



# ESCUELA DE INGENIERÍA DE SISTEMAS Y COMPUTACIÓN

## BASES DE DATOS

### PROYECTO

Febrero-Junio 2011

## BIBLIOTECA DIGITAL

Las bibliotecas digitales ofrecen recursos de información que están disponibles en formato digital (pdf, doc, etc.) y a las cuales se accede mediante un computador. En la Escuela de Ingeniería de Sistemas y Computación se generan documentos digitales resultado del trabajo conjunto entre profesores y estudiantes. Cada semestre, los estudiantes de último año deben presentar un trabajo de grado o una tesis de maestría o doctorado, dependiendo del programa de formación al cual pertenece el estudiante. Por su parte, los profesores producen notas de clase, artículos y en general resultados de sus trabajo en la Escuela. Para poner a disposición de los estudiantes estos documentos, la Escuela ha pensado en diseñar un repositorio que sirva de base a una biblioteca digital, gestionada por ésta. Las tareas más importantes en la biblioteca digital son, la gestión de usuarios, la catalogación de documentos, las consultas de usuario y los reportes.

### 1) Gestión de usuarios

Existen dos **tipos de usuarios** en la biblioteca: registrados y no registrados. Los registrados pueden tener 3 tipos de perfil (administrador, catalogador, normal). Cada perfil tiene permisos diferentes respecto a las operaciones que pueden hacer en la biblioteca.

El **administrador** puede gestionar las operaciones y la información relacionada con los usuarios de la biblioteca. El administrador asigna los perfiles de los usuarios. Cada perfil define ciertos privilegios para el catalogador, usuario registrado (normal) y usuario no registrado.

Los **usuarios normales** pueden consultar la biblioteca y descargar documentos digitales. Los usuarios no registrados son invitados y solo pueden consultar la biblioteca pero no pueden descargar ningún documento.

El registro como usuario normal, incluye datos personales, áreas de interés y datos demográficos. El registro se hace diligenciando un formulario que contiene la siguiente información: login, contraseña, verificación de la contraseña, pregunta secreta, respuesta secreta y los siguientes datos personales: Nombres, apellidos, género, fecha de nacimiento, correo electrónico, nivel de escolaridad, vínculo con Univalle (estudiante de pregrado, estudiante de postgrado, egresado, profesor activo, jubilado, ninguno). Información adicional sobre áreas de interés en computación se almacena también. Esto con el propósito de notificarle al usuario cuando se catalogan nuevos documentos relacionados con alguna de sus áreas de interés. Un usuario puede registrar varias áreas de interés. Una vez registrado puede acceder a la biblioteca autenticándose directamente.

Los **catalogadores** son los responsables de catalogar los documentos y almacenarlos en el repositorio. Los **documentos digitales**, son catalogados por personas entrenadas con este fin y a las cuales se les ha asignado este perfil de CATALOGADOR. Todos los catalogadores están registrados en la biblioteca digital. Por último, un usuario **no registrado** puede entrar a la

biblioteca y buscar documentos pero no puede descargarlos.

## 2) Catalogación de documentos

La catalogación es el proceso de que permite administrar los documentos en la biblioteca digital. Este proceso consiste en registrar ciertos metadatos de los documentos que serán almacenados en la biblioteca dependiendo de su tipo. El *tipo de Material*, solo incluye dos campos: Nombre y Descripción del material. Los tipos pueden ser Libro, artículo, memoria, trabajo de grado, tesis de maestría, tesis de doctorado, material de clase y otros que puedan aparecer en el futuro.

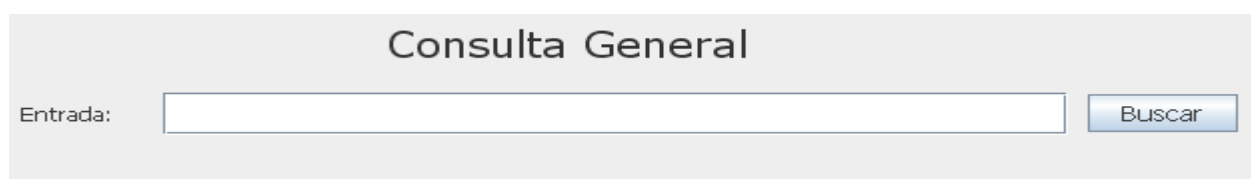
Los datos que se almacenan sobre los documentos son los siguientes:

- **Datos básicos**, que incluyen campos como: tipo de material, número de identificación, título principal, título secundario y/o Traducido, Editorial, Fecha Publicación, Fecha Catalogación, Fecha Creación, Idioma, Derechos de Autor y una breve descripción o resumen del material.
- **Autores**, incluyendo la información del *autor* del material, que debe estar registrado como autor. En caso de no existir el autor, se debe agregar como un nuevo autor a la base de datos, ingresando los siguientes campos: Nombre del Autor, Apellido, Correo y Acrónimo.
- **Palabras Clave**, que corresponde el material a catalogar. En caso de no existir la palabra clave, se puede agregar como una nueva *Palabra Clave* a la base de datos, ingresando los siguientes campos: Nombre y Descripción.
- **Áreas**, se debe seleccionar el *área de conocimiento* al cual corresponde el material a catalogar. Las áreas en Ciencias de la Computación aparecen en la Figura 1. Un área de conocimiento puede tener sub-áreas relacionadas.

Finalmente, se debe subir y almacenar el archivo correspondiente y los siguientes datos: Formato del archivo (.pdf, .doc, .jpg .odt), Tamaño del archivo en bytes, Resolución (píxeles) y el software recomendado para editar el material. La resolución y el software recomendado son datos opcionales.

## 3) Consultas de usuario

La biblioteca se diseña con el propósito de permitir que los usuarios consulten y descarguen documentos. Dos tipos de consultas se ofrecen: General y avanzada. En la **consulta general** el usuario solo necesita digitar una frase que contiene uno o varios términos de consulta. En este caso el usuario puede digitar cualquier metadato como el título, autor, palabra clave. El sistema debe ser capaz de buscar los documentos relacionados con esta entrada. Los resultados que se despliegan contienen los nombres de los documentos encontrados y el autor. En la figura 1 se muestra un bosquejo de la búsqueda general.



Consulta General

Entrada:

Buscar

Figura 1: Un bosquejo Consulta general

Para el caso de la **consulta avanzada** el autor especifica por cuáles campos desea buscar. El usuario puede buscar por título, autor, materia o área de conocimiento y/o la combinación de las anteriores. El usuario también puede especificar otros parámetros como la fecha de publicación,

el formato del archivo y el idioma. En la figura 2 se muestra un bosquejo de la consulta avanzada. **Esta imagen es solo un bosquejo y no significa que la interfaz de usuario tenga que ser así.**

El formulario se titula "Consulta Avanzada". Tiene seis campos de entrada: "Título", "Autor" y "Palabra Clave" son campos de texto largos; "Area", "Idioma" y "Formato archivo" son menús desplegables con "Todos" seleccionado; "Fecha Publicación" es un menú desplegable con "Cualquiera" seleccionado. A la derecha de los campos de texto hay tres botones con flecha hacia abajo que dicen "Con todas las palabras". En la parte inferior izquierda hay un botón "Buscar".

Figura 2: Un bosquejo de la consulta avanzada

Si el usuario busca por el título, también puede especificar las siguientes restricciones: "Con todas las palabras, con algunas de estas palabras, sin estas palabras". Lo mismo aplica para búsquedas por autor y palabras clave. El usuario puede especificar el área de conocimiento, por defecto se busca documentos de todas las áreas. En la fecha de publicación el usuario puede seleccionar los siguientes valores: "Cualquier fecha, En el último mes, Hace 6 meses, Hace un año, Hace 2 años, Hace 5 años". El usuario también puede especificar el idioma que puede ser: "Todos los idiomas, Español, Inglés, francés, alemán, etc". Lo mismo aplica al formato de los archivos. Después de especificar las restricciones anteriores, el usuario lanza la consulta y el sistema debe recuperar documentos teniendo en cuenta los parámetros especificados.

#### 4) Reportes y estadísticas

Cada vez que un usuario accede a la biblioteca, es necesario almacenar información de lo que descarga para poder generar estadísticas. Se almacena la fecha y la hora en la cual el usuario descargó el documento y la identificación del documento. Con base en esta información se pueden saber cuáles son los documentos que más se descargan, cuáles son las áreas en las cuales los usuarios muestran mayor interés, cuáles son las franjas horarias de más tráfico en la biblioteca. Algunas de las estadísticas que se requiere son las siguientes:

1. Documentos existentes por área en la biblioteca
2. Total de usuarios registrados por tipo
3. Total de consultas por día, área, tipo de usuario

El propósito de este proyecto, para el caso de la asignatura de Bases de Datos, es diseñar e implementar un sistema que use una base de datos de soporte a la gestión de la biblioteca digital. A nivel general, las funciones que el sistema de cumplir son:

1. Registrar, modificar, eliminar y consultar usuarios en los diferentes perfiles
2. Catalogar documentos en la biblioteca digital.
3. Registrar y modificar nuevas áreas de las Ciencias de la Computación
4. Registrar y modificar nuevas palabras clave y autores
5. Llevar registro de descargas de los documentos por usuario
6. Soportar la consulta general y avanzada
7. Generar, al menos los siguientes reportes en formato PDF
  - Número de documentos por área/autor/tipo de documento/formato
  - Documentos más descargados por mes

- Usuarios registrados por mes
- Documentos catalogados por área y por mes
- Número de documentos descargados por usuarios y área de conocimiento.

Recuerde que todas las consultas y reportes deben ser parametrizados, es decir, que debe permitir al usuario seleccionar diferentes opciones, y de acuerdo con estas generar los reportes o consultas.

### EVALUACION DEL PROYECTO

Entrega	Valor	Fecha Entrega
Modelo Entidad relación y modelo relacional de datos	15%	Marzo 18
Implementación de la estructura de la base de datos en SQL	10%	Abril 8
Implementación de la aplicación y sustentación  Gestión de usuarios y catalogación: 15% Consulta general: 15% Consulta avanzada: 15% Estadísticas y reportes: 15% Sustentación individual 15%  La sustentación se hace a nivel de grupo, sin embargo cada uno de los integrantes debe estar en capacidad de responder cualquier pregunta.	75%	Mayo 3 (Gestión de usuarios y catalogación)  Mayo 17 (Consulta general y avanzada)  Mayo 27 ( Reportes y sustentación final)

## ANEXOS

En la tabla 2 se muestra un listado de las áreas de conocimiento en Ciencias de la Computación. Las áreas pueden tener subáreas. Por ejemplo, el área **Construcción de Aplicaciones Web** es una sub-área de **Computación Centrada en la Red**.

Cada fila de la tabla 2 representa un área. Las filas tienen 4 campos separados por un “;”. El primer campo, es un identificador único del área, el segundo, es el nombre del área y tercero es una descripción ( sin información en la tabla). La relación entre áreas y subáreas se representa usando el cuarto campo de cada fila. Por ejemplo, el área **Construcción de Aplicaciones Web**, cuyo identificador de área es 2, tiene como valor 1 en el cuarto campo. Esto significa que es una sub-área de **Computación Centrada en la Red**.

1;	Computación centrada en la red";	"";	
2;	Construcción de Aplicaciones Web";	"";	1
3;	Métodos numéricos computacionales";	"";	
4;	Modelos de simulación";	"";	3
5;	Investigación de operaciones";	"";	3
6;	Ingeniería de software";	"";	
7;	Diseño de software";	"";	6
8;	Gestión de información";	"";	
9;	Hipermedia e hipertexto";	"";	8
10;	Modelo de sistemas de información";	"";	8
12;	Lenguajes de consulta de bases de datos";	"";	8
13;	Bibliotecas digitales";	"";	8
14;	Recuperación de información";	"";	8
15;	Redes de comunicación";	"";	1
16;	Diseño físico de bases de datos";	"";	8
18;	Lenguajes de programación";	"";	
19;	Programación orientada a objetos";	"";	18
20;	Sistemas inteligentes";	"";	
21;	Búsqueda por satisfacción de restricciones";	"";	20
25;	Entornos de desarrollo de software";	"";	6
26;	Diseño de bases de datos relacionales";	"";	8
27;	Validación de software";	"";	6
28;	Computación Web cliente servidor";	"";	1
29;	Tecnologías multimedia";	"";	1
30;	Computación inalámbrica";	"";	1
31;	Procesamiento de lenguaje natural";	"";	20
32;	Sistemas multimedia";	"";	8
33;	Fundamentos de programación";	"";	
34;	Programación orientada por eventos";	"";	33
35;	Robótica";	"";	20
36;	Algoritmos y complejidad";	"";	
37;	Algoritmos paralelos";	"";	36
38;	Interacción humano-computador";	"";	
39;	Aspectos de comunicación";	"";	38
40;	Desarrollo de software centrado en humano";	"";	38
41;	Diseño de interfaces gráficas de usuario";	"";	38
42;	Procesamiento de transacciones";	"";	8
43;	Minería de datos";	"";	8
44;	Computación visual y gráfica";	"";	
45;	Rendering avanzado";	"";	44
46;	Animación por computador";	"";	44
47;	Realidad virtual";	"";	44
48;	Arquitectura y organización";	"";	

49;"Arquitectura para redes en sistemas distribuidos";";48  
 50;"Análisis numérico";";3  
 51;"Modelamiento de datos";";8  
 52;"Bases de datos relacionales";";8  
 53;"Bases de datos distribuidas";";8  
 54;"Seguridad en redes";";1  
 55;"Bioinformática";";  
 56;"Estructuras discretas";";  
 57;"Grafos y árboles";";56  
 58;"Sistemas de bases de datos";";8  
 59;"Representación de conocimiento";";20  
 60;"Probabilidad discreta";";56  
 61;"Visualización";";44  
 62;"Sistemas operativos";";  
 63;"Sistemas en tiempo real";";62  
 64;"Tolerancia a fallas";";62  
 65;"Algoritmos criptográficos";";36  
 66;"Máquinas virtuales";";18  
 67;"Programación funcional";";18  
 68;"Aprendizaje - Redes neuronales";";20

*Tabla 2. Anexo, áreas de conocimiento*