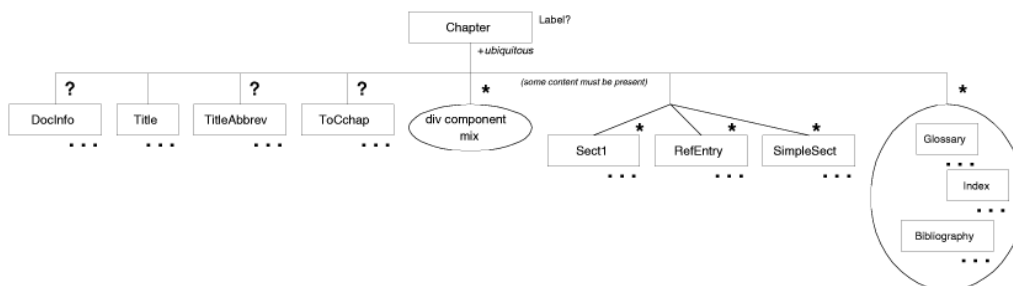


Figura 9. Diagrama de la estructura Chapter

El elemento Chapter contiene un DocInfo opcional, un Title, un TitleAbbrev opcional, un ToCchap para mantener en la tabla de contenido el nivel del capítulo, y cualquier número de elementos que puedan estar en el nivel del elemento %divcomponente.mix;

cualquier número de elementos Sect1, RefEntry, o SimpleSet, y cero o mas Indices, Glossary, y Bibliography.

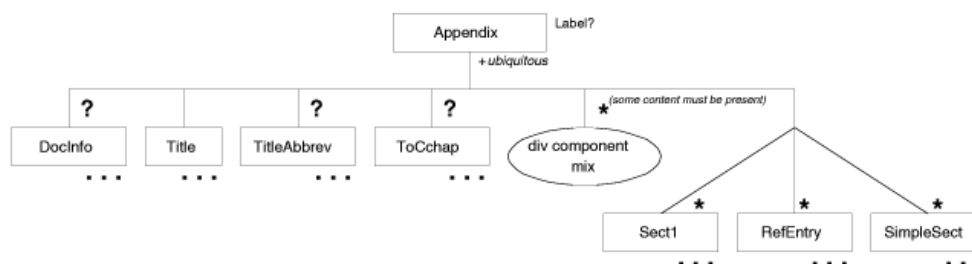


Figura 10. Diagrama de la estructura de Appendix

Appendix tiene el mismo modelo de contenido que Chapter, excepto que no puede contener Index, Glossary, y Bibliography en el final de DocInfo.



Figura 11. Diagrama de la estructura de DocInfo

DocInfo contiene metadatos del componente Book en el cual aparece. DocInfo se parece a BookInfo en que su contenido tiene un orden determinado: contiene cero o mas elementos Graphics que ilustran el componente, un Title, un TitleAbbrev opcional, un SubTitle, cero o mas AuthorGroups, cero o mas Abstracts, Un RevHistory opcional, y cero o mas LegalNotice.

3.1.2. Propósito general de los elementos Sect

Cinco niveles de secciones normales están disponibles: Sect1, Sect2, Sect3, Sect4 y Sect5. Estas pueden estar anidadas en ese orden. Además, SimpleSect puede ser usado dentro de cualquiera de los niveles numerados, y BridgeHead puede ser usado para simular la estructura de una sección. Por ejemplo, la misma información puede ser organizada así:

Sect1

Sect2

Sect3

o:

Sect1

Sect2

SimpleSect

o:

Sect1

Sect2

BridgeHead

El elemento Section contienen, en un determinado orden, un Title, un TitleAbbrev opcional, cualquier número de elementos %nav.class; cualquier número de objetos del nivel del elemento de el %divcomponent.mix; y opcional también un grupo del siguiente nivel por debajo del de la sección, un grupo de RefEntry, o un grupo de SimpleSects, seguido por cualquier número de elementos %nav.class. Una sección debe contener al menos un contenido de objetos del nivel o una subsección.

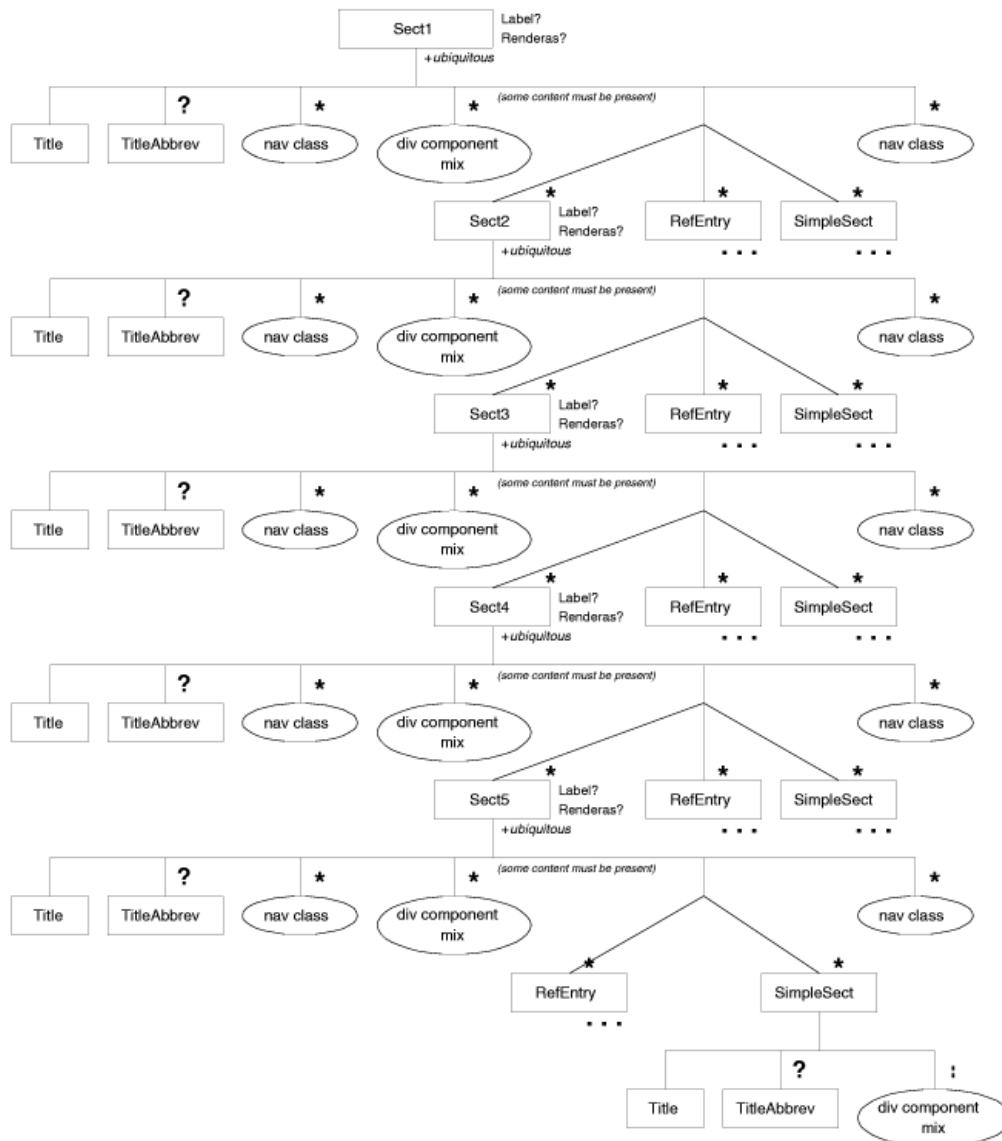


Figura 12. Diagrama de la estructura de Sections

Los elementos Section tienen un atributo Label opcional y un atributo Renderas opcional, los cuales indican la apariencia que se quiere para el título de la sección.

Los elementos SimpleSects son secciones atómicas que no pueden contener subsecciones, si en embargo, pueden contener elementos BridgeHeads.

Los elementos SimpleSect contienen un Title, un TitleAbbrev, y uno o mas objetos del nivel del elemento de la %divcomponent.mix.

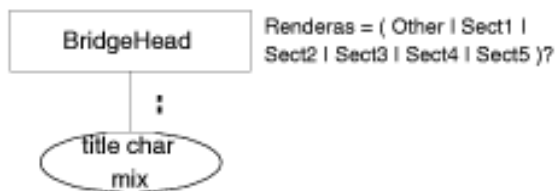


Figura 13. Diagrama de la estructura de BridgeHeads

El elemento BridgeHead posee un Title para un punto particular en el texto sin cambiar la jerarquía del documento. El BridgeHeads no podría aparecer en una tabla de contenidos. Contiene uno o mas elementos de la %title.char.mix, un atributo Renderas opcional, que indica la apariencia deseada del contenido de BridgeHead.

3.1.3. Tablas de contenido y listas de títulos

DocBook ofrece marcas para guardar información de navegación como Tablas de contenido e índices. Esta información puede ser manual, o puede ser generada, organizada en formato SGML, y usada para publicar un documento en SGML.

Un elemento ToC, o tabla de Contenido puede ser un componente por si solo o puede pertenecer a un elemento Book. ToC se subdivide para seguir la estructura de un Book: continuando con un DocInfo opcional, Title, TitleAbbrev. Un ToC puede tener cualquier número de ToCFronts, los cuales son las entradas para la parte inicial. Siguiendo el

ToCFronts, un ToC puede tener también uno o mas ToCparts (entradas para Parts) o ToCchaps (entradas para capítulos y apéndice), y puede tener cualquier número de ToCbacks (entradas para la parte final). Un elemento ToCpart comienza con uno o mas ToCentries (una apariencia para la entrada de la tabla de contenidos), entonces contiene cualquier número de ToCchaps. ToCentry tiene un atributo PageNum, que puede tener el valor físico de un número de página. Un ToCchap comienza con uno o mas ToCentries, entonces puede tener cualquier número de ToClevels, los cuales son entradas para Sect1s. Un ToClevel comienza con uno o mas ToCentries, entonces puede tener cualquier número de ToClevels2, y así sucesivamente hasta Toclevels5; que puede tener solo una o mas ToCentries. De esta manera, si tenemos una tabla de contenido que muestra las cabeceras de las secciones, las entradas del segundo nivel están anidadas dentro de las entradas del primer nivel, y así sucesivamente. Podríamos hacer un link de todo o parte del contenido de un ToCentry.

```
<toc>
```

```
<title>Table of Contents</title>
```

```
<tocchap>
```

```
<tocentry>Acknowledgements</tocentry>
```

```
<toclevel1>
```

```
<tocentry>How to Get the DocBook DTD Online</tocentry>
```

```
</toclevel1></tocchap>
```

```
<tocchap>
```

```
<tocentry>Development of the DocBook DTD</tocentry>
```

</tocchap>

</toc>

Un LoT es como un ToC pero que es usado para listas de tablas, figuras, o cosas por el estilo, y no tiene jerarquía. Un LoT contiene LoTentries, que podrían contener links, solo como ToCentries. LoTentries tiene un atributo PageNum y un atributo SrcCredit, este último para dar crédito a la fuente de la cual proviene la imagen.

3.1.4. Bibliografía

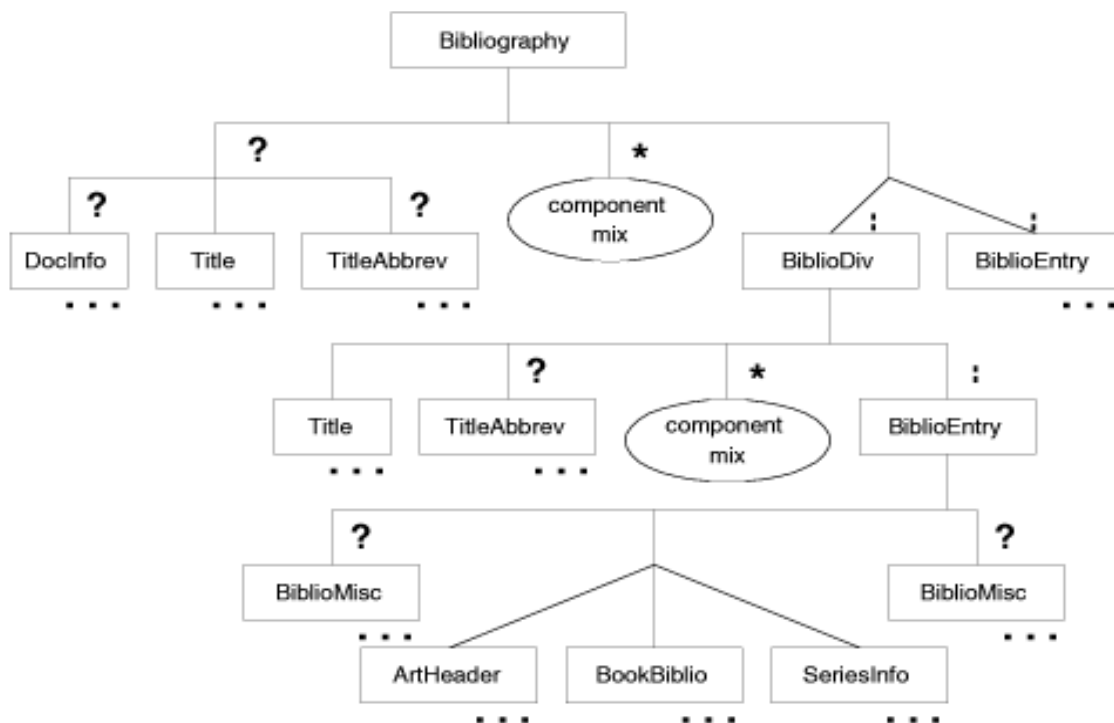


Figura 14. Diagrama de la estructura de Bibliography

El elemento Bibliography contiene un DocInfo opcional, Title y TitleAbbrev, cualquier número de objetos del nivel del elemento del %component.mix, y también uno o mas BiblioDiv o uno o mas elementos BiblioEntry.

BiblioDiv contiene un Title, un TitleAbbrev opcional, cualquier número de objetos del nivel del elemento %component.mix, y uno o mas elementos BiblioEntry.

BiblioEntry contiene un BiblioMisc opcional, seguido por un ArtHeader, BookBiblio, o SeriesInfo, seguido por un BiblioMisc opcional.

3.1.5.Glosary

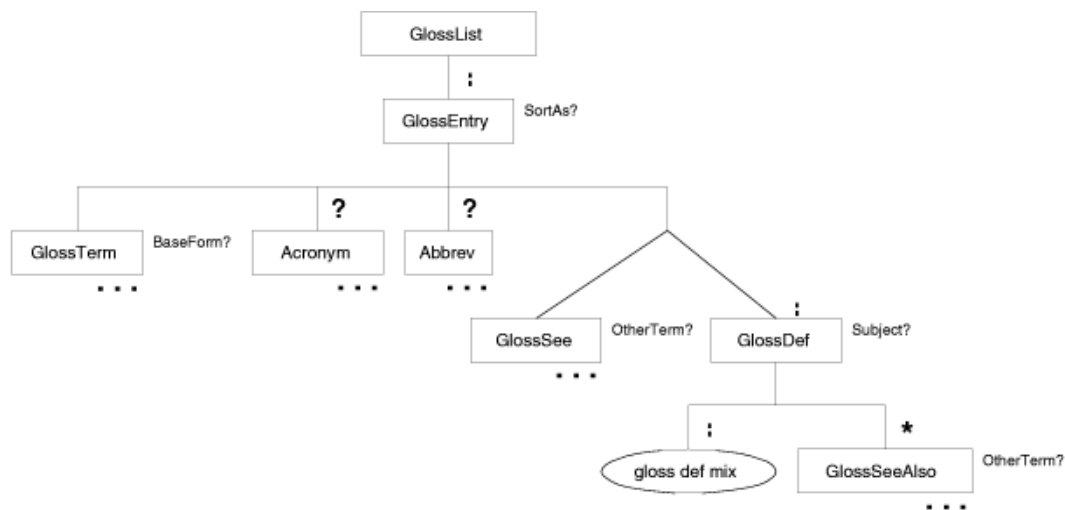


Figura 15. Diagrama de la estructura de Glossary

3.2.LA INTERFACE DE LYX

3.2.1.Operaciones básicas sobre archivos

LyX puede realizar la mayoría de las cosas que se pueden realizar con un procesador de texto. Separará las palabras y justificará los párrafos automáticamente. Basta acceder a un par de menús para ver cómo la mayor parte de los comandos simples (Archivo, Salir, Edición, Pegar, Imprimir) tienen los nombres que esperamos que tengan, están en el menú donde esperamos que estén, y funcionan tal y como esperamos que funcionen. A continuación una descripción de cómo realizar algunas acciones sencillas.

En el menú Archivo encontramos las nueve operaciones para cualquier procesador de palabras, todas ellas realizan las mismas tareas que en otro procesador de texto con pocas diferencias.

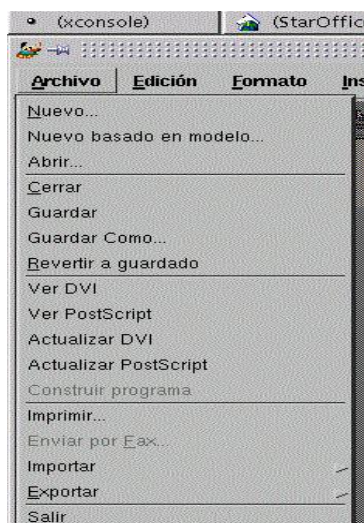


Figura 16. Menú Archivo

Opciones	Funcionalidad
Nuevo	Crear un nuevo archivo
Nuevo basado en modelo	Crear un nuevo archivo y la plantilla a usar
Abrir	Abre un archivo existente
Cerrar	Cierra archivo activo
Guardar	Guarda los cambios del archivo activo
Guardar Como	Guarda los cambios y pide nombre y ruta si no se ha guardado antes
Revertir a guardado	Es usado para volver a la última versión guardada antes de realizar los últimos cambios
Imprimir	Imprime el archivo
Salir	Termina la sesión con Lyx

Tabla 6. Funciones del menú Archivo

En el menú Editar encontramos:

Deshacer: LyX tiene capacidad para “deshacer infinitas veces”, lo que significa que puedes deshacer todo lo que hayas hecho desde que empezaste la sesión actual, aplicando una y otra vez Edición\Deshacer. Si se deshace demasiado, elegimos simplemente Edición\Rehacer para recuperar los cambios. Actualmente, el comando deshacer está limitado a 100 pasos. Tampoco funciona para todo (por ejemplo, en los cambios de formato de documento).

Cortar/Pegar/Copiar: Utilizan Edición\Cortar, Edición/Pegar, y Edición\Copiar para cortar, pegar y copiar. O pega automáticamente el texto seleccionado con el botón central del ratón.

Buscar/Reemplazar: Utilizan Edición\Buscar y Edición/Reemplazar para realizar una búsqueda sensible a las mayúsculas. En el menú que se despliega a tal efecto, podemos desplazarnos hacia delante y hacia atrás en la búsqueda mediante las flechas, y reemplazar aquellas palabras que hayas encontrado con el botón Reemplazar. La mayoría de los menús contextuales de LyX (incluyendo los de Buscar y Reemplazar, Índice General y Formato, así como los de matemáticas) son ventanas que pueden ser apartadas, en vez de cerradas¹.

Formato de caracteres: Se puede resaltar texto (lo que normalmente significa poner los caracteres en cursiva), ponerlo en negrita, o en Estilo Nombre (habitualmente en minúsculas, para nombres propios de personas) desde los botones interruptor en el menú Formato.

Barra de herramientas: Sus botones (justo debajo de los menús permiten realizar las funciones más usuales, como Pegar e Imprimir. Si se mantiene el cursor del ratón sobre alguno de los botones de la barra, una pequeña nota amarilla nos informará sobre la función concreta del botón.

Minibuffer: La franja gris en la parte de abajo de la ventana de LyX recibe el nombre de minibuffer. Se encarga de mostrarte toda clase de información útil. Por ejemplo, cuando guardas, te dice el nombre del fichero que acabas de guardar. También muestra algunos mensajes de error.

¹ Unos pocos menús como Archivo\Abrir, no te dejarán escribir nada en la ventana principal hasta que los cierres. Asegúrate de que el foco está en la ventana correcta cuando estés tratando de escribir en la ventana principal de LyX o introduciendo un comando en alguna ventana de diálogo.

3.2.2.WYSIWYM: El espacio en blanco en LyX

Una de las cosas más difíciles para los nuevos usuarios es acostumbrarse a la forma en que LyX maneja el espacio en blanco. Por mucho que se pulse Retorno de carro, sólo se conseguirá una única línea en blanco. Por mucho que se pulse la Barra espaciadora, sólo se conseguirá un único espacio en blanco. En una línea vacía LyX no permite poner ni siquiera un espacio. El Tabulador no adelantará ningún espacio, de hecho no hay tabulación. Tampoco hay ninguna regla en la parte superior de la página que permita definir tabulaciones o márgenes.

Muchos procesadores de texto comerciales están basados en el principio WYSIWYG: “lo que ves es lo que obtienes”. LyX, por el contrario, está basado en el principio “lo que ves es lo que quieres decir”. Se escribe lo que se quiere decir, y LyX se preocupará de la composición para que el resultado final quede bien. Un Retorno de carro gramaticalmente separa párrafos, y de la misma forma un espacio separa palabras, así que no hay ninguna razón para poner varios seguidos; un Tabulador no tiene función gramatical alguna, así que LyX no los usa. Con LyX se empleará más tiempo en el contenido del documento, y menos en la forma.

LyX tiene (muchas) formas de ajustar al detalle el formato del documento. Incluye espaciado vertical y horizontal —mucho más potentes y versátiles que múltiples espacios o líneas en blanco— y formas de cambiar tamaño y estilo de letra y alineación

de párrafos a mano. La idea es que se pueda escribir todo el documento concentrándose en el contenido, y solamente preocuparse de ajustar los detalles al final.

3.2.3.Entornos

La mayor parte del documento está formada por texto normal. Los títulos de sección (capítulos, subsecciones, etc.) permiten al lector saber que se va a tratar un nuevo concepto o idea. Ciertos tipos de documentos tienen entornos especiales. Un artículo de periódico tendrá un resumen y un título. Una carta no tendrá nada de eso, pero probablemente contendrá un entorno para la dirección del remitente.

Los entornos son una parte importante en la filosofía “lo que ves es lo que quieres decir” de LyX. Un entorno dado puede requerir un cierto estilo o tamaño de letra, sangrado, espaciado y otras características. Este problema se agrava cuando el formato exacto de un entorno puede cambiar: un periódico puede usar letra en negrita, de 18 puntos con párrafos centrados para los títulos, mientras que otros pueden usar párrafos justificados con letra cursiva de 15 puntos; idiomas distintos pueden tener diferentes convenios para el sangrado; y los formatos de bibliografía pueden variar ampliamente. LyX te evita tener que aprender todos los diferentes estilos de formato.

La caja de Entorno se sitúa al final a la izquierda en la barra de herramientas (justo debajo del menú Archivo). Indica qué entorno estás usando en cada momento. Mientras escribías tu primer documento, decía “Standard” (normal), que es el entorno por defecto

para texto. Podemos usar varios entornos en el un documento, pulsando sobre el icono con “la flecha hacia abajo” justo a la derecha de la caja de Entorno.

3.2.4.Secciones y Subsecciones

Si se quiere incluir una sección pulse en el menú entorno No se tiene que seleccionar la línea². LyX numera la sección y escribe el encabezado (título) en un tipo de letra mayor, por supuesto, el encabezado aparecerá de esta forma en el fichero dvi y en el documento impreso. Si se pulsa el Retorno de carro la caja de entorno cambia de “Section” a “Standard”. Se asume que los títulos de sección, como muchos entornos, terminan cuando introduces un Retorno de carro. Desde luego, el entorno Standard puede continuar a lo largo de varios párrafos. Los entornos de listas tampoco terminan con Retorno de carro. Siempre se puede saber el entorno en el que se está, mirando la caja de Entorno.

Niveles más profundos de sección son la subsubsección (Subsubsection), el párrafo (Paragraph), y el subpárrafo (Subparagraph). Si elegimos en el menú entorno una Subsection (subsección) la numera automáticamente teniendo en cuenta que pertenece a una seccion y la escribe con un tamaño de letra mayor que el de texto regular pero menor que el de un título de sección. Los títulos de párrafo y de subpárrafo no están numerados por defecto, y los subpárrafos están sangrados; Los encabezados de capítulo

² Si no hay nada seleccionado, LyX cambia el párrafo en el que estás escribiendo ahora al entorno elegido. Alternativamente, se puede cambiar varios párrafos seleccionándolos antes de elegir el nuevo entorno.

(Chapter) son realmente el nivel más alto de la jerarquía, por encima de las secciones, pero solamente se pueden utilizar en ciertos tipos (clases) de documentos.

Si se quieren usar secciones y subsecciones sin numerar, existen entornos para esto. Si cambias uno de los encabezados de sección al entorno `Section*`, LyX usará el mismo tamaño de letra que en las secciones normales, pero no la numerará. También están los entornos “no numerados” correspondientes a `Subsection` y `Subsubsection`.

3.2.5.Listas y sublistas

LyX tiene diferentes entornos para componer listas. Los variados entornos de listas evitan tener que pulsar el Tabulador un millón de veces cuando se está escribiendo un esquema, o de renumerar toda la lista cuando se quiere añadir un nuevo punto en mitad de ella. Así se pone más atención en el contenido de la lista.

- Una exposición de diapositivas podría usar las listas simples (etiquetadas con bolos) del entorno `Itemize` para describir los diferentes puntos.
- Un esquema usaría las listas numeradas (y sublistas etiquetadas con letras) del entorno `Enumerate` (enumeración).
- Un documento que describa varios paquetes de software usaría el entorno `Description` (descripción), en el que cada elemento de la lista comienza con una palabra en negrita.

- El entorno List (que no existe en LaTeX) es ligeramente diferente al entorno Description.

Los entornos de listas, al contrario que los encabezados, no terminan cuando se introduce un Retorno de carro. En vez de eso, LyX asume que se va a introducir el siguiente elemento de la lista. Si se desea más de un párrafo en un solo elemento de la lista, una forma de conseguirlo es mediante un Retorno de carro protegido, pulsando C–Retorno de carro. Para salir de la lista se selecciona el entorno Standard (o usar la combinación de teclas M–p s).

Mientras la lista esté seleccionada, se puede cambiar a los otros dos entornos, Description y List, para ver cómo son. Para ambos, cada elemento de la lista está compuesto por un término, que es la primera palabra del elemento, seguido de una definición, que es el resto del párrafo (hasta que se pulse Retorno de carro). El término se escribe en negrita (Description) o separado por un “Tabulador” (List) del resto del párrafo. Si quieres más de una palabra en el término, separa las palabras con Espacios protegidos, que se obtienen al pulsar C–Espacio y aparecen como pequeñas “ues” rosas.

3.2.6. Estrofas, Citas y otros

Hay dos entornos para separar las citas del texto que las rodea: Quote para citas cortas y Quotation para las más largas. Este entorno es el único sitio en LyX donde se permite usar varios espacios seguidos para permitir el sangrado del código. Puedes incluso

escribir poesía mediante el entorno Verse (estrofa), usando Retornos de carro para separar los versos, y C–Retorno de carro para separar líneas dentro de un verso.

3.2.7. Clases de texto y modelos

Diferentes tipos de documentos deben componerse de forma diferente. Por ejemplo, normalmente los libros se imprimen a doble cara, mientras que los artículos se imprimen a simple. Además, muchos documentos contienen entornos especiales: las cartas tienen entornos (como la dirección del remitente o la firma) que no tienen sentido en un libro o un artículo. Las clases de texto de LyX se encargan de estas grandes diferencias entre cada tipo de documento.

La clase Article (artículo) es la clase de texto que LyX usa por defecto, si se quiere usar otra clase de documento en el menú Clase dentro de Format/Documento, se elige la clase de documento que se va a utilizar.

El tamaño de letra, la impresión a una o dos columnas, o los encabezados de página son sólo algunas de las cosas en las que difiere el formato de composición de los distintos periódicos. Conforme la Era Digital ha ido madurando, éstos han empezado a aceptar presentaciones electrónicas, creando “ficheros de estilo” LaTeX para que los autores puedan enviar sus artículos correctamente maquetados. LyX también está preparado para esto. Así por ejemplo, ofrece soporte para composición (y entornos adicionales) para los

periódicos de la Sociedad Americana de Matemáticas mediante la clase de texto Article (AMS).

Estas son algunas otras clases de texto:

Clase	Descripción
Nombre	Comentarios
Article	Artículo (Simple cara, sin capítulos) (Ver Anexo A)
Article (AMS)	Formato y entornos de la Sociedad Americana de Matemáticas
Report	Informe (más extenso que el artículo a doble cara)
Book	Libro (informe + portada y contraportada) (Ver anexo A)
Slides	Transparencias (incluyendo FoilTex)
Letter	Carta (entornos adicionales para la dirección y la firma) (Ver anexo A)

Tabla 7. Clases de documentos

3.2.8. Título del documento

LyX (al igual que LaTeX) considera el título como una parte independiente, que puede incluir el título propiamente dicho, el autor, la fecha e incluso el resumen del documento³.

³ La clase carta no permite títulos

3.2.9. Etiquetas y referencias cruzadas

Se puede etiquetar una sección de un documento (o una subsección, o incluso, con menos frecuencia, un fragmento de texto cualquiera). Una vez que se realice se puede hacer referencia a esta sección desde otras partes del documento mediante referencias cruzadas. Podemos referirnos al número de sección o bien a la página donde aparece. Como sucedía con las secciones y las notas a pie de página, el propio LyX se encarga también de las referencias. La gestión automática de etiquetas y referencias cruzadas es una de las mayores ventajas de LyX (y LaTeX) sobre los procesadores de texto convencionales.

3.2.10. Notas a pie de página y notas al margen

Las notas a pie de página se pueden añadir usando el botón Insertar/ Nota a pie en la barra de herramientas⁴ o bien accediendo en el menú a Insertar\Nota al pie\. Pulsamos al final de la palabra “LyX” en cualquier parte de tu documento y en el botón Insertar Nota a pie. Una línea de pie de página se abrirá debajo de la línea en la que estabas escribiendo.

LyX se encarga de la numeración de las notas en el texto impreso. Si se añaden más notas, LyX las renumera. Como LyX se preocupa de esto, no hay necesidad de poner los números en el fichero LyX.

⁴ El botón muestra una flecha señalando texto en rojo, justo debajo de texto en negro

Una nota al pie puede ser cortada y pegada como texto normal. Además, se puede convertir texto normal en una nota, basta que se seleccione y se pulse el botón Insertar.

Las notas al margen se pueden añadir mediante el botón Insertar Nota al margen o bien el menú Insertar\Nota al Margen. Son como las notas a pie de página, salvo que:

- Los marcadores en pantalla dicen “margin” (margen) en vez de “foot”.
- Las notas se sitúan en el margen de la página, en vez de bajo el texto.
- No se numeran.
- Cuando una nota es plegada, se sitúa un signo de admiración en el margen, que no se verá en el texto impreso.

3.2.11. Bibliografía

Las bibliografías funcionan de manera similar a las referencias cruzadas. La bibliografía contiene una lista de referencias al final del documento que pueden ser referenciadas desde cualquier parte del texto. Al igual que los títulos de sección, LyX y LaTeX hacen el trabajo más fácil numerando automáticamente los elementos de la bibliografía y modificando las referencias cuando la numeración cambia. Al igual que los otros entornos, se pulsa el botón con una flecha y se activa el entorno Bibliography. De ahí en adelante cada párrafo que se escriba será una referencia.

3.2.12. Índice general

LyX facilita colocar un índice al principio del documento, haciendo click en el menú Insertar>Listas\Índice general, lo cual genera un índice al principio del documento, con todos los títulos de las secciones y subsecciones, con su numeración. No se permite editar el número de sección de ninguna forma, ya que LyX controla el numerado de secciones.

El índice no se imprime en la versión en pantalla del documento porque no se puede editar de ninguna manera. Sin embargo, se puede ver en una ventana separada pinchado con el ratón en el recuadro del índice o bien mediante Editar\Índice general⁵. Esta ventana es una herramienta muy práctica. Se puede usar para moverse a través del documento. Pulsando en una (sub)sección del índice se resaltará esa línea y el cursor se moverá a ese lugar del documento en la ventana de edición de LyX. También pueden usarse los cursores para moverse arriba y abajo en el índice.

Para deshacerse del índice, solo hay que borrar su marcador como cualquier otro carácter.

⁵ El comando del menú funcionará incluso si no se tiene recuadro de índice en el documento.

3.2.13. Soporte para Ecuaciones

LaTeX es utilizado por muchos científicos porque ofrece una gran calidad en el aspecto de las ecuaciones, evitando los caracteres de control usados por otros procesadores de texto y sus editores de ecuaciones. Sin embargo, muchos de estos científicos se sienten frustrados porque escribir ecuaciones con LaTeX se parece más a programar que a escribir. Afortunadamente, LyX tiene soporte WYSIWYM para las ecuaciones.

Este modo matemático tiene un par de diferencias con respecto al modo normal:

- El tipo de letra por defecto es de menor tamaño para unos pocos símbolos, como \sum y \int .
- Los subíndices y superíndices en las sumas y límites (no en las integrales) se escriben debajo, en lugar de seguir a los símbolos.
- El texto se centra

Aparte de estas diferencias, las expresiones en línea y en modo demostración son muy similares.

3.2.14. Importar documentos LaTeX: reLyX

Se Puede importar un fichero LaTeX usando el comando `Archivo\Importar\LaTeX`. Éste llamará a un programa Perl llamado reLyX, que creará un fichero `fich.lyx` a partir del

fichero `fich. tex` y lo abrirá. Si la traducción no funciona, se puede probar a ejecutar `reLyX` desde la línea de comandos.

`reLyX` traducirá la mayoría de los comandos legales de `LaTeX`, pero no todo. Dejará lo que no entienda en modo `TeX`, así que después de la traducción se puede buscar el texto en rojo y editarlo a mano para tratar de arreglarlo⁶.

3.2.15. BibTeX

`LyX` tiene un soporte de `BibTeX`, el cual permite construir bases de datos de referencias bibliográficas que pueden usarse en múltiples documentos. Selecciona en el menú `Insertar\Listas e Índice gral.\Referencia BibTeX` para incluir un fichero `bib`. Se hace click en el cuadro resultante de “Referencias `BibTeX` generadas” y se obtiene un menú `BibTeX`. En el campo `Base de Datos` se escribe aquello que se pondría dentro de las llaves del comando `\bibliography`.

Después de hacer esto, se puede hacer referencia a cualquier entrada de las bibliografías que se haya incluido mediante `Insertar\Referencia a Cita`. El programa se preocupará de ejecutar `BibTeX`.

⁶ Cuando `LyX` se instala, se crea un fichero ejecutable separado llamado `reLyX` en el mismo directorio que el propio `lyx` (i.e. `usr/local/bin/reLyX`)

3.2.16.Errores

A veces, al ejecutar LaTeX habrá errores, cosas que LyX o el propio LaTeX no entienden. Cuando esto sucede, LyX crea un recuadro de error (con la palabra “error” dentro). Pulsando sobre el recuadro se abrirá una ventana que muestra el mensaje de error concreto. Si se trata de algo se ha hecho mal con LyX, será un error de LyX.

4.HT://DIG

4.1.INDEXACION

4.1.1.Estructura de la Base de Datos

La base de datos va a estar compuesta, casi exclusivamente, por documentos en formato HTML correspondientes a los documentos que contiene la biblioteca. Estos documentos se encuentran alojadas en el servidor principal de la biblioteca.

4.1.2.Sistema Utilizado

El sistema utilizado es ht://Dig .

4.1.3.Resumen de funcionamiento

El sistema realiza esencialmente tres tareas necesarias para la creación de los índices y los mecanismos necesarios para la búsqueda en esos índices. Estas tareas, el sistema htdig las define como :

4.1.3.1.Digging

Antes de realizar una búsqueda es necesario que sea creada una base de datos que contenga los documentos a ser buscados. Estos se obtienen de las direcciones que se indican en la configuración de ht://Dig. En esta fase el programa actúa como un usuario de web excepto que sigue todos los 'links' que se indiquen en la página y se encuentren dentro del dominio indicado. El programa que realiza este proceso es ht://Dig.

4.1.3.2.Merging

Una vez que se ha creado la base documental, esta deberá ser convertida a un formato desde el cual pueda realizarse la búsqueda. Asimismo en esta etapa del proceso es posible agregar a la base de datos creada con anterioridad únicamente aquella información que haya sido cambiada o agregada. El programa que realiza este proceso es el htmerge.

4.1.3.3.Búsqueda

El proceso final es el de búsqueda sobre la base ya creada. Este proceso se realiza a través de un programa CGI que es invocado desde un formulario HTML. El programa htsearch es el encargado de realizar las búsquedas y el que realiza una salida en formato HTML que será vista por los usuarios del servicio y que contendrá los documentos que coincidan con las claves de búsqueda.

4.1.4.Programas Requeridos

El paquete ht://Dig fue desarrollado bajo Unix utilizando C++ y es necesario entonces un sistema Unix para su instalación y un compilador C y C++ (se utiliza el compilador C para la generación de algunas de las librerías que componen el paquete).

El sistema fue probado satisfactoriamente en las siguientes plataformas:

Linux

Sun Solaris 2.5 SPARC (using gcc/g++ 2.7.2)

Sun SunOS 4.1.4 SPARC (using gcc/gcc 2.7.0)

HP/UX A.09.01 (using gcc/g++ 2.6.0)

IRIX 5.3 (SGI C++ compiler. Don't know the version)

4.1.4.1.libg++

Si se va a compilar el paquete con GNU C++ es necesario instalar las librerías libg++ (que se distribuyen en forma separada del compilador GNU). Se pueden obtener estas librerías del repositorio de archivos GNU.

4.1.4.2.GNU 'make'

Si se utiliza Berkeley 'make' habrá problemas en la compilación del paquete, este no reconoce la sintaxis de los 'include' de los Makefiles de ht://Dig. Es preferible que instale GNU 'make' (o modifique los Makefiles necesarios con la sintaxis correcta).

4.1.4.3.Servidor Web Apache

Así como en todas las otras opciones del proyecto hemos utilizado el servidor web desarrollado por Apache, por ser:

- Potente
- Flexible.
- Gratuito (licencia GNU).
- Muy extendido (alrededor de un 50% de los servidores de WWW de Internet)

5.BIBLIOTECA VIRTUAL

Como resultado de estudios preliminares en la biblioteca, y de la investigación de la plataforma Linux (distribución SUSE 7.0), y las herramientas que éste facilita, creamos entonces un sistema llamado Biblioteca Virtual que no es mas que la combinación de dichas herramientas, para brindar a los estudiantes de la Corporación Universitaria Tecnológica de Bolívar, la posibilidad de hacer consultas a los documentos de la biblioteca.

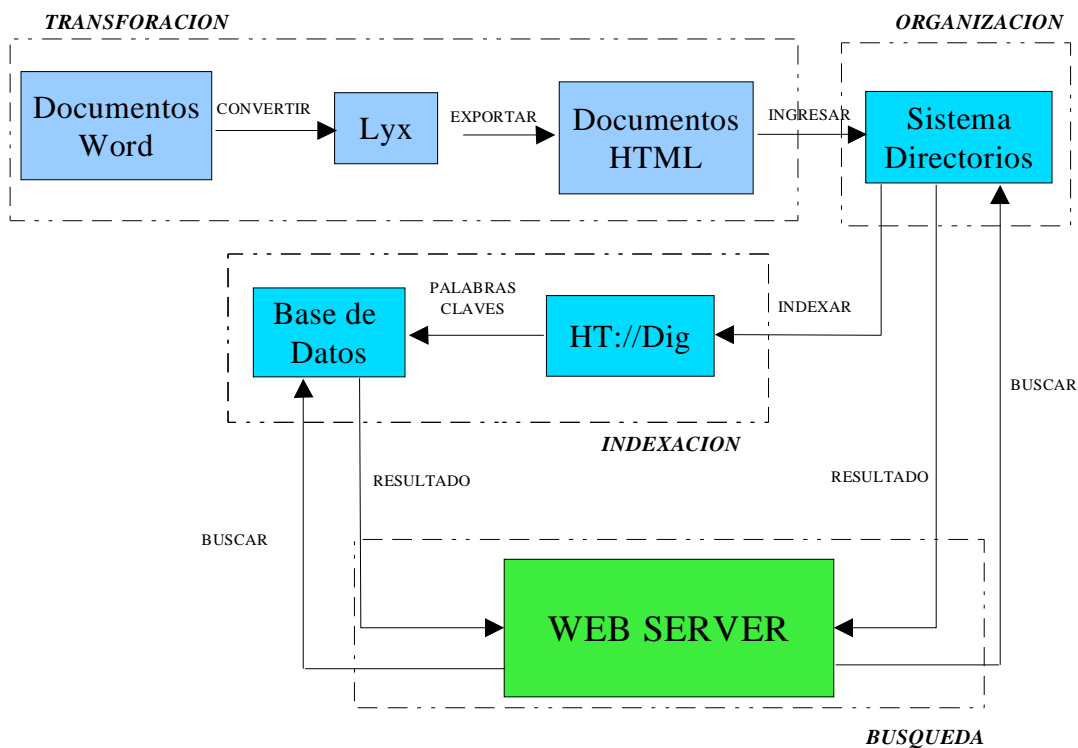


Figura 17. Esquema de la Biblioteca Virtual

Los cuatro procesos básicos del sistema son:

1. Transformación de los documentos al formato requerido
2. Clasificación y organización de los documentos
3. Indexación de los Documentos
4. Búsqueda

5.1.TRANSFORMACIÓN DE LOS DOCUMENTOS AL FORMATO REQUERIDO

Será el proceso mediante el cual, transformaremos los documentos del formato en que se encuentren a Latex, utilizando el editor de texto llamado LyX que nos facilitara exportarlos a HTML, entre otros.

Como nuestro fin era tener control total sobre la estructura y presentación de los documentos, decidimos utilizar un editor de texto de Linux llamado Lyx. Este trabaja con la estructura definida por la DTD DocBook (tratada en capítulo 3) de Linux, y unas hojas de estilo que nos permiten entonces concentrarnos en el contenido de los documentos y no en su presentación. Además nos permite exportar nuestros documentos a varios formatos: PDF, Postscript, HTML, y Latex. Los documentos exportados a HTML quedan fácilmente navegables desde su tabla de Contenido, por esto los documentos se encontrarán en HTML.

Si el documento que se va a ingresar a el sistema esta en formato *.DOC, debe ser transformado a un *.LYX de tal forma forma que pueda ser exportado a HTML (ver anexo A), en caso tal de que este en este formato solamente deberá ser clasificado para introducirlo en el directorio que le corresponda.

Otras ventajas de LyX que nos llevaron a escogerlo como editor de nuestros documentos son las siguientes:

- LyX posee soporte WYSIWYG para tablas
- LyX también permite incluir gráficos PostScript (o LaTeX puro).
- Soporte de control de versiones, usando RCS.
- LyX es altamente configurable. Todo, desde el aspecto de la ventana Hasta la forma de la salida, puede ser configurado de múltiples maneras. Gran parte de la configuración se lleva a cabo editando el fichero lyxrc. Actualmente, se edita el fichero lyxrc con un editor de texto. Los programadores esperan crear una interfaz para la configuración dentro de LyX.
- LyX está siendo desarrollado por un equipo de programadores en los cinco continentes. De esta forma, tiene mejor soporte para otros idiomas además del inglés (como holandés, alemán, griego, checo, turco, español) que muchos procesadores de texto. Puedes escribir documentos en otros idiomas, pero también se puede configurar LyX para que muestre los menús y los mensajes de error en otras lenguas.
- Los menús de LyX tienen asociadas combinaciones de teclas. Esto significa que se puede hacer Archivo\Abrir tecleando M-F seguido de O. Las asociaciones de teclas

también son configurables (y puede haber asociaciones incluso para algunos de los menús traducidos del inglés).

- LyX puede leer documentos LaTeX.
- Corrige los errores ortográficos de los documentos con Edición\Ortografía. El corrector ortográfico sólo comprueba desde el cursor hasta el final del documento.

5.2. CLASIFICACIÓN Y ORGANIZACIÓN DE LOS DOCUMENTOS

Este es el proceso que realizara la persona encargada de la administración de los documentos en la biblioteca, la cual tendrá que clasificarlos e introducirlos en el directorio que le corresponda .

Para agilizar y organizar la búsqueda de documentos, nuestro sistema exige que estos estén organizados en carpetas y subcarpetas como lo indica la figura 18. Habrá un directorio para cada facultad dentro del cual habrá un directorio para cada categoría (tesis, libro, artículo).

Para la clasificación e ingreso de los documentos a el sistema siga los siguientes pasos:

1. Identificar si el documento es de pregrado, postgrado o generales.
2. Si el documento es de pregrado o postgrado, identificar a que facultad o postgrado pertenece el documento y ubicarlo en el directorio correspondiente. En el directorio

generales irán los documentos que pueden ser de uso general, es decir, que pueden ser usados por todas las facultades y postgrados .

3. Dentro del directorio facultad, postgrado o generales ubicarlo según su categoría.

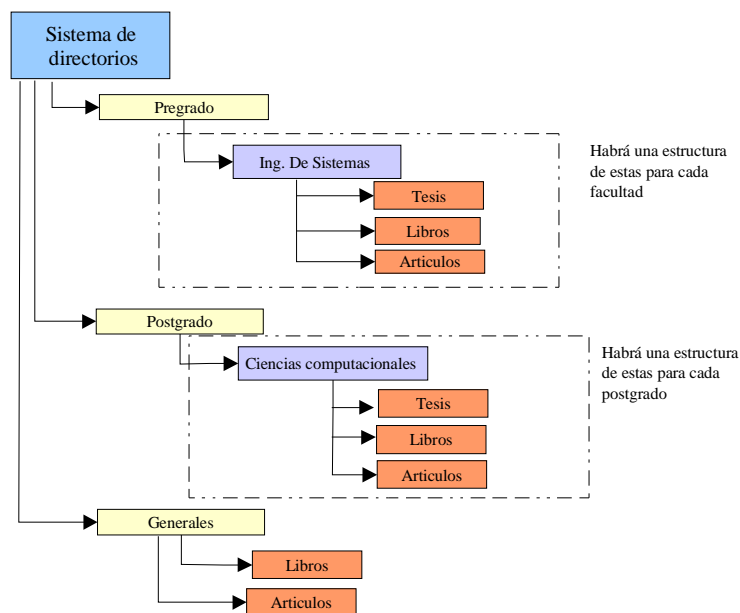


Figura 18. Sistema de directorios

5.3.INDEXACION DE LOS DOCUMENTOS

Para realizar la indexación y búsqueda de la información, escogimos el HT://Dig por ser una herramienta integrada de indexación y búsqueda que provee Linux, y porque es de fácil instalación y administración. Por otro lado necesitábamos un sistema flexible para permitir la adaptación de sus interfaces a nuestros propósitos, lo cual implicaba la configuración personal de las interfaces de búsqueda, el cambio de textos de salida y logos y demás elementos que intervienen en el proceso de búsqueda pero que,

fundamentalmente, ofreciera un resumen breve de cada uno de los documentos que coincidían con las claves buscadas.

Después de configurado el ht://Dig se deben indexarlos directorios corriendo el archivo *runding*. Cada vez que insertemos un nuevo documento a nuestro sistema de archivo tendremos que reindexarlo corriendo nuevamente el archivo mencionado.

Estas son algunas características del ht://Dig

- La instalación es sencilla (ver anexo B) y con unos cambios mínimos ya se puede tener un sistema de búsqueda completo y con buena respuesta en cuanto a velocidad y presentación de los datos buscados.
- Los procesos de actualización son automáticos y la instalación instala los scripts necesarios (incluso se proveen ejemplos de scripts que podemos incluir dentro del 'cron' de tareas del sistema).
- El paquete se ofrece completo y no es necesario instalar programas adicionales para su funcionamiento (con la salvedad de que es posible que no todos hayan instalado las librerías necesarias para compilar programas en C++).
- Es fácil de configurar para todos los dominios que manejamos y los datos son incluidos en una sola base de datos de forma que la búsqueda se realiza en todas las direcciones sin necesidad cambiar el formulario de búsqueda o indicar en él una base de datos diferente.

- Una importante ventaja de ht://Dig, para los webmaster de un sitio, es la posibilidad de enviar un correo electrónico en forma automática cuando sea necesario actualizar una página (htnotify). Por medio de META-TAGS es posible incluir en el diseño de la página una fecha de expiración, un casillero de correo electrónico y ht://Dig recordará 'automáticamente' pasada esta fecha, con un correo electrónico a la dirección indicada en el META-TAG, que es necesario actualizar la página. Cuando se manejan grandes volúmenes de información esto resulta particularmente útil a la hora de actualizaciones, modificaciones o cambios periódicos en las páginas que mantenemos.

5.4.BUSQUEDA

Las interfaces que facilitan la búsqueda están hechas en HTML y se utiliza el lenguaje PHP para poder organizar y mostrar los resultados. En la pagina principal se podrá encontrar información acerca de novedades, mejoras, información del proyecto entre otros.

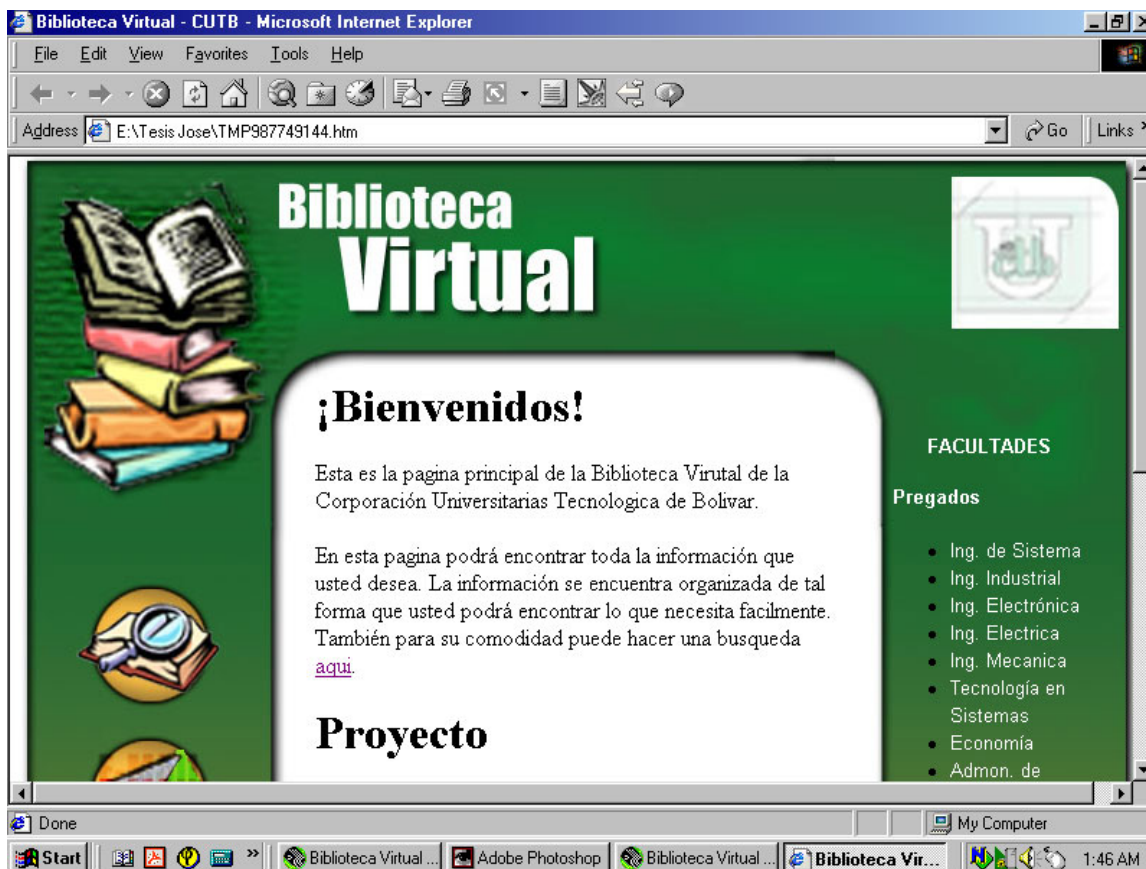


Figura 19. Interface pagina principal

Adicionalmente hay un enlace directo a la pagina de búsqueda. La pagina de búsqueda consta de una breve explicación de como realizar una búsqueda en sistema de la biblioteca virtual. El formulario de búsqueda consta de tres opciones las cuales nos dejan configurar como se realiza la búsqueda. Hay dos opciones las cuales son Buscar Todo y Buscar Cualquiera. En la primera opción se devuelven todos los documentos que contengan todas las palabras digitadas en la casilla de búsqueda. En la segunda opción se devuelven todos los documentos que contengan al menos una de las palabras digitadas en la casilla de búsqueda. Adicionalmente se le puede decir al motor de búsqueda como se

debe devolver los datos. Si deseamos que los ordene por titulo, hora o puntaje, y si deseamos ver una descripción de los documentos encontrados.

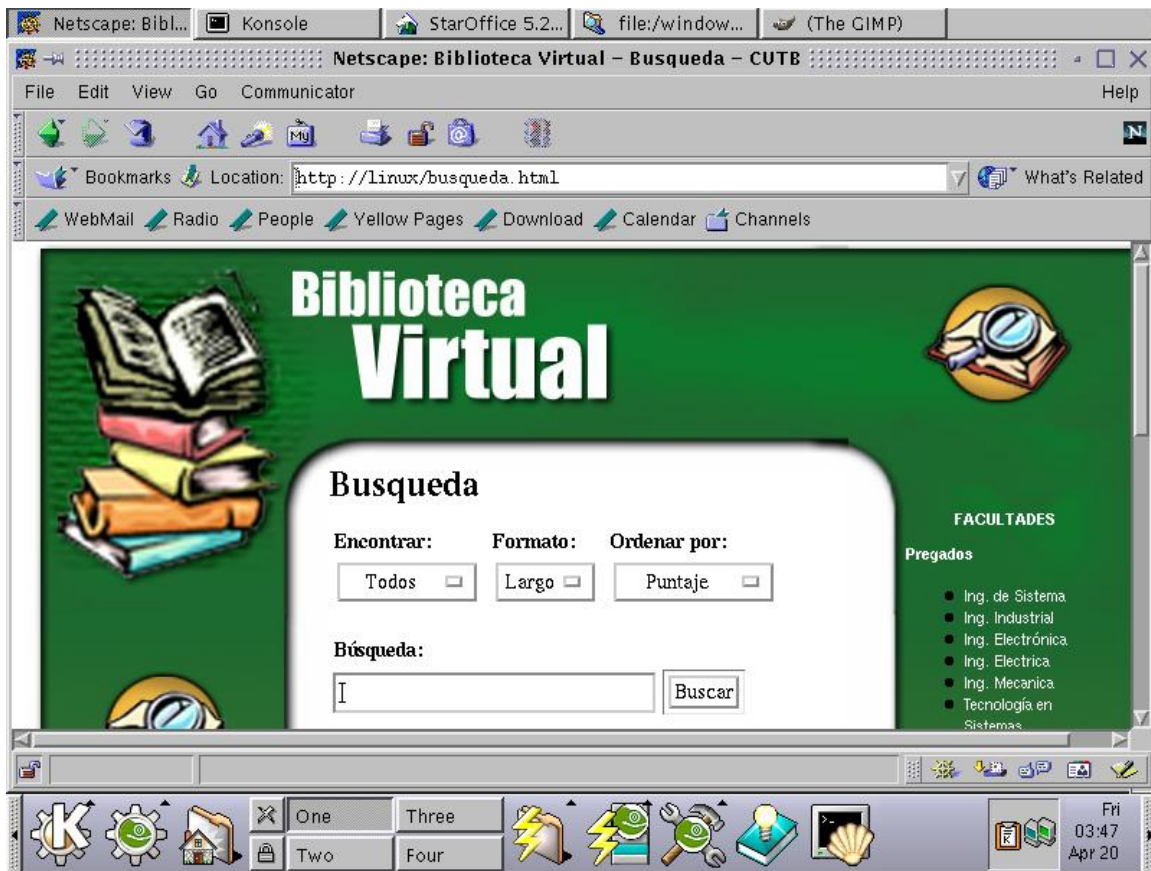


Figura 20. Interface de búsqueda

6.CONCLUSIONES

Los nuevos esquemas de bibliotecas, virtuales, son muy eficientes y aminoran en un 100% las desventajas presentadas por los sistemas convencionales (papel). Sin embargo a los usuarios de las bibliotecas virtuales, les gusta el hecho de tener la información de una manera más rápida en formato electrónico, pero aún les parece tedioso leer todo un texto desde su monitor, y se les hace necesario una copia impresa del documento. Pero, las bibliotecas convencionales gastan mucho dinero en comprar libros (de papel), y sus respectivas actualizaciones, y aún así no pueden satisfacer la demanda de los libros. Por otro lado cada vez atentamos mas contra el medio ambiente con la tala de árboles y la fabricación de CD's. Así que aunque todos nos estemos preparados para este cambio de esquema, si se hace necesario, si tenemos en cuenta que el impacto ambiental y los costos son menores si tenemos una Biblioteca Virtual.

Además del HTML, existen otros lenguajes de marcas que ofrecen ventajas sobre este. Ejemplo de estos el XML que también define marcas para la estructura de sus documentos pero además permite que el usuario de nombre a esas marcas para definir una estructura y tener control total sobre esta. Sin embargo, para los fines de este proyecto, el uso de XML, obligaría a las personas que quisieran colaborar con sus documentos a la biblioteca, a aprender dicho lenguaje que a todos no les puede parecer sencillo. Es por esto, que buscando una manera fácil de controlar la estructura y

presentación de los documentos, investigamos varias herramientas entre las cuales encontramos el LYX que es un editor de texto fácil de utilizar y además nos sirve para nuestros fines. De una manera sencilla también puede el usuario dejar su documento navegable transportándolo a HTML. Estas son las razones por las cuales decidimos tomar el HTML como formato electrónico para nuestros documentos, además de ser el lenguaje de marcas más conocido y soportado por todos actualmente.

Este proyecto es muy amplio. Tiene muchas alternativas pero aquí se han considerado sólo las más adecuadas a los objetivos e infraestructura de la universidad, que tiene en desarrollo el proyecto de montar su Intranet en LINUX. Por esta razón hemos utilizado la plataforma Linux y utilizado sus herramientas que son totalmente gratuitas.

Una vez que se vea el prototipo, se podrá ver la verdadera amplitud de este proyecto, lo que tendrá que esperar el tiempo necesario para tener suficiente información que los estudiantes puedan consultar. Como en todos los proyectos se presentaron algunos inconvenientes como algunas fallas en el editor de texto LYX, pero que no son tan graves como para desecharlo. Además la documentación de Linux y sus herramientas, por ser este un sistema operativo que ha sido desarrollado por muchas personas alrededor de todo el mundo, se encuentra en diferentes idiomas.

RECOMENDACIONES

- Mantener actualizadas las versiones de Linux y de las aplicaciones utilizadas para fines de la Biblioteca como el Ht://Dig, el LyX, y el apache.
- Asegurarse de que la partición destinada a la ubicación de la base de datos tenga espacio suficiente como para almacenar los datos luego de la indexación y para las actualizaciones de nuestros documentos de acuerdo a los volúmenes y frecuencias de actualización.
- Teniendo en cuenta que nuestro servidor prestaría dos servicios únicamente, las especificaciones del equipo para que el sistema funcione de una manera óptima son las siguientes:
 - Pentium III de 900 Mhz
 - 128 Mb en memoria RAM
 - Disco duro de 20 o mas Gb
- Seguir paso a paso el proceso de configuración del sistema y modificar luego de este proceso, los archivos necesarios.
- Como el editor de texto LyX facilita la conversión del texto en formato de LyX a HTML, quedando navegable a través de la tabla de contenido. Sugerimos que la universidad exija a los estudiantes el uso de LyX para la transcripción de sus trabajos

de grado, ya que los documentos de la Biblioteca Virtual deben tener la estructura y presentación que el LYX. Además de entregar también el documento navegable en HTML, lo cual se hace exportando el documento de LYX a HTML.

- Para el uso de La Biblioteca Virtual fuera de línea, se recomienda, darle una IP fija al servidor de esta.
- Fomentar la investigación de la plataforma LINUX, creando un seminario para los estudiantes de Ingeniería de Sistemas inicialmente, de manera que los alumnos de este conozcan profundamente sus herramientas y puedan contribuir con la documentación de Linux en español y el desarrollo de aplicaciones bajo este sistema operativo.
- Es necesaria, una capacitación para los empleados de la biblioteca, de modo que ellos conozcan el sistema y las herramientas a usar, para que así estén en capacidad de colaborar a los alumnos con la entrega de los documentos en HTML y además ayuden con la transcripción de los documentos que no se encuentren en HTML.
- Recomendamos tomar de la biblioteca, los trabajos de grados y artículos, que se encuentren en formato electrónico y hayan sido entregados a partir del año 2000, los cuales deben ser transformados a LYX y exportados a HTML.
- Le recomendamos utilizar la distribución de Linux SUSE 7.0, la cual contiene los paquetes HT://Dig, Apache, PHP y LYX. En caso de no tener esta recomendamos bajar de Internet los diferentes paquetes por separado e instalarlos y configurarlos en la versión de Linux que se tenga.

BIBLIOGRAFIA

HUGHES, Sterling. PHP: Developer's Cookbook. 1 ed. Indianápolis, Indiana: SAMS, 2000. 486p.

YOUNG, Michael J. Aprende XML ya: El libro que necesita para aprender a programar con XML. 2 ed. Madrid, España: Mc Graw Hill, 2000. 344p + 1 CD.

MOHR, James. LINUX: Recursos para el usuario. 1 ed. México: Prentice Hall, 1999. 788p + 1 CD.

ANONIMO. LINUX: Máxima seguridad. Edición especial. España: Prentice Hall, 2000. 780p + 1 CD.

HISPALINUX. Modificación actual 2001. Proyecto Lucas.

[Http://lucas.hispalinux.es/](http://lucas.hispalinux.es/)

UNIVERSIDAD DE ALICANTE. 1999. Biblioteca virtual de Cervantes.

[Http://cervantesvirtual.com](http://cervantesvirtual.com)

SCHERPBIER, Andrew. 2000. ht://Dig [WWW](#) Search Engine.

<http://www.htdig.org/>

ANEXOS

ANEXO A. COMO HACER TUS DOCUMENTOS EN LYX

A continuación presentamos un manual que podrá ayudar a cualquier usuario de LyX a editar sus documentos, que pueden ser: un trabajo de Grado, un artículo, o una carta. Este manual tienen en cuenta que al momento de visualizar una tesis, no es necesario que veamos portadas, guardas y cartas. Para nuestros fines es importante que el documento tenga el título, los autores, su director y su fecha de realización.

Es importante saber antes de comenzar el manual que para insertar una tilde al documento debe escribirla, pulsar la tecla backspace y después colocar la letra que irá tildada.

Las partes de este manual son las siguientes:

1. Como pasar tu tesis en LyX
2. Como escribir una carta usando modelos
3. Como redactar tus artículos con Lyx

1. COMO PASAR SU TESIS EN LYX

1.1 EDICION DE PRELIMINARES

Los preliminares, según el formato de Trabajo de Grado de la Universidad Tecnológica de Bolívar son: Glosario, Listas especiales (figuras, tablas, anexos), Resumen Analítico y Tabla de contenido.

Abra el editor de texto LyX y en el menú Archivo\Nuevo, crear un nuevo documento. Se abrirá una ventana que le indicará en donde está quedando guardado el documento, que por defecto se llama nuevo–archivo. Haga click en O:K y el cursor estará esperando para empezar a editar el documento.

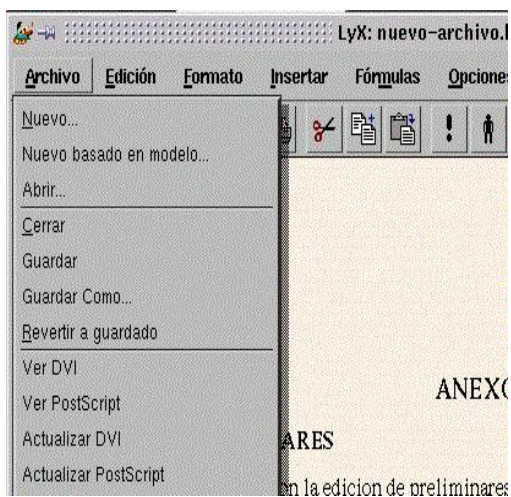


Figura 21. Formato para crear un nuevo documento

Ahora vamos al menú Formato\Documento, y verá una ventana como la siguiente:

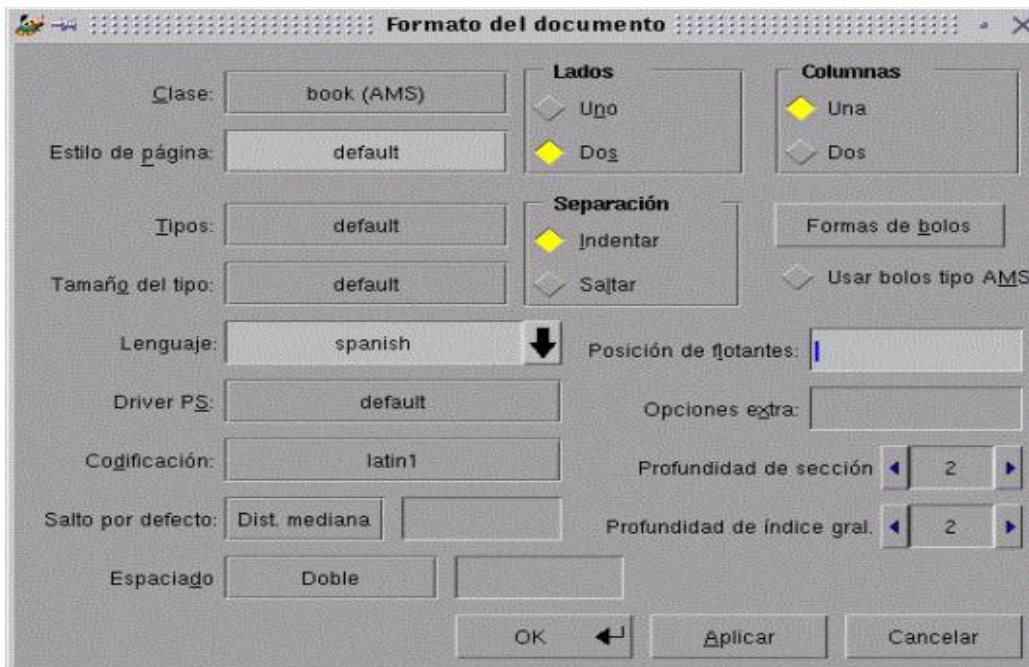


Figura 22. Formato para cambio de clase de un documento

En **clase**, podrá elegir entre una lista de diferentes tipos de documentos (libros, artículos, cartas, etc.) el tipo Book (AMS), que es el adecuado para sus propósitos. Las otras opciones como podrá ver, le permiten configurar el tipo y el tamaño de letra, interlineado, entre otras cosas. Cabe aclarar que estas opciones las podrá cambiar cuando termine el documento si es que no le ha gustado el tipo de letra que la plantilla use por defecto.

Entonces, podemos comenzar por ingresar el título haciendo click en el botón que contiene una Flecha mirando hacia abajo. Veremos una lista de entornos entre los cuales elegimos Title.

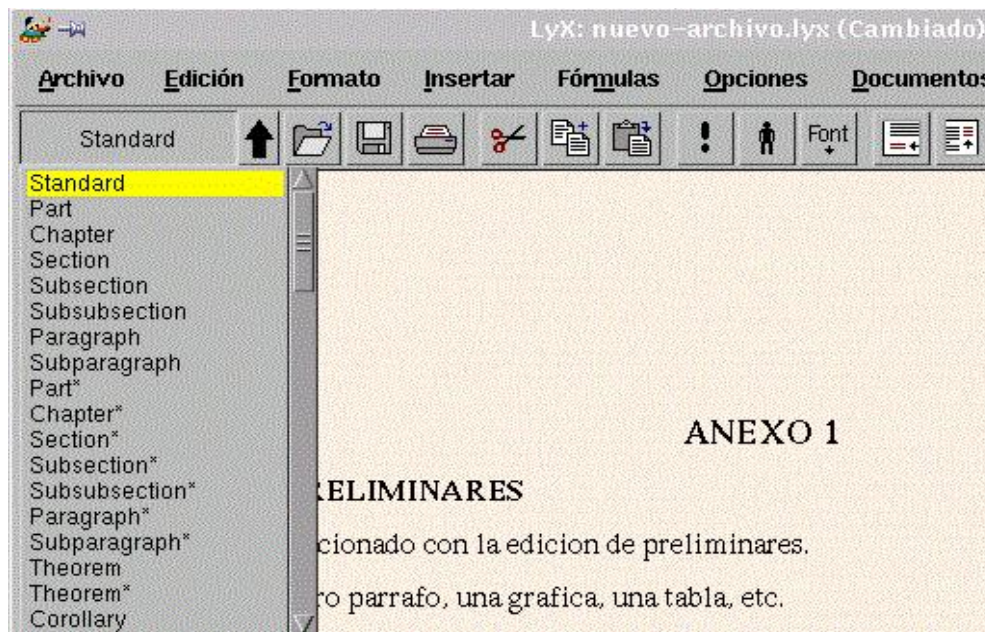


Figura 23. Formato para la elección de entornos

Escriba entonces el título y oprima la tecla Enter. Verá que inmediatamente en la ventana de entornos aparecerá standar que es el entorno que Lyx coloca por defecto. Continúe la edición de la portada, eligiendo de la lista de entornos Author y escribiendo los nombres del autor o autores del proyecto. Si usted cambia de entorno y quiere que la siguiente línea también tenga el mismo entorno, por ejemplo en el caso de los autores, si ingresamos uno y tecleamos Ctrl+Enter, la línea siguiente también tendrá el entorno Author. Así puede hacerlo con todos los entornos.

1.1.1 Glosario

Para insertar el Glosario diríjase a la lista de entornos y escoja el entorno Section*. Escriba aquí el título Glosario en mayúscula sostenida. Pulse enter y entonces busque ahora en la lista de entornos Description. Escriba entonces la primera palabra del glosario, pulse solo enter y continuará en el entorno Description hasta que usted le indique que debe utilizar algún otro entorno. Las palabras aparecerán en negrita y justo después de los dos puntos comenzarán a aparecer normales.

1.1.2 Listas de tablas y figuras

Estas listas, después de tener el documento listo y haber insertado todos estos elementos dentro del mismo, las generamos automáticamente haciendo uso de los comandos que lyx utiliza para este fin y que se encuentran en el menú Insertar>Listas e Índice Gral.

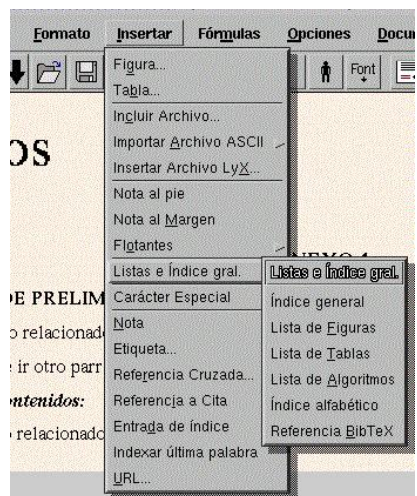


Figura 24. Formato para insertar tablas y listas

1.1.3 Resumen Analítico

En la lista de entorno elija Title y escriba Resumen. Después de teclear Enter entonces escriba el resumen de su trabajo permaneciendo en el entorno standar.

1.1.4 Tabla de Contenido

La tabla de contenido se encuentra al principio del documento pero ésta la vamos a incluir después que lo tengamos listo. Lyx le facilita hacer el índice con un solo comando. Elija en el menú Insertar\Listas e Índice gral.\Índice general. Aparecerá una caja (también conocida como recuadro) con las palabras “Índice general” en la primera línea del documento.

Esto puede no parecer muy útil. Sin embargo, si observa el fichero dvi, verá que se ha generado un índice con todas las secciones y subsecciones de su documento. Una vez más, si reordena las secciones o añade alguna, estos cambios se verán reflejados en el fichero dvi cuando lo actualices.

El índice no se imprime en la versión en pantalla del documento porque no puede editarlo de ninguna manera. Sin embargo, puede mostrarlo en una ventana separada pinchado con el ratón en el recuadro del índice o bien mediante Editar\Índice general. Esta ventana es una herramienta muy práctica. Puedes usarla para moverte a través de tu documento. Pulsando en una (sub)sección del índice se resaltará esa línea y el cursor se moverá a ese lugar el documento en la ventana de edición de LyX. También puede usar los cursores para moverse arriba y abajo en el índice. Puede que le resulte conveniente dejar esta ventana abierta a lo largo de las sesiones de edición.

Para deshacerse del índice, puede borrar su marcador como cualquier otro carácter.

1.2. EDICION DEL CUERPO DEL DOCUMENTO

1.2.1 Introducción

Para incluir la introducción elija el entorno Section*, la cual no va enumerada y escriba introducción. Después de teclear Enter entonces escriba todo el texto de su introducción dejándolo en el entorno Standar.

1.2.2 Capítulos

Cada vez que vaya a insertar un Capítulo escoja en la lista de entornos Chapter y escriba el título del capítulo. Seguido puede es escribir tantos párrafos necesite usando el entorno Standar, tecleando Enter entre párrafos para separarlos. Dentro de los capítulos podemos tener subcapítulos y subsubcapítulos. Estos los vamos a insertar usando los entornos Section y Subsection respectivamente dentro de Chapter. Recuerde que no necesita numerarlos porque LyX se encarga de esto. Si en un capítulo necesita insertar, una tabla, una gráfica, notas de pie de página, o notas al margen, en el menú Insertar puede encontrar los comandos para realizar dichas operaciones.

Si haces click en **Nota al Pie**, una línea de pie de página se abrirá debajo de la línea en la que estabas escribiendo. En el extremo izquierdo verás la palabra “foot” (pie) escrita en rojo sobre fondo gris. El resto de la línea está enmarcada en rojo; aquí es donde escribirá la nota. LyX sitúa el cursor al principio de la línea.

Si pulsa en la palabra “foot”. La línea de la nota desaparece, dejando “foot”, subrayada en rojo, mostrando el sitio donde aparecerá el marcador de la nota en el texto impreso. A esto se le denomina “plegar” la nota. Puede desplegar la nota en cualquier momento y volver a editar el texto si lo quiere pulsando de nuevo en el marcador rojo. La respuesta es que LyX se encarga de la numeración de las notas en el texto impreso, por lo tanto no tiene que numerarlas.

Para insertar las **Gráficas**, hacer click en Figura. Te aparecerá un recuadro en el que seleccionará el tipo de imagen que va a utilizar. Hace click en o.k y verá un cuadro como el siguiente:

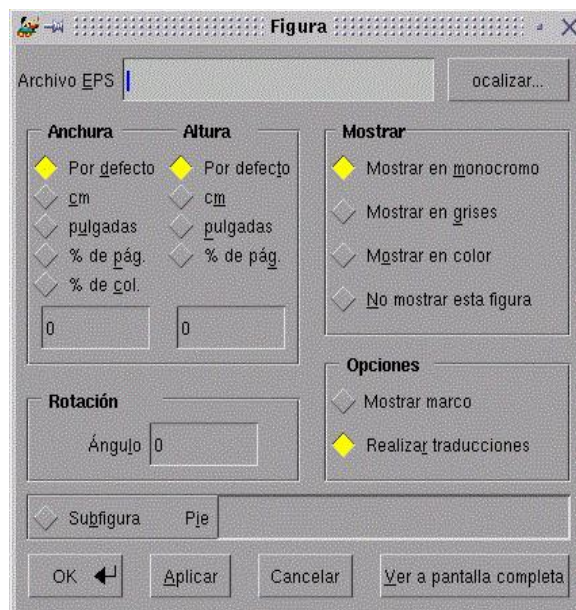


Figura 25. Formato para insertar imágenes

En este cuadro busque la imagen que desea insertar haciendo click en el botón Localizar, configure las otras propiedades que observa en el cuadro y entonces haga click en o.k.

Las gráficas deben estar en formatos *.ps ó *.eps, por lo tanto debe cerciorarse de que las gráficas estén en alguno de estos formatos al momento de insertarlas. Para esto Linux tiene sus editores de gráficos (por ejemplo el GIMP), que le permiten transformar sus imágenes al formato requerido.

Si queremos incluir **Tablas** en el menú insertar haga click en Tabla. Verá entonces un cuadro que le preguntará cuantas filas y cuantas columnas desea, le provee esta información y hace click en o.k. Verá entonces su tabla, a la cual le podrá cambiar ciertas propiedades haciendo click derecho sobre esta, en un cuadro como el siguiente:

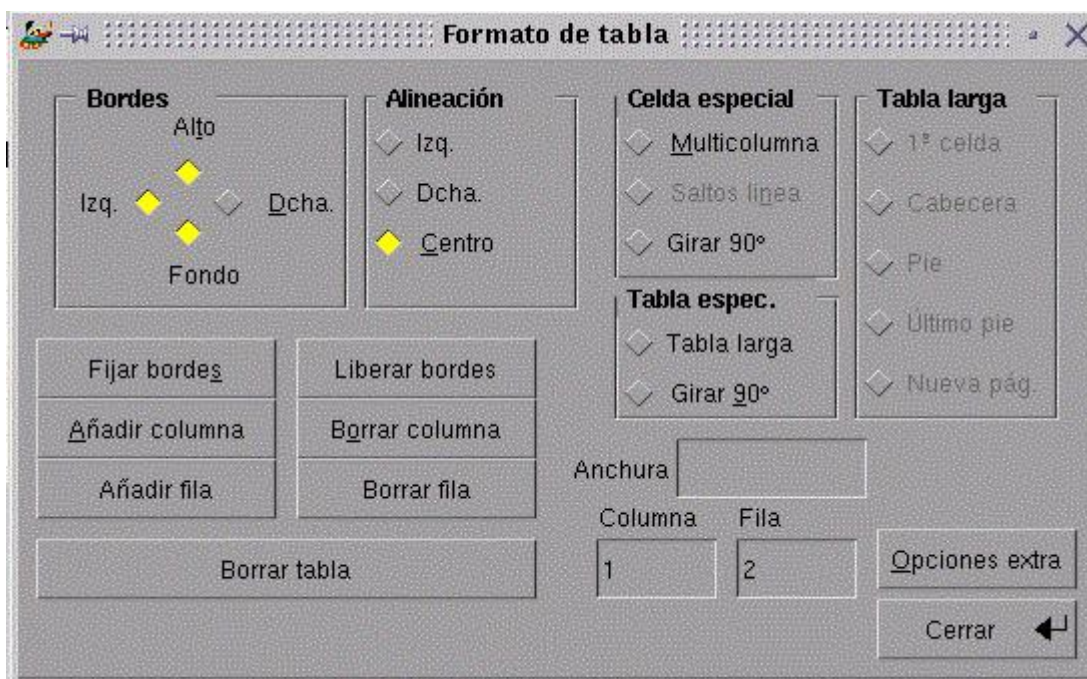


Figura 26. Formato de tablas

1.2.3 Conclusiones y Recomendaciones

Estas partes del documento las podemos insertar con un elemento Chapter debido a que necesitamos que estas estén enumeradas. Entonces en la lista de entornos, seleccione Chapter y escriba como título Conclusiones o Recomendaciones según el caso. La información correspondiente a la parte que esté introduciendo, insértela usando el

entorno Standar y si necesita alguno de los elementos como una lista con viñetas o definiciones, use los entornos Itemize y Description respectivamente.

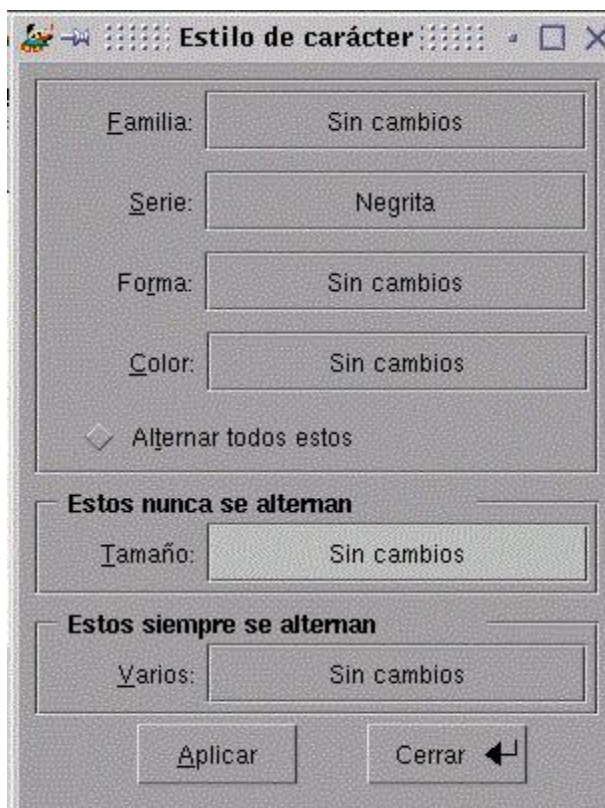


Figura 27. Formato para el cambio de tipo de letras

Si lo que desea es cambiar las propiedades a las letras, hágalo igual que en otro procesador de texto común. Seleccione las celdas a las cuales le desea hacer los cambios, vaya al menú Formato\Caracter y realice los cambios en la siguiente ventana.

Estas opciones las puede cambiar haciendo click en los espacios en los cuales dice por defecto “Sin cambios”, y escogiendo de la lista que se le desplegará la opción que desee.

Cabe anotar que esto aplica para cualquier parte del texto no solo para las tablas.

Las expresiones matemáticas se escriben en modo matemático o de fórmulas. Para entrar en dicho modo, sólo tiene que pinchar en el botón de la barra de herramientas con un $\frac{a+b}{c}$ escrito en azul. LyX abrirá un pequeño cuadro azul, con un rectángulo magenta a su alrededor. El cuadrado azul es el punto de inserción, que le indica que está esperando a que inserte algo, y el rectángulo indica que está en el modo matemático. LyX ha situado el cursor en el cuadro azul, así que introduzca su fórmula. La expresión se escribe en azul, y el cuadro azul desaparece tan pronto como el punto de inserción deja de estar vacío. Ahora pulsa Esc para dejar el modo matemático (nota: pinchar en el botón modo de fórmulas otra vez no te servirá). El rectángulo magenta desaparece, dejando el cursor a la derecha de la expresión. Las expresiones matemáticas se escriben en modo matemático o de fórmulas.

1.3 EDICION DE LOS COMPLEMENTARIOS

1.3.1 Bibliografía

La bibliografía contiene una lista de referencias al final del documento que pueden ser referenciadas desde cualquier parte del texto. Al igual que los títulos de sección, LyX y LaTeX hacen tu trabajo más fácil numerando automáticamente los elementos de la bibliografía y modificando las referencias cuando la numeración cambia.

Vaya al final del documento y active el entorno Bibliography. Ahora, cada párrafo que escriba será una referencia. LyX pone automáticamente un número encerrado en un recuadro antes de cada referencia. Pinche con el ratón en el recuadro, y se abrirá una ventana de diálogo Elemento de Bibliografía. El primer campo, la clave, le sirve para referirse a esta entrada desde el documento LyX. Por defecto es un número.

1.3.2 Anexos

Debido a que los anexos no tiene numeración, vamos a introducir los anexos utilizando el entorno Part*, escribe el título Anexos y continúa introduciendo los diferentes anexos utilizando los entornos Chapter*, Section*, Subsection*, Subsubsection*, según los vaya necesitando. Utilizamos estos entornos porque LyX no les coloca numeración y nosotros le podemos colocar la que deseemos, y aparecerán en la tabla de contenido cuando la genere automáticamente.

1.4 CONVERTIR DOCUMENTOS LYX A OTROS FORMATOS DE TEXTO

Para convertir documentos LyX a otros formatos de texto selecciona Archivo\Exportar\ como LaTeX, DVI, PostScrip, Ascii, HTML. Se creará un fichero cualquiera .tex a partir del fichero cualquiera .lyx que se esté editando. Al fin y al cabo, el programa siempre genera ficheros temporales cuando visualiza o imprime los documentos, así que no tiene ningún problema para hacer esto.

1.5 DE .DOC A .LYX

Si usted ya tiene su documento en formato .doc y quiere pasarlo a LyX deberá seguir los siguientes pasos:

1. Guarde su documento .doc como un archivo de texto. Escoja el menú Archivo → Guardar Como. En el recuadro Tipo de Archivo seleccione Documento Texto.
2. En LyX escoja el menú Insertar → Archivo ASCII → Como Párrafo. Seleccione el nombre del Archivo texto y presione abrir.
3. Con las técnicas ya mencionadas recorra el documento y va dándole formato a las diferentes secciones del documento seleccionado y buscando en la casilla de entorno el tipo de formato que corresponde a la sección seleccionada.

2. COMO ESCRIBIR UNA CARTA UTILIZANDO LOS MODELOS

Una de las clases de texto más populares es la carta. Una forma de escribir una carta sería abrir un Nuevo archivo, y elegir en Formato\Documento (ver figura X) y escoger en la lista de Clases Letter. Aunque esta es la manera más obvia de hacerlo, supone trabajo de más. Cada vez que escribe una carta de negocios pone su dirección, la del destinatario, el cuerpo, la firma, etc. Por tanto, LyX ofrece un modelo para cartas, que contiene un ejemplo de carta; una vez que tiene el modelo, sólo tiene que sustituir un par de cosas cada vez que quiera escribir una nueva carta.

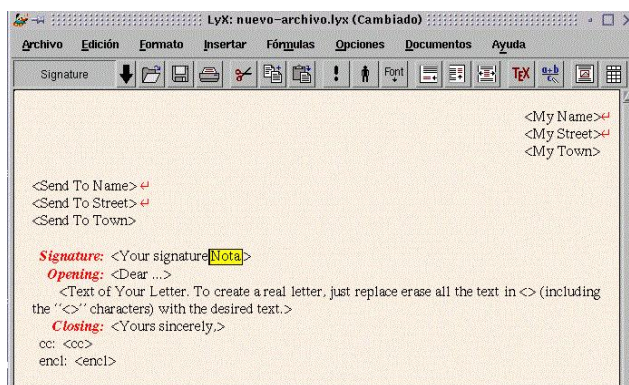


Figura 28. Formato de carta

Abra un archivo nuevo con Archivo\Nuevo basado en Modelo. Tras decidir un nombre para el nuevo archivo, elija latex_letter.lyx en el menú Seleccionar Modelo. Guarde e imprima el fichero para ver cómo se componen los distintos entornos, o simplemente vaya al menú archivo y escoja la opción Ver Postscript o DVI.

En el menú Entorno, verá algunos entornos, como My Address (dirección del remitente), que no están disponibles en otras clases. Otros, como Quote y Description, son familiares. Puede jugar con ellos para ver cómo funcionan. Comprobará por ejemplo, que en el entorno Signature (firma) la palabra “Signature:” en rojo antecede al texto de la firma. Esta palabra no se muestra en la verdadera carta, como podrá ver si la imprime. Sólo está ahí para que sepa dónde va la firma. Tenga en cuenta también que no importa dónde esté situada la línea Signature. Recuerda que LyX es WYSIWYM, así que puede poner el entorno Signature en el lugar que quiera, él sabe que en la salida impresa la firma debe ir al final.

Un modelo es simplemente un fichero de LyX. Esto quiere decir que puede completarlo con su dirección y su firma y guardarlo como un nuevo modelo. A partir de ahora, siempre que quiera escribir una carta ahorrará tiempo usando tu nueva plantilla.

Los modelos pueden ahorrar muchísimo tiempo, así que le aconsejamos usarlos siempre que pueda. Además, le pueden ayudar a usar algunas de las clases de texto más elaboradas y complejas.

3. COMO REDACTAR SUS ARTICULOS CON LYX

Si quiere redactar un artículo, cree un nuevo archivo desde Archivo\Nuevo basado en Modelo y aquí escoja el modelo aapapper.lyx.

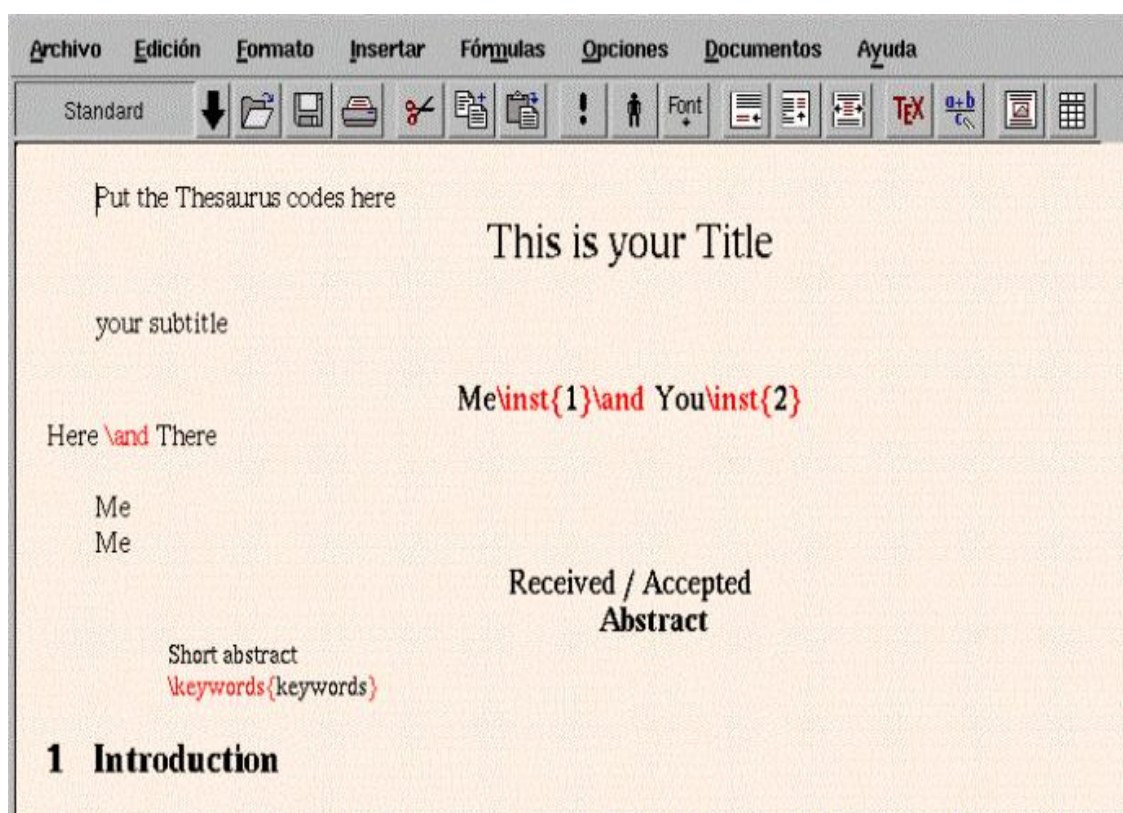


Figura 29. Formato para la realizacion de articulos

Lo primero que debemos hacer es llenar la información de la Cabecera, que consiste en ocho entradas algunas de las cuales son opcionales. Estas son:

- Thesaurus (requerido)
- Title (requerido): donde dice This is your Title
- Subtitle (opcional)
- Author (requerido): se encuentran en donde dice Me\inst{1}\andYou\inst{2}. Borre esa línea y escriba su nombre teniendo en cuenta que el entorno debe ser Author.
- Address(requerido): En donde dice Here\andThere. Borre la línea y escriba la dirección a la que va dirigida teniendo en cuenta que el entorno debe ser Address.
- OffPrints (opcional)
- Email (opcional)
- Date (requerido): se encuentra en donde dice Received/Accept. Fecha en que fue recibido y la fecha en la que fue aceptado.

Si usted desea colocar una nota de pie de página para indicar el estado de su dirección actual, insértela normalmente usando Insertar\Nota al pie.

Además hay dos comandos adicionales que no aparecerán que son:

- \and utilizado para separa los nombres de los autores en el caso en el que haya mas de uno e instituciones.
- \inst {<nr>} utilizado para marcar los correspondientes pares Author/Institucion. Las instituciones están numeradas secuencialmente como aparecen en el campo Addres. Entonces, usted debe colocar un marcador para cada Author.

En ambos casos, los comandos deben estar incluidos en código látex.

El Abstract está inmediatamente después de la cabecera en donde usted puede comenzar a escribir su resumen del artículo.

Sigue entonces la introducción y el cuerpo del artículo, dentro del cual puede introducir cualquiera de los siguientes entornos:

- Standar
- Section
- Subsection
- Subsubsection
- Itemize
- Enumerate
- Description
- Caption
- Abstract
- Acknowledgment
- Bibliogray
- Latex

Cualquiera de estos elementos puede ser usado de la manera que hemos estado viendo. Seleccionándolos de la lista de entornos y escribiendo el contenido de cada uno. También podemos insertar Tablas y figuras flotantes usando los comandos del menú Insertar.

ANEXO B. INSTALACION Y CONFIGURACION DE HT://DIG

1. INSTALACION DE HT://DIG

Una vez que el paquete ha sido descomprimido y ubicado en un directorio se puede proceder a la configuración previa al proceso de compilación e instalación en su ubicación final.

Editar el archivo CONFIG y modificar los valores de las variables donde vayamos a ubicar los archivos luego de compilados y aquellos directorios donde se guardaran los datos indexados y los archivos de configuración.

El directorio que asumiremos como principal en la instalación, en nuestro caso, es /usr/local/htdig (en SuSE 7.0 la ruta de instalación es /opt/www/htdig) y teniendo esto en cuenta modificaremos las variables del archivo de configuración para reflejar estos cambios.

VARIABLE	DESCRIPCION
DEST = /usr/local/htdig (en SuSE 7.0 /opt/www/htdig)	Esta variable es el directorio principal donde se instalara ht://Dig y otras variables tomarán este valor como referencia.
BIN_DIR = \$(DEST)/bin	Todos los ejecutables de ht://Dig se almacenaran en este directorio.
CONFIG_DIR = \$(DEST)/conf	Aqui se ubicarán los archivos de configuración.
COMMON_DIR = \$(DEST)/common	Este directorio almacenará archivos compartidos por las bases de datos y sus procesos.
DATABASE_DIR = \$(DEST)/db	Aquí se almacenaran las bases de datos. Este directorio deberá ubicarse en una partición con suficiente espacio como para almacenar los datos de las bases.
DEFAULT_CONFIG_FILE = \$(CONFIG_DIR)/htdig.conf	Este archivo es el que todas las aplicaciones buscarán para su funcionamiento.
CGIBIN_DIR = /usr/local/httpd/cgi-bin	En este directorio el servidor WEB buscará los programas CGI. Aquí es donde el programa htsearch será ubicado luego del proceso de instalación.
IMAGE_DIR = /usr/local/httpd/icons	Definir este directorio como el lugar accesible desde el servidor WEB donde se almacenarán las imágenes utilizadas por ht://Dig. Estas imágenes luego podrán ser modificadas de acuerdo a nuestras necesidades y gusto.
IMAGE_URL_PREFIX = /	Este es el URL que apunta al directorio especificado por la variable IMAGE_DIR
SEARCH_DIR = /usr/local/httpd/htdocs SEARCH_FORM = search.html	El nombre del archivo que contendrá la página con el formulario de búsqueda.

Tabla 8. Variables del Archivo CONFIG

Luego de finalizar la modificación del archivo CONFIG se puede correr el programa %/configure que tratará de determinar el tipo de sistema operativo instalado y setear correctamente las variables y los Makefiles necesarios. Luego de correr este programa ya podemos compilar el paquete con...%make. En caso de haber modificado alguno de los fuentes del programa es necesario ejecutar un 'make depend' antes de correr el 'make' final.

Si todo ha ido bien, se pueden instalar los binarios y archivos de configuración con `%make install`. Esto realiza varias acciones. Crea los directorios necesarios que se indicaron en el archivo `CONFIG`, instala los binarios resultado del proceso de compilación.

`htdig`

`htmerge`

`htfuzzy`

`htnotify`

También copia el programa encargado de realizar las búsquedas (`htsearch`) en el directorio indicado en el archivo `CONFIG` (directorio `CGIBIN_DIR`).

Luego de este proceso varios archivos que se han instalado, pueden ser reconfigurados de acuerdo a nuestras necesidades. Algunos de los mas importantes:

- ***CONFIG_DIR/htdig.conf***. Un archivo de configuración con los elementos mínimos para una configuración y creación de una base de datos.
- ***SEARCH_FORM***. Una página con un formulario de búsqueda.

- *COMMON_DIR/footer.html*,
COMMON_DIR/header.html,
COMMON_DIR/nomatch.html,
COMMON_DIR/syntax.html.

Estas páginas se muestran como encabezado (*footer.html*), pie (*header.html*), mensaje de que no se ha encontrado ningún documento que contenga las claves indicadas (*nomatch.html*) y error de sintáxis en la indicación de claves de búsqueda que contengan operadores lógicos (*syntax.html*).

- *BIN_DIR/rundig*. Un programa que creará nuestra primera base de datos de acuerdo con los datos proporcionados en el archivo *CONFIG* antes del proceso de compilación.

2. CONFIGURACION DE NUESTRO SISTEMA

Solamente resta ahora modificar el archivo de configuración ubicado en *CONFIG_DIR/htdig.conf* (en nuestro caso, de acuerdo a los datos indicados antes de la compilación en el archivo *CONFIG* este se halla ubicado en */usr/local/htdig/conf/htdig.conf*).

Aquí algunos de los valores mas importantes a cambiar en este archivo. Los datos indicados son los necesarios para crear la base de datos de nuestro servidor principal y el servidor de Bases de Datos.

```
# Directorio principal donde se ubicaran las bases de
# datos y los archivos necesarios para los procesos de
# indexación y búsqueda.
#
database_dir: /usr/local/htdig/db

# El URL principal donde htdig comenzará los procesos
# de obtención de páginas. Se puede indicar mas de un
# URL separados por un espacio.
#
start_url: http://nuestrodominio/index.html

# Si hay páginas que se desean excluir de los procesos
# de indexado se pueden indicar en esta sección. Los
# valores que se indican aquí son cadenas de caracteres,
# si algún URL coincide con esta cadena no será indexado.
# También pueden utilizarse las posibilidades ofrecidas
# por los META-TAGS reconocidos por ht://Dig para exclusión
# de páginas.
exclude_urls: /cgi-bin/ .cgi
```

Una vez modificados estos valores ya es posible indexar nuestro sistema ejecutando `/usr/local/htdig/bin/rundig` (ese script de ejemplo creará las bases de datos necesarias y las almacenará en `/usr/local/htdig/db`).

Ya es posible cargar la página que contiene el formulario de búsqueda, realizar una búsqueda y ver el resultado.

ANEXO C. CONFIGURACION E INSTALACION DEL SERVIDOR WEB

APACHE

1. INSTALACION

Para realizar la instalación es necesario tener los paquetes del servidor web Apache.

[Http://www.apache.org/](http://www.apache.org/)

Si queremos más flexibilidad (relacionar con bases de datos, server-side includes ...):
apache-modules (si queremos modulos para ampliar: mod_counter, mod_sp), php (si
queremos el módulo de PHP/FI), libfcgi1 (si queremos soporte para FASTCGI's).

Se instala en el sistema (en SuSE con 'rpm'). Si tuviéramos una distribución RedHat
serían otros paquetes pero el proceso es el mismo.

Y ya está! La instalación es fácil y rápida (solo queda configurarlo, aunque hay
configuración por defecto y preparar el sistema para lanzarlo en el arranque – ya incluye
script para hacerlo)

Es posible coger también el código fuente original (apache.tar.gz) y compilarlo (si no nos
fiamos o queremos tocarlo), los paquetes para Linux lo traen precompilado.

2. CONFIGURACION DEL SERVIDOR

Los ficheros de configuración del servidor Apache se encuentran en `/etc/httpd/` . Hay tres ficheros importantes: `httpd.conf`, `srm.conf`, `access.conf` , de los cuales explicaremos su configuración por separado.

2.1. Configuración del servidor: `httpd.conf`

Describe como se comporta el demonio de `httpd`:

Forma de lanzarse: `inetd` o `standalone`

Puerto donde se lanza el servidor

Usuario y grupo al que pertenece (IMPORTANTE! NUNCA como `root`, diferencia de demonio y programa que atiende peticiones)

Donde están los ficheros de configuración (se le llama con `-f`) y donde se guardan los logs (IMPORTANTE para depurar)

Módulos que carga

Tiempos del servidor: de `time-out`, de petición (más de una por conexión)

Número de servidores a lanzar (los que atienden la petición)

Configuración de PROXY (servicio distinto a WWW no lo tratamos)

Host virtual

Es importante porque es la configuración del servidor completo

Modificaciones aquí cambian MUCHO como funciona así que hay que entenderlas (al menos un poco)

2.2 Configuración del servidor: srm.conf

Define los nombres (URLs) y como se traducirán las peticiones del cliente

DocumentRoot: donde está el raíz (/) del servidor

UserDir: traducción de http://maquina/ usuario (en general a public_html)

DirectoryIndex: los ficheros que se utilizan para index (mentalidad antigua de servidor de ficheros, aún persiste)

Iconos para generar índices 'al vuelo'

Ficheros que no indexa (IMPORTANTE! Agujero que da información al exterior)

Redirección de documentos que antes estaban

Acciones del servidor (añadidos) comportamiento frente a .cgi o tipos mime predefinidos.

(todo comentado al principio porque todos pueden ser posibles agujeros de seguridad)

Apache es muy versatil y permite hacer muchas cosas....

... pero hay que tener cuidado con que no haga cosas que no queremos (revisar bien todo antes de lanzarlo)

2.3 Configuración del servidor: access.conf

Define los servicios que ofrece y bajo qué circunstancias (algunos ya viene SIEMPRE):

Document-Root (tiene que coincidir con el de srm.conf): no deja Overrides (para que un usuario con .htaccess pueda modificar el comportamiento), no tiene opciones activadas de enlace simbólicos ni ejecución de CGI's en el espacio de usuario..

Cgi-bin, desactiva todas las opciones para que no haya agujeros

TODOS los directorios a los que el Apache tiene acceso puede (y debe) ser configurado aquí.

Sintaxis: Directory directorio + Opciones + acceso.

Conviene mirar la documentación para entender todas las opciones. Estas son algunas:

IncludesNoExec: desactiva server side includes

SymLinksIfOwnerMatch

Indexes

FollowSymlinks

AllowOverride

Includes

ExecCGI

allow

deny

1.8 Comportamiento del servidor

El servidor escucha en un puerto (o alguien escucha por él: inetd) y cuando le llega una petición hace un fork a un hijo (que atiende la solicitud) .

La atención de la solicitud puede ser: coger fichero del disco duro, preparar mensaje (server-includes), ejecutar programa (CGIs)... siempre guarda logs (útil para fallos o problemas de seguridad).

Si no tiene más trabajo que hacer muere (en Apache se puede limitar el número de peticiones que puede recibir un hijo).

El padre crea hijos indiscriminadamente (no mira si la petición es válida lo mira el hijo), con lo que conviene tener limitado el número máximo de hijos (para no cargar la máquina).

1.9 Comportamiento del cliente

El cliente (navegador) no es más que (ahora) una interfaz a un protocolo (HTTP: RFC2068 el 1.1, RFC1945 el 1.0).

Realiza la conexión al puerto 80 (aunque el servidor de www podría lanzarse en cualquier puerto) y con la directiva GET pide lo que necesita.

El servidor le responde y él se lo representa a el usuario (gráficamente o no) .