

Bibliotecas digitales: Greenstone

10

Este capítulo se publica bajo licencia Creative Commons 2.5 Reconocimiento-NoComercial-SinObraDerivada 2.5 Spain



10.1. LAS BIBLIOTECAS DIGITALES

Greenstone es una aplicación pensada para la creación de bibliotecas digitales, accesibles a través de un navegador. Se puede utilizar tanto en entornos abiertos, como en entornos de intranet. Sin embargo, para comprender adecuadamente las capacidades de Greenstone, así como los contextos más adecuados de aplicación, resulta necesario establecer los elementos que definen a una biblioteca digital:

- Una biblioteca digital es un servicio de información cuyo contenido fundamental es una colección de documentos. La colección de documentos debe ofrecer un nivel de homogeneidad notable, y debe estar orientada a solucionar los problemas de información y documentación de los usuarios de la misma.
- La colección de documentos debe ser gestionada según unos principios rigurosos de selección, organización y clasificación de los documentos atendiendo a su contenido informativo.
- Los principales métodos de acceso a la información son las búsquedas a texto completo del contenido de los documentos, mediante un motor de búsqueda, y los mecanismos de navegación, basados en clasificaciones. Estos mecanismos deben crearse y desarrollarse teniendo en cuenta las características de los usuarios.

De los párrafos anteriores puede deducirse que no todo lo que se llama “biblioteca digital”, que se encuentra en Internet, es una biblioteca digital. En numerosas ocasiones se emplea, no muy correctamente, como sinónimo de portales verticales, de directorios temáticos o de bases de datos especializadas. Esto no quiere decir que no se pueda integrar una biblioteca digital en un portal, o en otro tipo de servicio de información digital. En cualquier caso, una biblioteca digital debe permitir que el usuario realice búsquedas dentro del contenido de los documentos que forman parte de la colección. Las búsquedas dentro de los documentos se llevan a cabo sobre el contenido textual de los mismos, lo cual

supone que es necesario disponer de una herramienta software que permita acceder al texto de los documentos de la colección, lo analice y lo procese, mediante técnicas de indexación, para crear índices que hagan posible la búsqueda por medio de términos y frases contenidos en los mismos.

Cabe plantear en qué entornos resulta interesante implantar un servicio de biblioteca digital. En primer lugar, los usuarios deben acceder a un conjunto numeroso de documentos, que contienen principalmente información textual. El contenido de los documentos resulta homogéneo, y éstos pueden clasificarse, usando diferentes criterios, en virtud de ese contenido. Los usuarios necesitan acceder a los documentos importantes para sus necesidades, bien mediante la formulación de ecuaciones de búsqueda, bien mediante la utilización de criterios de clasificación significativos, y necesitan un listado de respuestas que ofrezca una valoración de los resultados obtenidos, lo que se denomina relevancia. Por ejemplo, un servicio técnico con gran cantidad de guías y documentos en formato pdf, que necesita localizar aspectos específicos de los productos, o un servicio de formación que disponga de gran cantidad de manuales de aplicaciones en formatos de procesador de texto, o en HTML, o similares, y los quiere hacer disponibles a sus usuarios. También sería de aplicación en un entorno de intranet en el que los usuarios utilicen ampliamente documentos ofimáticos, que aportan a repositorios centralizados, y que deben luego recuperar atendiendo a su contenido, como informes, proyectos, hojas de cálculo, etc. Las herramientas comerciales que ofrecen las prestaciones necesarias tienen un precio muy elevado, y sus capacidades son similares, en gran número de casos, a las disponibles en un software libre como el que ocupa este capítulo.

10.2. GREENSTONE

Greenstone es un paquete de software que permite la creación y utilización de una biblioteca digital, con sus correspondientes colecciones de documentos. Se distribuye bajo licencia GNU. Su desarrollo lo lleva a cabo un equipo de investigadores de la *University of Waikato* (Nueva Zelanda), y es la plataforma que da soporte a un creciente número de bibliotecas digitales. Su fiabilidad y nivel de desarrollo y de mantenimiento ha llevado a la UNESCO a incluirlo en su programa de aplicaciones informáticas para servicios de información y documentación. Se puede obtener descargándolo de *Sourceforge*, o mediante un CD-ROM que distribuye la UNESCO. El web de referencia se encuentra en <http://www.greenstone.org>. La documentación de referencia es bastante completa, pero sólo cubre los aspectos generales del software (Figura 10.1).

La arquitectura de Greenstone puede parecer, a primera vista, complicada. En primer lugar, un motor de indexación y recuperación de información textual, llamado MG, que utiliza el modelo vectorial para el tratamiento de la información. La explicación del modelo vectorial sobrepasa el objetivo y planteamiento de este libro, y basta saber que es empleado por herramientas comerciales de alto coste. Este motor es alimentado por un conjunto de scripts en Perl, encargados de preprocesar los documentos, mediante los llamados plugins, pequeños programas que transforman los documentos originales en documentos XML, codificados en Unicode, lo que asegura la capacidad de la aplicación para

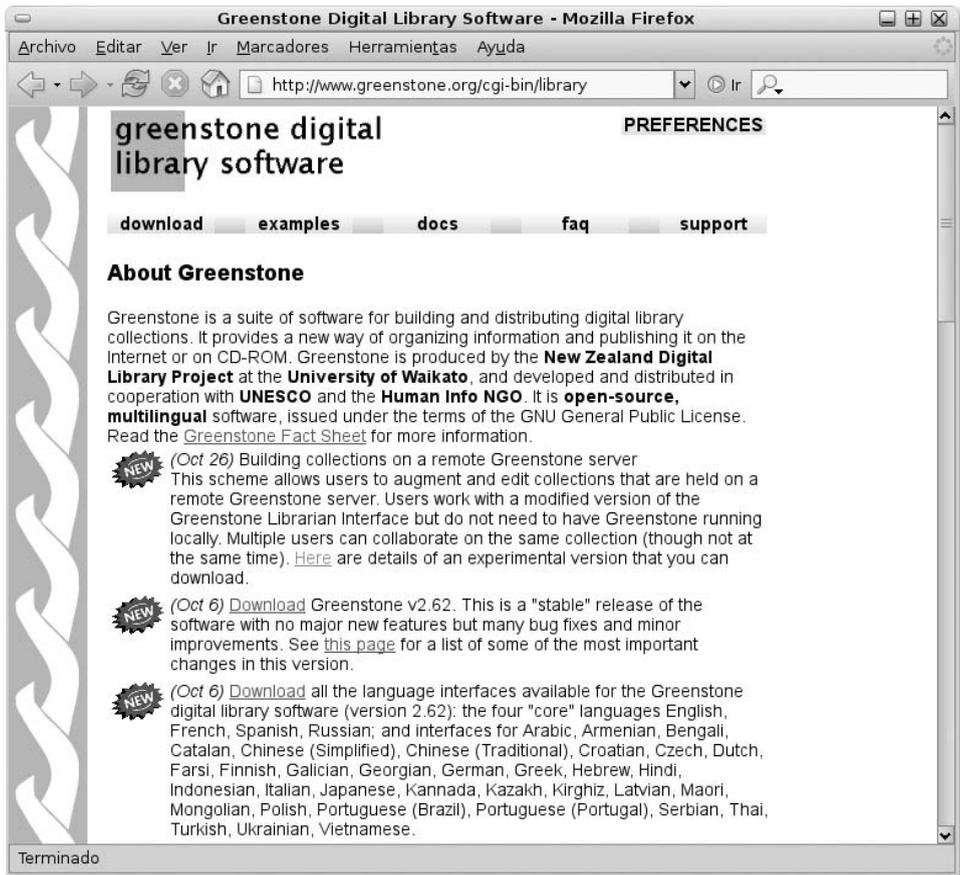


Figura 10.1. Web de Greenstone.

tratar cualquier idioma. La información es tratada y almacenada mediante el sistema de gestión de bases de datos GDBM (*GNU DataBase Manager*), que también es software libre. Además, el paquete incluye una versión especial de MG, el MG+/MGPP, para cuando sean necesarios formularios con múltiples campos para las interfaces de búsqueda. En la versión 2.61 se añadió, además, el motor *Lucene*, desarrollado por *Apache Foundation*. El proceso básico respondería a la selección de los documentos a tratar, su preprocesamiento y conversión en XML mediante los plugins de Perl adecuados, la extracción de metadatos de los documentos (asignados automática o manualmente), la indización con el motor elegido y el almacenamiento con GDBM. El proceso de recuperación usaría el motor correspondiente para localizar y acceder a los documentos relevantes. Para el usuario final, el proceso resulta transparente. Para el responsable de las colecciones, el proceso es más complejo, y ofrece múltiples posibilidades, relacionadas con el tratamiento de la información textual y el desarrollo de las clasificaciones y presentaciones de los resultados. Desde su versión 2.41, se ha incorporado una interfaz gráfica en Java, GLI (*Greenstone Librarian Interface*) que facilita el proceso de creación y administración de las colecciones.

10.3. REQUERIMIENTOS Y VERSIONES

En el momento de redactar este apartado la versión disponible de Greenstone es la 2.62, que se encuentra disponible ya compilada para plataformas Windows, Linux (RedHat/Fedora, Debian, Suse y Slackware), Sun Solaris y Macintosh OS/10. Para otras plataformas UNIX, es necesario obtener el código fuente y compilarlo, por ejemplo con el compilador GCC, también software libre. Se encuentra disponible la versión 3, de prueba, desarrollada en Java, y que necesita para su funcionamiento un servidor de aplicaciones como Tomcat. La versión experimental no se trata en este texto. Los requerimientos previos para una instalación de Greenstone son:

- Un servidor web (Apache, por supuesto, aunque funciona correctamente con PWS y con IIS de Microsoft, y se ha documentado su instalación con otros servidores libres).
- Un intérprete de Perl versión 5 o superior (Greenstone recomienda ActiveState Perl, pero funciona con la distribución estándar de Perl).
- Java Runtime Environment, JRE, de Sun, en su versión 1.4.2 como mínimo (para poder usar GJI, casi imprescindible).
- ImageMagick, un paquete de software libre para el procesamiento de ficheros gráficos, en el caso de que se trabaje con colecciones de imágenes.

GBDM viene integrado en las versiones para Windows y Macintosh, ya que se encuentra preinstalado en las diferentes distribuciones de Linux. En lo que respecta a prestaciones, si bien son mejores en una máquina equipada con Linux que en una plataforma Windows, la facilidad de instalación en esta última ha favorecido la instalación de servidores que combinan el software propietario con las herramientas de software libre.

Tabla 10.1. Recursos web sobre Greenstone

Web de referencia	http://www.greenstone.org
Descarga	http://www.greenstone.org/cgi-bin/library?e=p-en-home-utfZz-8&a=p&p=download
Documentación	http://www.greenstone.org/cgi-bin/library?e=p-en-download-utfZz-8&a=p&p=docs
Recursos: comunidades de usuarios	http://greenstone.docunautica.com/ http://www.wrlc.org/dcpc/UserInterface/interface.htm

10.4. PROCESO DE INSTALACIÓN

Una vez obtenida la versión correspondiente a la plataforma sobre la que se desea instalar, el procedimiento de instalación requiere lanzar un ejecutable (en Windows), o un script del shell (para Linux/Unix). Sin embargo, es necesario hacer una puntualización sobre la cuenta de usuario. Al tratarse de un servicio que se ejecuta a petición del servidor

web, la instalación debe hacerse usando una cuenta con derechos para modificar determinados ficheros de configuración y permisos de ejecución. Esto supone que, en plataformas Linux, la instalación debe hacerse desde una cuenta de administrador (*root*), y que en plataformas Windows debe ser el administrador, o un usuario del grupo de administradores. Aunque es posible instalar Greenstone no siendo administrador, luego comenzarán a aparecer problemas de permisos (por ejemplo, en la modificación del fichero `httpd.conf`), así que es preferible montarlo correctamente desde su inicio. El proceso de instalación sigue los siguientes pasos:

1. En Linux, descomprima el fichero `gsdl-2.62-unix.tar.gz`, y ejecute el script `Install.sh`. En Windows, doble pulsación sobre el fichero `gsdl-2.62-win32` (el número indica la versión de Greenstone). Hay diferencias entre el instalador de Linux (textual) y el de Windows (gráfico), aunque hay que responder a cuestiones similares.
2. Se pone en marcha el proceso de instalación, que en primer lugar le pedirá el idioma que va a usar. La versión estándar incorpora inglés, francés, español y ruso. Se le informará de la licencia GNU, y se le pedirá que la acepte.
3. Posteriormente, se le pide que indique el directorio en el que se instalarán Greenstone y las colecciones de documentos que vaya creando (viene predefinido `C:\Program Files\gsdl\` para Windows, y `/usr/local/gsdll` para GNU/Linux). No es necesario situarlo dentro de los directorios del servidor web, ya que luego se configurará mediante las directivas `Alias` y `ScriptAlias` (ver más adelante). Puede variar el directorio según le interese (Figura 10.2).



Figura 10.2. Instalación de Greenstone: definición de directorio.

4. Una vez indicado el directorio, se le ofrecerán las cuatro opciones de instalación (Figura 10.3). Seleccione aquellas que necesite, recordando que, si desea montar un verdadero servidor, deberá marcar la Biblioteca Web:

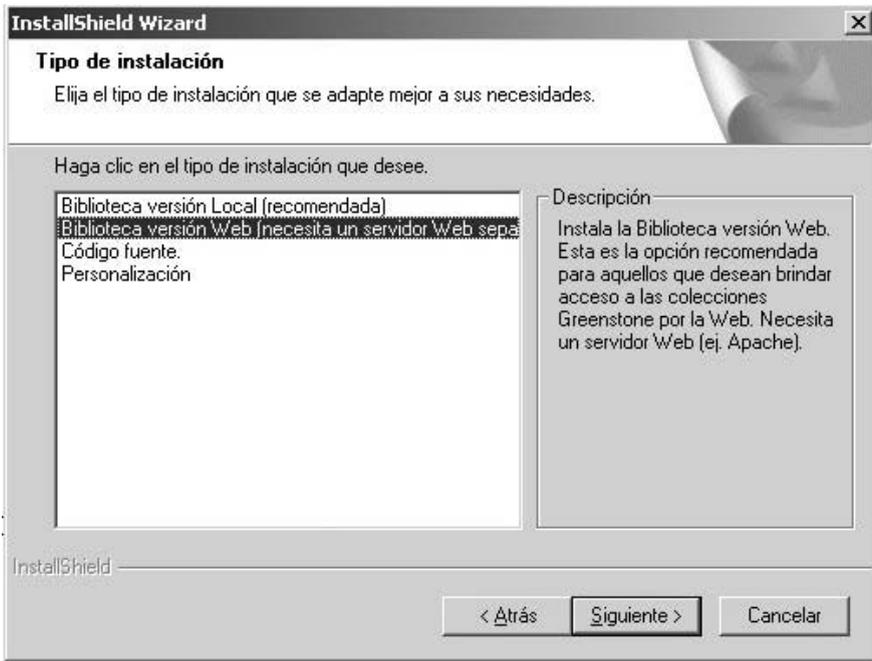


Figura 10.3. Instalación de Greenstone: opciones de instalación.

- a) Biblioteca local: corresponde a una instalación para un usuario local (aunque también es accesible por red), en la que Greenstone instala un servidor web mínimo para acceder a las colecciones. Sólo interesa para experimentar en la creación y configuración de colecciones, no sirve para dar un servicio real en el web. Esta opción sólo está disponible en las versiones para plataformas Windows.
 - b) Biblioteca web: ésta es la versión adecuada para dar un servicio fiable a través de un servidor web, como Apache. Es altamente recomendable que el servidor web esté instalado previamente.
 - c) Código fuente: como su nombre indica, si necesita compilar Greenstone para una máquina específica, ésta es la opción adecuada. Recuerde que necesitará un compilador, y luego terminar el proceso de configuración del servidor web que se trate.
 - d) Personalización: permite seleccionar los componentes a instalar (por ejemplo, la Biblioteca y el Código fuente, etc.).
5. La instalación le pedirá ahora una clave de administración para Greenstone, que será necesaria para acceder luego a opciones de colecciones, usuarios, etc. Viene

predefinida como “admin”, así que cámbiela inmediatamente, y recuérdela para evitar problemas posteriores. Tras ello, Greenstone comenzará a crear los directorios, registrar los componentes y copiar los ficheros necesarios para su funcionamiento.

- Una vez terminado, la ventana de instalación le informará de los cambios que debe introducir en el fichero de configuración de Apache para que se pueda acceder correctamente a Greenstone (Figura 10.4). No hace falta que los copie urgentemente, ya que el proceso los ha copiado en el fichero de texto `server.txt`, en el directorio raíz de Greenstone. En caso de usar otro servidor web diferente, consulte la *Guía de Instalación de Greenstone*. Tras ello, la instalación básica termina, y es necesario continuar con varias tareas de configuración. Aunque ya podría poner en marcha Greenstone, primero debe ajustar algunos parámetros para que funcione correctamente.



Figura 10.4. Modificaciones a efectuar en la configuración del servidor web.

- En primer lugar, localice el fichero `server.txt`, mencionado en el párrafo anterior. Dentro del mismo encontrará las siguientes líneas:

```
ScriptAlias /gsdl/cgi-bin "C:/gsdl/cgi-bin/"
  <Directory "C:/gsdl/cgi-bin">
    Options None
    AllowOverride None
  </Directory>
```

```
Alias /gsdl "C:/gsdl/"
<Directory "C:/gsdl">
  Options Indexes MultiViews FollowSymLinks
  AllowOverride None
  Order allow,deny
  Allow from all
</Directory>
```

8. Cópialas en el Portapapeles. Después, localice el fichero httpd.conf, que contiene la configuración de Apache, y que se encuentra en el directorio /conf/, dentro del directorio raíz de Apache. Edítelo (es un fichero de texto), y busque la sección que comienza por la etiqueta **<IfModule mod_alias.c>**. Al final de la misma, añada el texto, y guarde los cambios. Reinicie el servidor Apache, y ya podrá acceder a su Greenstone usando el URL `http://nombre_del_servidor/gsdl/cgi-bin/library`. Si la instalación ha sido correcta, podrá ver la página principal (Figura 10.5). Sin embargo, este URL puede dar muchas pistas a terceros... así que lo mejor será definir un alias que simplifique el URL a escribir por los usuarios, y que limite el acceso al contenido de los directorios de Greenstone.

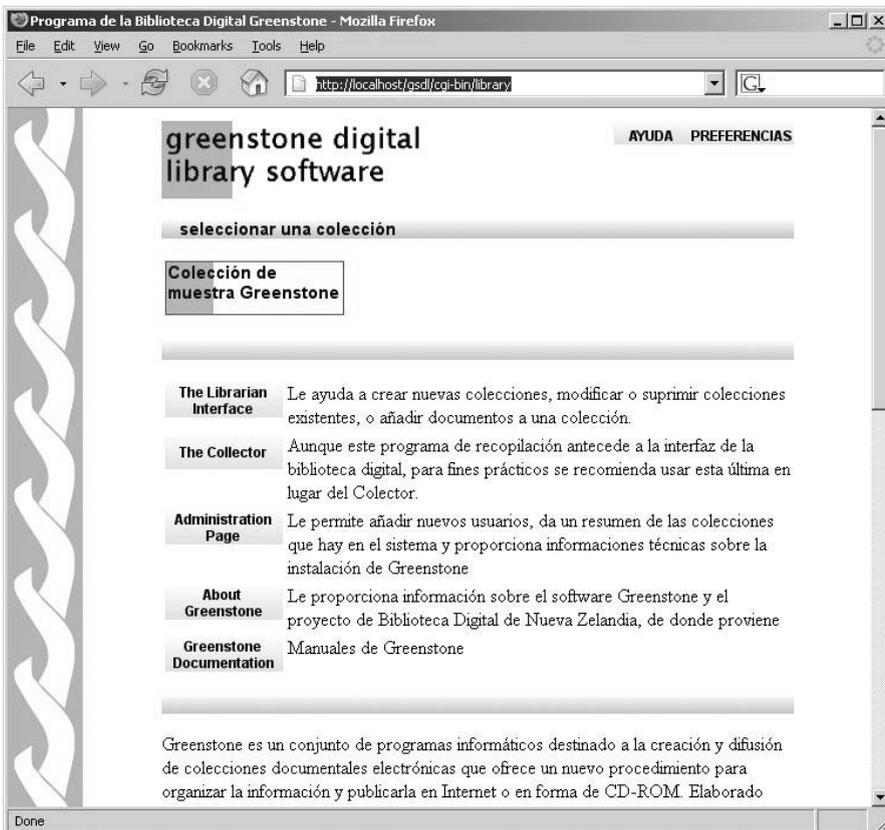


Figura 10.5. Página principal de Greenstone.

9. En primer lugar, vuelva a editar el fichero `httpd.conf`. Vaya de nuevo a la sección `<IfModule mod_alias.c>`, localice las líneas de las secciones `ScriptAlias` y `Alias`, que se han indicado en los puntos anteriores, y sustitúyalas por las siguientes:

```
ScriptAlias /gsdl/cgi-bin "C:/gsdl/cgi-bin/"
  <Directory "C:/gsdl/cgi-bin">
    Options None
    Order allow,deny
    Allow from all
  </Directory>

Alias /gsdl/ "C:/gsdl/"
  <Directory "C:/gsdl/">
    Order allow,deny
    Deny from all
    <FilesMatch "\.(gif|jpe|jpg|png|css|mov|mpeg|ps|pdf|doc|rtf|jar|class)$">
      Order allow,deny
      Allow from all
    </FilesMatch>
  </Directory>
```

10. En segundo lugar, como el URL es largo y parece complicado, se va a definir una redirección sencilla. Por ejemplo, el usuario podrá acceder a Greenstone usando la dirección `http://nombre_del_servidor/greenstone`. Esto se hace añadiendo un parámetro `Redirect` en el mismo fichero `httpd.conf`. Edite el fichero, y localice la sección **Redirect**. Esta sección está situada justo después de la sección de **ScriptAlias** y **Alias**. Añada la línea **Redirect /greenstone http://nombre_del_servidor/gsdl/cgi-bin/library**. Reinicie Apache, y podrá comprobar cómo el servidor permite usar el URL `http://nombre_del_servidor/greenstone`, y cómo los usuarios no pueden acceder al contenido de los directorios. Con estas modificaciones ha añadido un nivel básico de seguridad a su instalación de Greenstone.

10.5. ACCESO A LAS COLECCIONES

Una vez instalado y configurado, es el momento de revisar brevemente los procesos y técnicas de localización y acceso a la información que deben seguir los usuarios, antes de estudiar el procedimiento para crear colecciones. Vuelva a la página principal de su servidor (Figura 10.5). En la parte superior de la página encontrará la opción **Seleccionar una colección**. Al ser una instalación nueva, la única colección disponible es la *Colección de muestra de Greenstone*. Conforme vaya creando colecciones, aparecerán en el mismo espacio de la página principal. Debajo de este espacio, se muestra un conjunto de botones que le ofrecen varias opciones. En realidad, de éstos sólo interesa **Administration Page**, que da acceso a información sobre la versión, ficheros de configuración, colecciones y usuarios de Greenstone, y para acceder a varias de ellas volverá a necesitar la clave de administración.

Para consultar una colección, en este caso la Colección de muestra, pulse sobre el icono o enlace a la misma. Se mostrará entonces la página principal de la colección (Figura 10.6), que recibe el nombre **Acerca de**. Las páginas principales de la colecciones son similares en

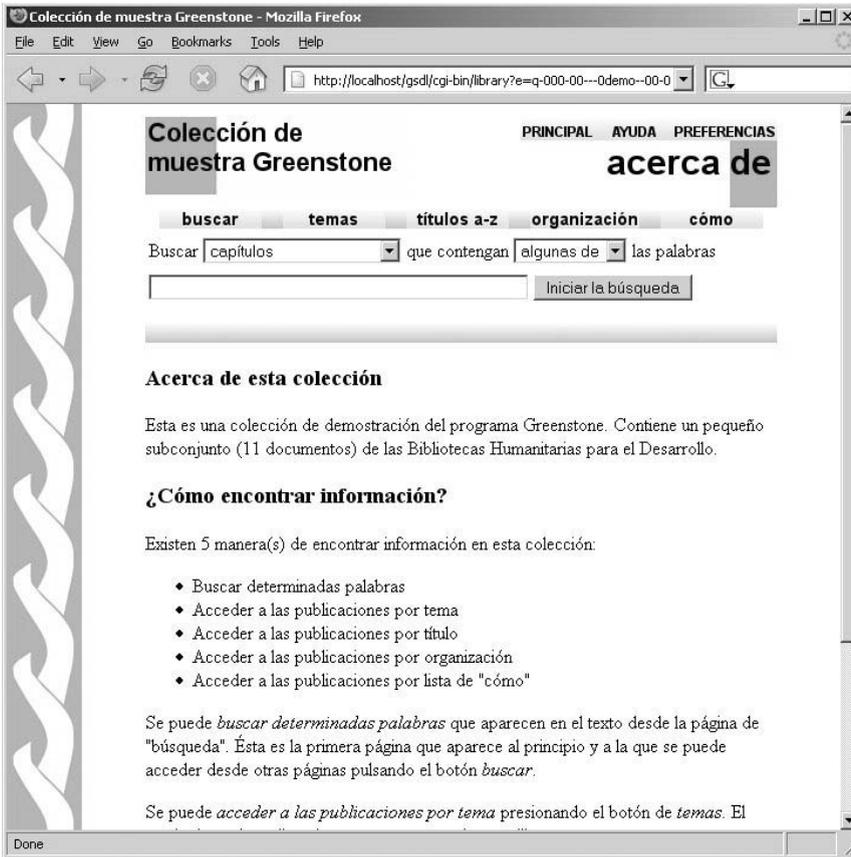


Figura 10.6. Acerca de la página principal de cada colección.

su versión básica. En primer lugar, muestran el icono de la colección, en la esquina superior izquierda. En la parte superior derecha, **Principal** le llevará a la página principal de su instalación de Greenstone, **Ayuda** le ofrecerá indicaciones de cómo buscar a texto completo o mediante los clasificadores, y **Preferencias** permitirá al usuario fijar características como la formulación de ecuaciones de búsqueda complejas, o el histórico de búsquedas.

Debajo se encuentra la barra de menú. El primero de ellos corresponde a la opción **Buscar**, y utiliza los cuadros de diálogo situados inmediatamente debajo para buscar cualquier palabra o palabras introducidas por el usuario, en el texto de todos los documentos incluidos en la colección. Las otras opciones reciben el nombre de clasificadores, son definidas durante el proceso de diseño y creación de la colección (que se verá más adelante), y pueden variar en cada colección. Los clasificadores muestran los documentos existentes en la colección, clasificados y ordenados según unos criterios preestablecidos, y se usan para buscar mediante navegación. Por último, el texto de ayuda que ocupa el resto de la página es generado automáticamente por Greenstone.

La interfaz de búsqueda ofrece unas opciones (capítulos, títulos, documentos, algunas palabras, todas las palabras) que se definen en el proceso de diseño y creación de la co-

lección. Tras formular una ecuación de búsqueda, el resultado es un listado de documentos que cumplen las condiciones establecidas (Figura 10.7), ordenados por relevancia (es decir, los más pertinentes aparecen en las primeras posiciones, al igual que en los motores de búsqueda en Internet). La pulsación sobre los resultados lleva al documento original, con su texto e imágenes, si las hay. El formato del listado de respuestas, así como el aspecto de la presentación del documento original y de su contenido, también pueden establecerse durante el proceso de creación de la colección.

The screenshot shows the Greenstone search interface. At the top left is the title 'Colección de muestra Greenstone'. To the right are links for 'PRINCIPAL', 'AYUDA', and 'PREFERENCIAS'. A large 'buscar' button is on the right. Below this is a navigation bar with tabs: 'buscar', 'temas', 'títulos a-z', 'organización', and 'cómo'. The search bar contains the text 'Buscar documentos enteros que contengan algunas de las palabras' with a dropdown menu set to 'documentos enteros'. Below the search bar is a text input field containing 'rice' and a button labeled 'Iniciar la búsqueda'. A section titled 'resultados' follows, showing 'Palabras contadas: rice: 75' and '6 documentos corresponden a la consulta.' Below this is a list of six search results, each with a magnifying glass icon and a link to the document.

Palabras contadas: rice: 75
6 documentos corresponden a la consulta.

- [The Courier - N°159 - Sept- Oct 1996 Dossier Investing in People Country Reports: Mali ; Western Samoa](#)
- [The Water Buffalo: New Prospects For An Underutilized Animal \(1984\)](#)
- [Microlivestock - Little-Known Small Animals with a Promising Economic Future](#)
- [Farming snails 1: Learning about snails; Building a pen; Food and shelter plants](#)
- [The Courier - N°158 - July - August 1996 Dossier Communication and the media - Country report Cape Verde](#)
- [Managing Tropical Animal Resources - Crocodiles as a Resource for the Tropics](#)

Figura 10.7. Listado de respuestas.

La función de los clasificadores para buscar la información es diferente. Ofrecen una ordenación y una clasificación de todos los documentos, según unos criterios establecidos al crear la colección. Por ejemplo, puede fijarse un clasificador por autores, lo que daría como resultado un listado de todos los documentos, ordenados alfabéticamente por autor, o por títulos. Observe el clasificador Temas. La pulsación sobre este clasificador permite mostrar una clasificación de tipo jerárquico, con temas o categorías anidadas. De esta forma se pueden crear clasificaciones complejas, en las cuales situar a los documentos, y que permiten ajustar la organización de los mismos a las categorías y ordenaciones a las que estén acostumbrados los usuarios. Como el creador de las colecciones puede establecer las clasificaciones que quiera utilizar para organizar sus documentos, puede optar por utilizar clasificaciones ya existentes, o tesauros, o ajustar taxonomías específicas que se ajusten a las necesidades de información de los usuarios. El potente conjunto que forman el

motor de búsqueda textual y la combinación de clasificadores convierten a Greenstone en una herramienta adecuada para organizaciones que deban ofrecer a sus usuarios organización, búsqueda y acceso a información contenida en documentos digitales (Figura 10.8).

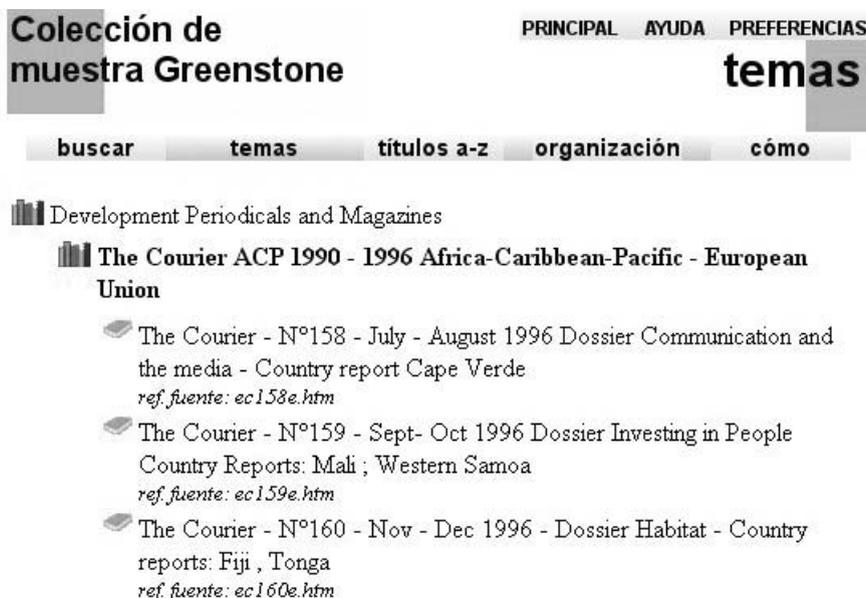


Figura 10.8. El clasificador jerárquico Temas de la Colección de muestra.

10.6. INTERFACES DE CREACIÓN DE COLECCIONES

Una vez instalado Greenstone, y revisados los principios básicos de consulta y uso, es posible comenzar con el proceso de creación y configuración de las colecciones. Greenstone ofrece para ello tres vías diferentes:

1. La interfaz *Collector*, a la que se accedía mediante un cliente web, desde la página principal del servidor Greenstone. Ha sido desactivada a partir de la versión 2.61. Si bien permitía el control de usuarios y la creación de colecciones directamente en el servidor, era necesario conocer con todo detalle las opciones de configuración disponibles, lo que limitaba su utilidad.
2. La interfaz de línea de órdenes o comandos, que debe ejecutarse desde una sesión de terminal (Linux shell, DOS, Telnet o similar), y que exige conocer la estructura de directorios de Greentone. Además, debe tener una cuenta de usuario con permisos de ejecución de ficheros en el servidor. Se trata de dar una secuencia de órdenes que ejecuta varios scripts en Perl, y también es necesario preparar externamente los ficheros de configuración y de metadatos correspondientes a la colección. Estos requerimientos limitan su utilización a administradores expertos en las capacidades de la herramienta.

3. La interfaz GLI, *Greenstone Librarian Interface* (o Interfaz de Bibliotecario de Greenstone), se trata de una aplicación en Java, de entorno gráfico, que se incorporó en la versión 2.41. Desarrolla con detalle las opciones de configuración, y facilita el trabajo de configuración de la colección. Se ejecuta en modo local, aunque en la última revisión de Greenstone se ha añadido un applet de Java que realiza las mismas funciones a través del servidor web, lo que permite el trabajo remoto. En este libro, todas las colecciones que se expliquen van a ser creadas y modificadas utilizando GLI.

En la interfaz que sea utilizada para la configuración y creación de colecciones, la secuencia de acciones que debe seguirse corresponde a las siguientes fases:

1. Creación de la colección: definición del nombre de la colección, descripción del contenido de la misma y responsable de su gestión.
2. Selección de los documentos: especificación de los documentos que se van a incluir en la colección y de la ubicación de los mismos.
3. Configuración de la colección: parámetros básicos de tratamiento, indización y presentación de la colección. Definición de la interfaz de usuario. Greenstone ofrece una configuración estándar, que puede ser modificada.
4. Construcción de la colección: de acuerdo con los parámetros establecidos en la fase anterior, los documentos son importados, procesados por los plugins correspondientes, transformados en XML e indizados. Se crean los índices textuales y los vectores, y las estructuras de organización y acceso, llamadas clasificadores.
5. Consulta de la colección: procesos de localización y acceso a los documentos, utilizando la interfaz disponible, mediante búsqueda en índices textuales o utilización de los clasificadores.

10.7. CREACIÓN DE UNA COLECCIÓN DE DOCUMENTOS

La primera colección que va a crear será una básica que recoja documentos ofimáticos en formatos Word y PDF, usando la configuración estándar que ofrece Greenstone. Una vez funcione correctamente, en una segunda fase se mejorará la configuración, para ajustarla a las necesidades de los usuarios de la misma. En primer lugar, es necesario disponer de un conjunto de documentos en estos formatos, y se pueden utilizar algunos previamente disponibles. En segundo lugar, hay que comprobar que ha sido instalado GLI, y que tiene privilegios de ejecución para ponerlo en marcha. Para ello, busque en los menús de aplicaciones de su entorno gráfico el correspondiente a Greenstone, y dentro del mismo, el icono correspondiente a *Greenstone Librarian Interface/Interfaz de Bibliotecario de Greenstone*. En entornos Windows lo encontrará en el menú; en entornos Linux/Unix, hay que ejecutar el script `gli.sh`, situado dentro del directorio `usr/local/gsd/gli`, con la orden `./gli.sh`. Para su ejecución es necesario que el sistema tenga instalado el *Java Runtime Environment*.

Una vez en marcha GLI, revise las características y los componentes de la interfaz que ofrece. En primer lugar, en la parte superior muestra una barra con los menús correspondientes a **Archivo** y **Editar**. En el espacio de trabajo de la ventana, muestra cinco pes-

tañas, **Bajar**, **Reunir**, **Enriquecer**, **Diseño** y **Crear**, que corresponden a las fases de creación de la colección, y que activarán sus opciones conforme vaya avanzando el proceso. En cada una de las pestañas, el espacio de trabajo inferior se divide en dos paneles, izquierdo y derecho, que mostrarán diferentes menús y opciones de trabajo y configuración. Para comenzar, y tener pleno acceso a toda la funcionalidad de GLI, es necesario abrir el menú **Archivo**, opción **Preferencias**. En la pestaña **Modo**, cambie la selección a **Avanzado**, ya que de esta forma, y como le indica el texto de ayuda inferior, tendrá acceso a todas las funciones (Figura 10.9).

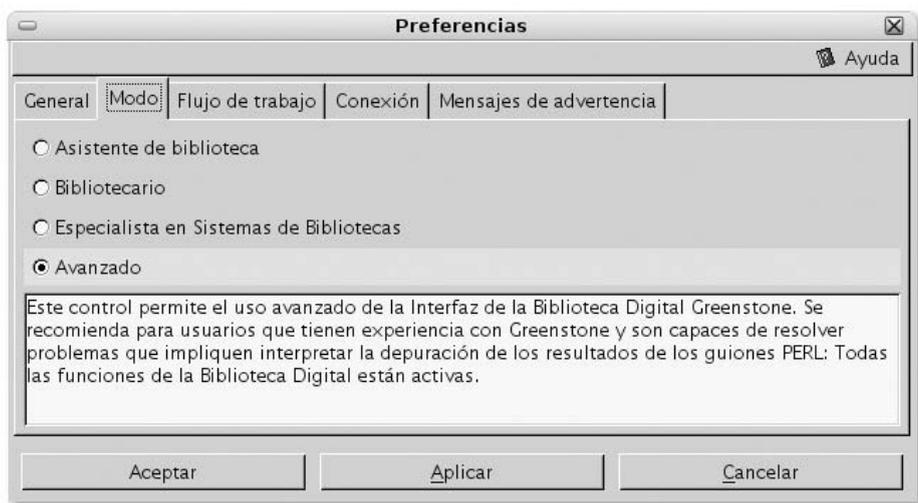


Figura 10.9. Configuración del Modo Avanzado.

El resto de opciones de esta ventana permiten definir el correo electrónico del creador de la colección y el idioma de la interfaz, el modo de conexión al servidor Greenstone (en principio, no será necesario cambiar nada), y los mensajes de advertencia que GLI mostrará al usuario. Una vez cambiado el modo, puede empezar la creación de la colección:

1. Menú **Archivo**, opción **Nueva**. Se muestra la ventana **Crear una nueva colección**. En los campos correspondientes introduzca el título de la misma, así como un breve resumen del contenido, que aparecerá en la página de inicio de la colección. En el parámetro **Construir esta colección con base en**, seleccione **Nueva colección**. Pulse sobre el botón **Aceptar** (Figura 10.10).
2. GLI pregunta qué conjunto de metadatos se va a asignar a la colección, y ofrece dos opciones. Por ahora, no se va a seleccionar ninguno de ellos. Pulse sobre el botón **Aceptar** (Figura 10.11). Pulse de nuevo **Aceptar** sobre el aviso que aparece ahora. Tras ello, GLI lanza los procesos necesarios para crear la estructura de la colección, y queda a la espera de que se le indiquen los documentos que debe procesar.

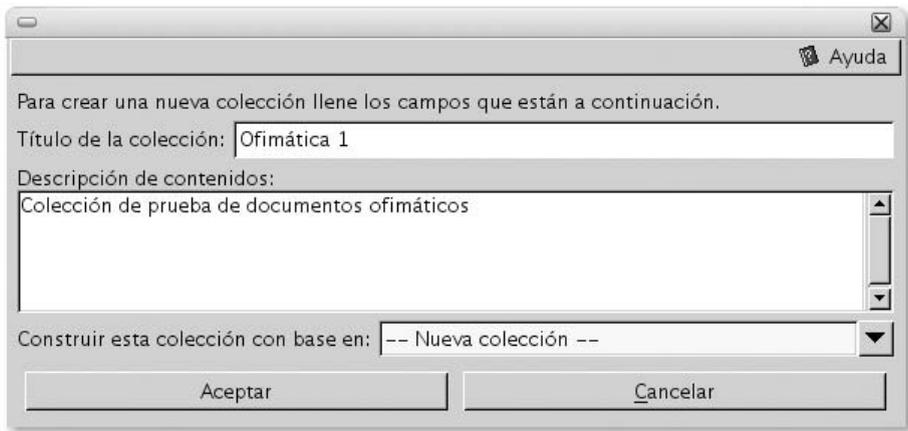


Figura 10.10. Diálogo de creación de colección.

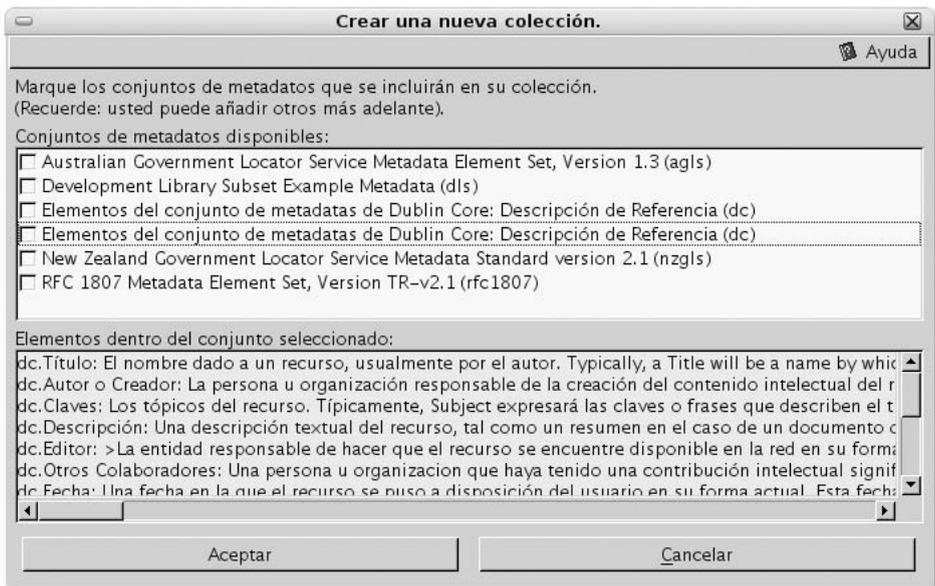


Figura 10.11. Definición de metadatos.

3. En la interfaz general de GLI, ahora todas las pestañas están activas. En primer lugar, hay que indicar a GLI qué documentos se van a incluir en la colección. En el espacio de trabajo bajo la pestaña **Reunir**, se muestran dos paneles de trabajo. El panel izquierdo permite explorar y localizar los documentos en el espacio de archivos local, en sus unidades, directorios y subdirectorios. Una vez haya localizado los documentos que quiere incluir en la colección, lo que tiene que hacer es pulsar sobre los mismos, y arrastrarlos al panel derecho. Greenstone hace una co-

pia de los documentos originales, y la incluye en sus directorios de trabajo. En este ejemplo, se han incluido ficheros word, pdf, txt, rtf y excel (Figura 10.12). Una vez copiados los documentos que se quieren incluir en la colección, pulse sobre la pestaña **Enriquecer**.



Figura 10.12. Selección de documentos para la colección.

4. Si se hubiesen definido metadatos, éste sería el momento de introducirlos. Como no ha sido así, las opciones para ello no están activas. Los metadatos se añadirán en la modificación posterior. Pulse sobre la pestaña **Diseño**.
5. La pestaña **Diseño** muestra un panel izquierdo que permitirá acceder a las opciones y parámetros de configuración que controlan los plugins a usar en cada caso, los índices, tipos de búsqueda, clasificadores, formatos. En el panel derecho se mostrarán las opciones y parámetros para cada una de ellas. En este momento, sólo van a interesar modificar la configuración estándar en lo referente a los plugins. Para ello, pulse sobre la opción **Plugins de documento**. Se muestran en el panel izquierdo los plugins predefinidos. Ésta es la configuración predefinida de Greenstone, que deberá modificar cuando desarrolle colecciones más complejas. Para ajustar la configuración necesaria para los tipos de documentos que se han seleccionado en el ejemplo, pulse sobre **plugin ZIPPlug**, y luego sobre el botón **Eliminar plugin**, de la parte inferior del panel derecho. Repita el proceso con **EMAILPlug** y **PSPlug**. Ahora, en el área de **Controles de edición**, en el menú

desplegable de plugins, seleccione **ExcelPlug**, y pulse sobre el botón **Añadir plugin**. Se abre una ventana de diálogo para configurar las opciones del plugin, en la que, por ahora, sólo debe pulsar sobre el botón **Aceptar**. La configuración final de plugins se muestra en la Figura 10.13.



Figura 10.13. Panel de configuración de plugins.

6. Por ahora, no se van a modificar otras opciones ni parámetros de configuración. Pulse sobre la pestaña **Crear**, y luego sobre el botón **Crear colección**, situado en la parte inferior de la pantalla. El desarrollo del proceso de importación y tratamiento de documentos se explica detalladamente en el fichero de eventos (log), que se va mostrando (Figura 10.14). En la parte superior de la ventana, dos indicadores de proceso informan sobre el mismo. Tras la finalización del proceso, se muestra con un diálogo el fin del procedimiento. No hay que preocuparse ahora si la revisión muestra que algún documento ha sido rechazado, o ha dado algún problema.
7. Ahora utilice el navegador para acceder a su Greenstone, y consulte la colección formada. Para ello, use el enlace textual que se muestra con el nombre de la base

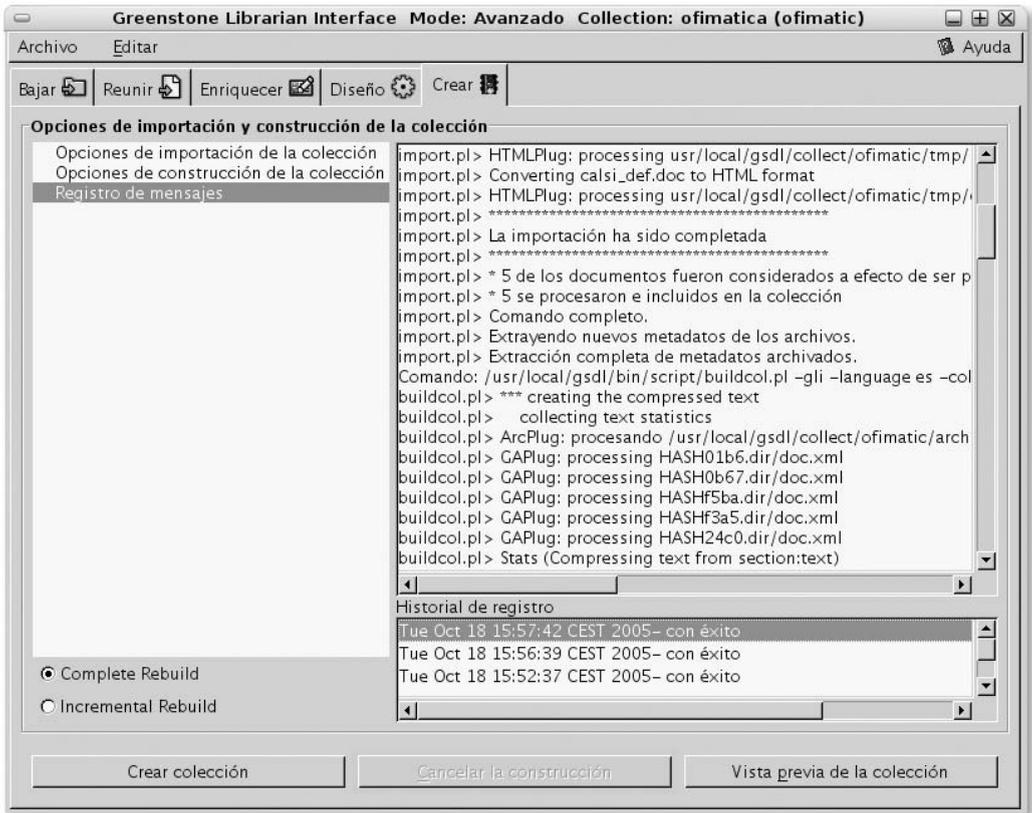


Figura 10.14. Información sobre el proceso de creación de la colección.

de datos. De esta forma se accede a la página principal de la colección. Observe que Greenstone ha creado directamente un índice de texto, que incluye todo el contenido de los documentos, y que ofrece un campo de consulta. Además, ha creado dos clasificadores de manera automática, a través de los cuales puede acceder directamente a los documentos (Figura 10.15). Observe que puede acceder, usando los iconos correspondientes, tanto al documento en el formato original, como al texto y contenido que ha extraído Greenstone de cada uno de ellos.

Sin embargo, observe que han aparecido problemas. En nuestro ejemplo, dos documentos han sido rechazados, y los clasificadores no aportan valor añadido al usuario final. Por otra parte, aparecen caracteres extraños en la páginas de la colección, en aquellos que deberían aparecer acentuados. A primera vista, el principal valor que aporta la herramienta es la creación de índice de texto, que permite disponer de un motor de búsqueda local muy potente y gratuito. Sin embargo, es posible aumentar las prestaciones y solucionar estos problemas con una adecuada configuración de los parámetros de Greenstone, a través del GLI. La mejora de la colección, aplicando parámetros de configuración más adecuados, será tratada en el siguiente apartado.

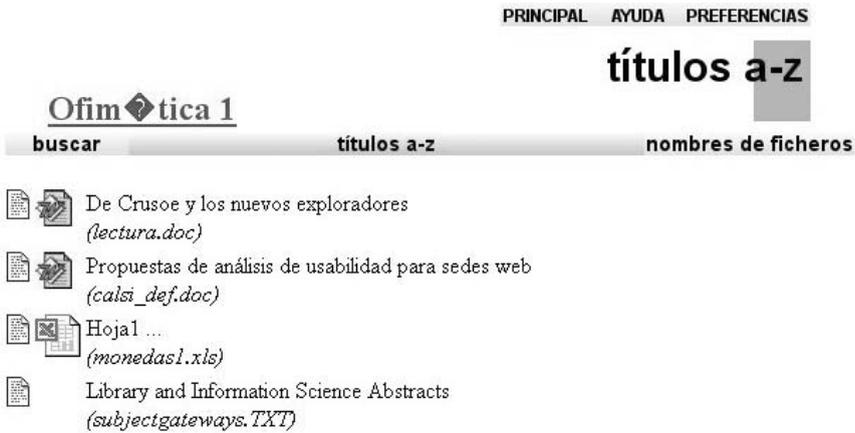


Figura 10.15. Presentación básica de los documentos por clasificador.

10.8. CONFIGURACIÓN AVANZADA DE LA COLECCIÓN

Las colecciones se construyen para los usuarios. En consecuencia, los contenidos de las mismas, y su organización, deben crearse pensando en el contexto del usuario final, sus técnicas y costumbres de acceso a la información. Si los usuarios acostumbran a organizar sus documentos por fechas, debería ofrecerse un clasificador por fecha del documento; si atienden a autores, debería disponerse de un índice o clasificador por autores, etc. Como puede imaginar, esto supone que, antes de sentarse delante de Greenstone, es necesario un trabajo previo de análisis, mediante el cual definir las características de la colección, los usuarios (reales y potenciales), y las formas más adecuadas de organizar la información para facilitar el acceso de los usuarios a ella.

En el ejemplo que se está tratando, resulta claro que, si se encuentra en un entorno ofimático de documentos digitales, además de disponer de un motor de búsqueda a texto completo, a los usuarios les va a interesar saber quiénes son los autores de los documentos, cuál es su título, una breve descripción de sus contenidos, disponer de una clasificación temática, etc. Por otra parte, es posible que interese modificar el formato de presentación de resultados, para ajustarlo a las necesidades reales de los usuarios. Greenstone es capaz de detectar automáticamente algunos metadatos, dependiendo del tipo de documento que se le suministre (en este ejemplo, ha identificado como título la primera línea o párrafo de cada documento), pero en otros casos es necesario incorporar metadatos externos a los documentos, como por ejemplo un resumen, o unos descriptores. Tomando como punto de partida estos metadatos, la aplicación permitirá generar índices de texto específicos, además del general, nuevos clasificadores, etc. Es recomendable que, antes de proceder a la creación de la colección, se disponga de un diseño previo, que reúna las siguientes cuestiones:

- Objetivos de la colección.
- Características y necesidades de información de los usuarios.

- Tipos de documentos que va a contener, y plugins necesarios para su procesamiento.
- Índices de texto necesarios para las búsquedas.
- Clasificadores para la organización de los documentos.
- Formatos de presentación de los resultados de las búsquedas y de los clasificadores.

Para desarrollar la colección de ejemplo siguiendo estos criterios, es necesario volver a poner en marcha GLI. Para abrir la colección, acuda al menú **Archivo**, opción **Abrir**. Se muestra un cuadro de diálogo con el listado de las colecciones disponibles en Greenstone. Seleccione su colección de ejemplo, y pulse sobre el botón **Aceptar**. GLI mostrará las pestañas de trabajo activas, y le dejará en **Reunir**, por si desea añadir nuevos documentos a la colección, siguiendo el procedimiento explicado en el apartado anterior. Si desea eliminar alguno de los documentos, bastará con seleccionarlo en el panel derecho, pulsar botón derecho del ratón, y seleccionar la opción **Suprimir** del menú emergente. Si es necesario, puede añadir directorios completos, arrastrando el icono correspondiente, y Greenstone incluirá el directorio y todos sus contenidos, creando estructuras jerárquicas dentro de la colección. Sin embargo, en este momento van a interesar las opciones que ofrecen las pestañas **Enriquecer** y **Diseño**. **Enriquecer** va a permitir añadir metadatos a los documentos, mientras que **Diseño** ofrecerá todas las opciones referidas a índices de texto, clasificadores y formatos de presentación. Todas las opciones de configuración son introducidas por GLI en el fichero `collect.cfg`. Cada colección tiene su propio fichero `collect.cfg`, que se localiza en la ruta (directorio raíz de Greenstone)/collect/(nombre de la colección)/etc/. En este fichero habrá órdenes de configuración que describan los datos generales de la colección, los plugins asignados, algunas opciones nuevas de configuración. Deberá volver a crear la colección tras las modificaciones.

Antes de pasar al siguiente apartado, es interesante hacer mención de la pestaña **Bajar** de la funcionalidad que ofrece. Como se ha podido comprobar, Greenstone incluye un plugin para documentos HTML, y es capaz de tratar documentos locales y externos. Para poder integrar los documentos o páginas web externas, situados en otros servidores, debe indicarse en la ventana **Bajar** el servidor a copiar, mediante el parámetro **URL fuente**, la profundidad de la descarga y los límites de tipos de ficheros a recuperar. La pulsación del botón **Bajar** lanzará una conexión que traerá a nuestro servidor una copia de todos los documentos y ficheros web. Para incorporar éstos a una colección, pulse sobre la pestaña **Reunir**, y en el panel de la izquierda, sobre la opción **Archivos descargados**. A partir de esto, ya puede seguir el procedimiento general de creación de colecciones, recordando que deberá seleccionar y configurar el plugin HTMLPlug.

10.8.1. ASIGNACIÓN DE METADATOS A DOCUMENTOS

Dependiendo del tipo de documento que procese Greenstone es capaz de detectar automáticamente metadatos dentro del documento, como título o autor. Esto será más sencillo procesando mensajes de correo electrónico o bases de datos que documentos de texto completo. En cualquier caso, se pueden asignar metadatos externos a los documentos, tanto de forma manual, como aprovechando ficheros externos ya creados, siempre y cuando sigan una DTD o XML Schema estándar de metadatos. La instalación de Greenstone incorpora varios con-

juntos de metadatos ya creados (ficheros con terminación .mds), que pueden aplicarse a los documentos. El más interesante es el correspondiente a Dublin Core (contenido en el fichero dublin.mds), ya que corresponde a la norma ISO Conjunto de metadatos Dublin Core (<http://www.dublincore.org>). Esta norma ha sido ampliamente aceptada como referencia de metadatos a escala mundial, y puede ser incorporada a documentos XML, e incluida en RDF (*Resource Description Framework*) sin problemas. Para asignar metadatos a los documentos de la colección de ejemplo, debe seguirse el siguiente procedimiento:

1. Pulse sobre la pestaña **Enriquecer**. El panel izquierdo muestra los documentos existentes en la colección, mientras que el derecho se muestra en gris, y será el usado para asignar los metadatos a cada uno de ellos.
2. Para asignar un conjunto de metadatos a la colección, vaya al menú **Diseño**, sección **Conjuntos de metadatos**. Pulse sobre el botón **Añadir conjunto de metadatos**. En la ventana de diálogo que se muestra, seleccione el fichero **dublin.mds**, y pulse sobre **Añadir...**
3. Ahora, el panel derecho de **Enriquecer** se ha activado, y se muestra el conjunto de metadatos seleccionado para los documentos (Figura 10.16). Los diferentes campos muestran una etiqueta que comienza por dc. (correspondiente a Dublin Core), seguida del metadato que se trate (título, autor, descripción...). El conjunto de metadatos Dublin Core se compone de 15 elementos, pero no es necesario completar todos y cada uno de ellos. Situando el cursor encima de cada etiqueta de la columna **Elemento**, se muestra una ayuda emergente sobre el contenido y la finalidad de cada metadato.

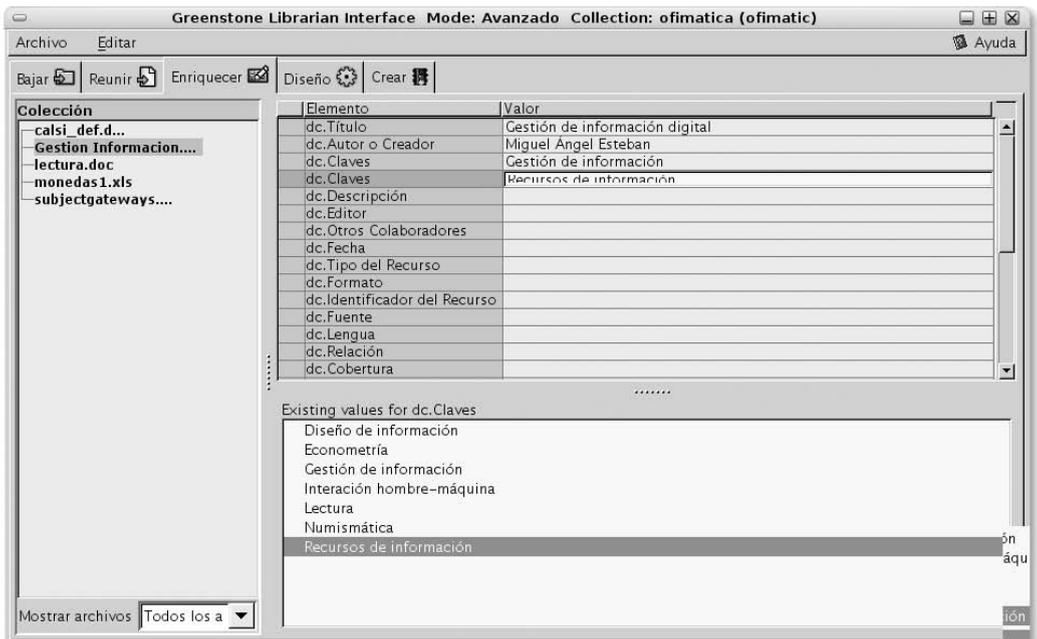


Figura 10.16. Panel derecho activado para la introducción de metadatos.

- Para comenzar a asignar los metadatos correspondientes a cada documento, seleccione éste en el panel de la izquierda. En el panel derecho, seleccione el metadato que desea introducir (por ejemplo, dc.Título). Sitúe el cursor en el campo e introduzca el título del documento. Conforme vaya agregando los títulos de los diferentes documentos, se irán mostrando en el espacio dedicado a **Valores existentes** (Figura 10.17).

Elemento	Valor
dc.Título	Propuestas de análisis de usabilidad para sedes web
dc.Autor o Creador	Jesús Tramullas
dc.Claves	Diseño de información
dc.Descripción	
dc.Editor	
dc.Otros Colaboradores	
dc.Fecha	
dc.Tipo del Recurso	
dc.Formato	
dc.Identificador del Re...	
dc.Fuente	
dc.Lengua	
dc.Relación	
dc.Cobertura	
dc.Derechos	
ex.Encoding	utf8
ex.FileFormat	Word

Existing values for dc.Autor o Creador

- Jesús Tramullas
- John Smith
- Manuel Martín-Bueno
- Miguel Ángel Esteban

Figura 10.17. Introducción de metadatos para un documento.

- Para continuar introduciendo otros metadatos, sólo hay que seguir el mismo procedimiento. En este ejemplo, se han introducido metadatos para autor, claves (palabras clave) y descripción (o resumen). Es importante recordar que la forma correcta de asignar metadatos es crear una entrada independiente para cada uno de ellos. Por ejemplo, para claves se debería introducir una palabra clave, y pulsar *Retorno*, introducir otra nueva palabra clave, y así sucesivamente. GLI creará automáticamente una entrada independiente para cada una de ellas (Figura 10.18). El mismo proceso debe seguirse con metadatos que puedan ofrecer valores múltiples, como autor, editor, otros colaboradores... De esta forma se podrán ajustar con mayor precisión los clasificadores, usar clasificaciones de contenido, como tesauros y ontologías, para explorar los documentos y facilitar el acceso a la información a los usuarios. En el caso de que quiera establecer una estructura jerárquica de metadatos, introduzca todos en una única entrada, separándolos con el símbolo “|”, en la forma “unoldos”.

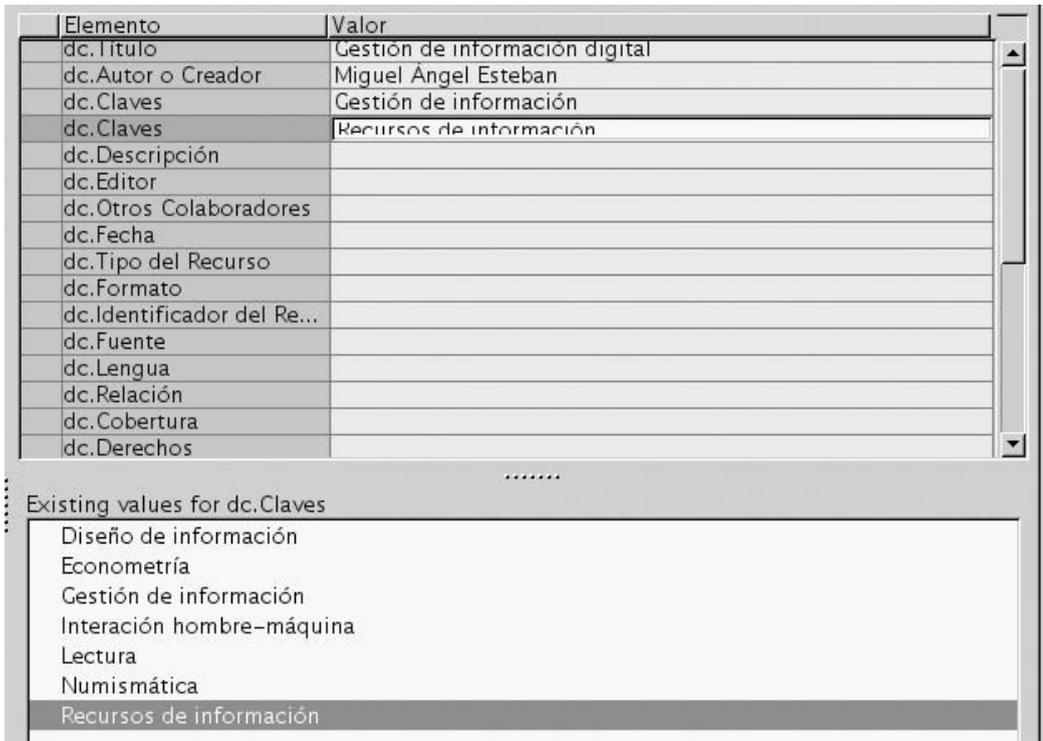


Figura 10.18. Entradas independientes para cada metadato del mismo tipo.

- Por último, se añade una descripción a cada documento. Para ello, se sitúa el cursor en el campo `dc.Descripción`. Evidentemente, una descripción o resumen ocupan más espacio que el disponible en el campo **Valor**. Para solucionar este inconveniente, haga doble clic sobre el campo. Se abre una ventana especial, en la que puede introducir una mayor cantidad de texto (Figura 10.19). Pulse **Aceptar** una vez introducido el resumen correspondiente.
- Si selecciona dos o más documentos en el panel izquierdo simultáneamente, el panel derecho mostrará los metadatos combinados de todos ellos. Mostrará en negro los metadatos comunes a los documentos, y en gris aquellos que no lo son. Ésta puede ser una forma interesante de asignar o modificar metadatos rápidamente a documentos que ofrezcan características similares de autoría, contenido, fechas de referencia, etc.
- En el caso de que la colección incluyese directorios y subdirectorios, se pueden asignar metadatos a los propios directorios, además de a los documentos. En este caso, al existir una jerarquía, se heredarían jerárquicamente los metadatos asignados.

Desde la versión 2.51 se ha incorporado un editor de metadatos, llamado *Greenstone Editor for Metadata Sets*, o GEMS. Esta aplicación Java, disponible en el directorio `/gli` de la instalación, facilita la creación de conjuntos de metadatos definidos por el usuario, que luego estarán disponibles para ser asignados a las colecciones, a través de GLI.

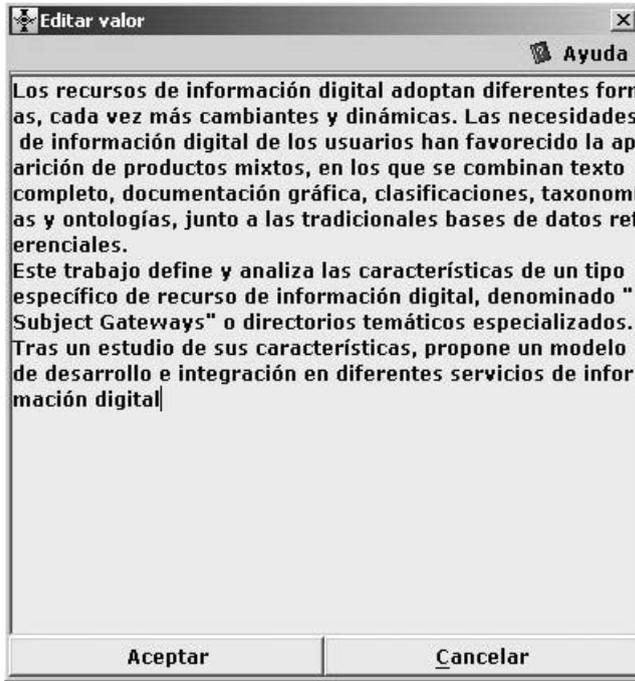


Figura 10.19. Ventana de introducción de gran cantidad de texto, para el metadato dc.Descripción.

10.8.2. OPCIONES GENERALES DE LA COLECCIÓN

Una vez introducidos los metadatos correspondientes a los documentos que forman la colección, es necesario pasar a la configuración detallada de la misma. Como se ha indicado en apartados anteriores, las opciones y parámetros necesarios se encuentran bajo la pestaña **Diseño**. En el panel izquierdo se muestra el menú de configuración. La selección de las diferentes entradas mostrará en el panel derecho las opciones y parámetros correspondientes a cada una de ellas. Para empezar, es necesario cambiar las opciones generales.

1. Pulse sobre la entrada **General**. Se muestran los datos básicos que se introdujeron al crear la colección. Habrá observado que, en comparación con las colecciones de muestra, la que ha creado carece de icono gráfico de identificación, y que los caracteres acentuados pueden representarse incorrectamente. En primer lugar, compruebe que la dirección de correo electrónico del creador son correctas. En ocasiones las palabras acentuadas no se muestran correctamente. Si es necesario, utilice los códigos de HTML para caracteres acentuados.
2. Puede definir los ficheros de imágenes que va a usar como iconos para la colección. Para este ejemplo, se ha creado un sencillo logo con la herramienta gráfica The GIMP (<http://www.gimp.org>), que también es software libre, y ofrece versiones para Linux/UNIX y Windows. Pulse sobre los botones **Explorar** para localizar

los ficheros gráficos en su estructura de directorios, y selecciónelos siguiendo los pasos normales. Greenstone admite formatos GIF, JPEG y PNG.

3. En el apartado **Descripción de la colección** puede introducir el texto que considere oportuno para orientar al usuario sobre el contenido y la organización de la colección y de sus documentos. Puede utilizar cualquier combinación de texto y HTML. En el ejemplo, se ha creado una tabla y se ha incorporado un texto breve de explicación (Figura 10.20).

Opciones generales

La sección de diseño de la Interfaz de la Biblioteca Digital le permite controlar varios aspectos relativos a la apariencia de su colección. El diseño se divide en varias secciones. Esta sección contiene las opciones y ajustes generales. Para ver una sección diferente, haga click en el nombre de la sección en la lista que se encuentra a la izquierda.

E-mail del creador de la colección: tramullas@unizar.es

E-mail del administrador de la colección: tramullas@unizar.es

Título de la colección: ofimatica

Carpeta de la colección: ofimatic

URL del icono de la página 'acerca de':

URL del icono de la 'página principal':

La colección será de acceso público

Descripción de la colección:

```
<table width="80%" border="0" cellspacing="0" cellpadding="1" bgcolor="red">
<tr>
<td>
</td>
</tr>
<table width="100%" border="0" cellspacing="0" cellpadding="12" bgcolor="white">
<tr>
<td>
<p>Colección de ejemplo de documentos ofimáticos en la que se han incluido documentos en formatos word, excel, pdf y texto. Incluye un conjunto de metadatos Dublin Core</p>
<p>Se han definido clasificadores propios por autores, títulos y palabras clave</p>
</td>
</tr>
</table>
</td>
</tr>
</table>
</td>
</tr>
```

Figura 10.20. Panel de configuración de Opciones generales.

4. Una vez seguidos los pasos anteriores, y para comprobar los cambios efectuados, vaya a **Archivo**, opción **Guardar**. Visualice el resultado a través de su cliente web (Figura 10.21). Podrá observar cómo se han solucionado los problemas previos: aparece un icono propio de la colección, y el párrafo de ayuda definido en GLI.

Este primer paso hacia la adecuada configuración de la colección hace posible crear una primera imagen general. Sin embargo, esta apariencia puede ser modificada más profundamente utilizando los ficheros macro que sirven para generar de manera dinámica las páginas que sirve Greenstone. Los ficheros de macro de Greenstone son aquellos que, situados dentro del directorio de la aplicación, tienen la terminación `.dm`. La programación de las macros de Greenstone sobrepasa el objetivo de este capítulo, pero el administrador encontrará la información necesaria en la *Greenstone Developer's Guide*.



Figura 10.21. La nueva presentación de la página principal de la colección.

10.8.3. SELECCIÓN Y CONFIGURACIÓN DE PLUGINS

El procesamiento de documentos para su inclusión en las colecciones tiene como elemento clave la selección y configuración de los plugins que han de tratarlos. Greenstone puede trabajar con gran cantidad de documentos digitales, en diferentes formatos, e integrarlos en una colección, mediante la adecuada configuración de plugins. Como se indicó al explicar el esquema de funcionamiento de Greenstone, los documentos originales, en diferentes formatos, son procesados y transformados en XML por programas externos a Greenstone. Estos programas pueden haber sido creados por el equipo de desarrollo de la aplicación, o bien ser programas de terceros, pero todos ellos son también software libre. Con esta arquitectura, la aparición de nuevos formatos ofimáticos, o la ampliación de prestaciones a otros tipos de documentos aún no contemplados, sólo depende de la disponibilidad del plugin necesario. De la misma manera, la aparición de una nueva versión de un plugin ya existente puede ser integrada sin dificultad en las instalaciones previas.

La construcción de una colección requiere que sean definidos claramente los tipos de documentos que se van a incluir, para seleccionar los plugins necesarios. Los plugins pueden ser configurados usando parámetros especiales, si se necesita. Cada vez que se crea una colección, Greenstone asigna un conjunto básico de plugins, que suele ser necesario modificar, para ajustarlo según las características de los documentos a procesar. En la colección de documentos ofimáticos que se está configurando se han incluido documentos en formatos word, excel, txt y pdf.

1. En la pestaña **Diseño** pulse sobre **Plugins de documento**, en el menú del panel izquierdo.
2. En el panel derecho se muestra el listado de plugins asignados. Observe que en el ejemplo se había incluido previamente el plugin para documentos Excel, ExcelPlug, ya que en la colección de ejemplo se añadió un documento de este tipo, y se habían eliminado los plugins EMAILPlug y ZIPPlug. La Figura 10.22 muestra el estándar antes de las modificaciones.



Figura 10.22. Configuración estándar de plugins.

3. En la asignación de plugins, dos de ellos se muestran en el panel en la parte inferior, separados del resto por una línea horizontal. Se trata de ArcPlug y RecPlug. Ambos son imprescindibles para el funcionamiento de Greenstone, de tal forma que sólo se puede cambiar su configuración, pero no se pueden eliminar del listado. Además, existe un plugin básico, BasPlug, que no aparece en el listado, pero que es necesario para el procesamiento de los documentos. A ello hay que añadir GAPIug, que aunque puede eliminarse del listado, hay que recomendar que se mantenga. El creador debe tener en consideración que el procesamiento de los documentos puede necesitar de más de un plugin, es decir, que primero se pase a un formato, y luego a otro. Por ejemplo, ExcelPlug transforma el documento Excel en texto plano o en HTML, y, en consecuencia, se necesitará tam-

bién incluir HTMLPlug o TextPlug. Ésta es la razón por la que se mantiene HTMLPlug en el ejemplo (Figura 10.23), aunque no haya documentos HTML en la colección.

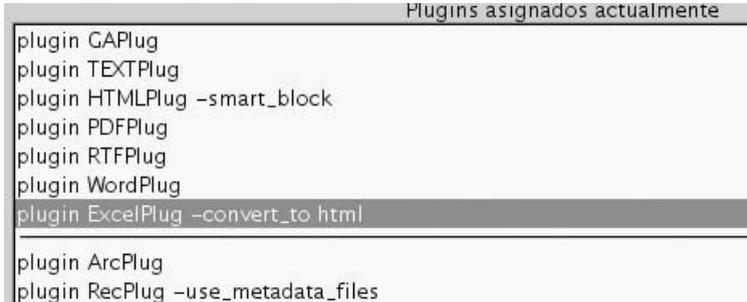


Figura 10.23. Listado de plugins de la colección de ejemplo.

- El procedimiento para asignar nuevos plugins es sencillo. Primero, en el área **Controles de edición**, en la parte inferior del panel, despliegue el menú **Seleccione el plugin que se va a añadir**. En el listado que se muestra entonces aparecen todos los plugins disponibles en la instalación de Greenstone (Figura 10.24). Por ejemplo, para añadir un plugin capaz de procesar imágenes, seleccione **ImagePlug** del listado de plugins, y pulse sobre **Añadir plugin**.

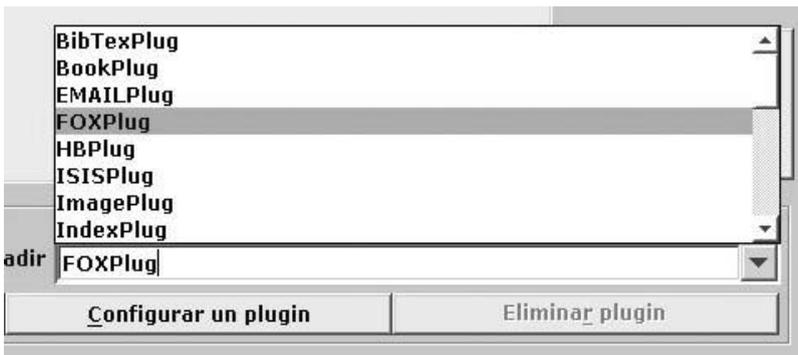


Figura 10.24. Menú desplegable de selección de plugins.

- Se abre la ventana de configuración del plugin. Puede cambiarse la configuración de los plugins a través de este tipo de diálogos. Para ello, hay que seleccionar el plugin, y pulsar sobre el botón **Configurar un plugin**. Se muestra un conjunto de parámetros que pueden modificarse para ajustar las opciones de importación y construcción de la colección (Figura 10.25). Pulse el botón **Aceptar** para asignar el plugin al fichero de configuración.

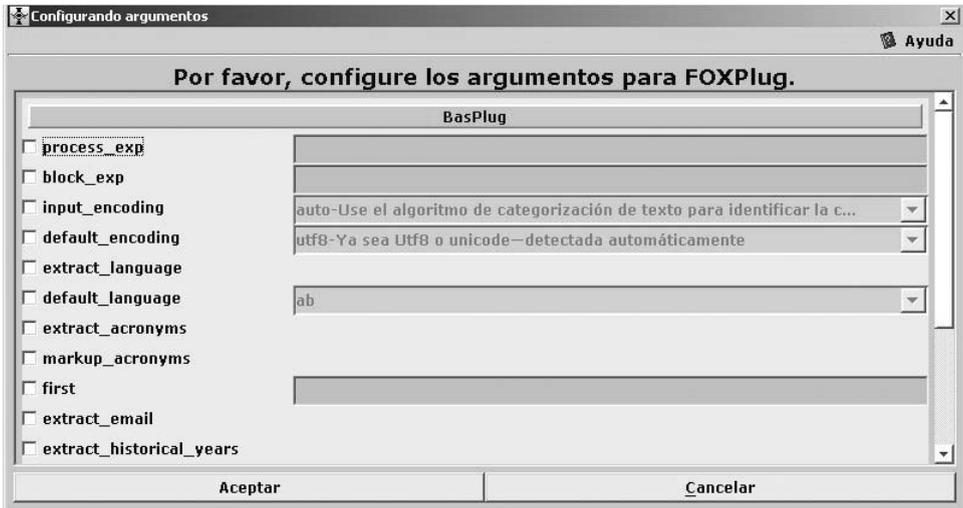


Figura 10.25. Ventana de diálogo para configurar el plugin.

6. Compare los parámetros de configuración de diferentes plugins. En cada uno de ellos podrá encontrar, en primer lugar, parámetros del propio plugin; después puede mostrar parámetros que se pasan a otro plugin para que lo procese posteriormente; tras estos parámetros para BasPlug, y por último, parámetros propios del usuario, bajo la etiqueta **A la medida, Parámetros personalizados**. Por ejemplo, si selecciona PDFPlug, y pulsa el botón **Configurar un plugin**, la ventana de diálogo le mostrará la posibilidad de configurar parámetros para PDFPlug, ConvertToPlug, HTMLPlug, BasPlug y A la medida (Figura 10.26).

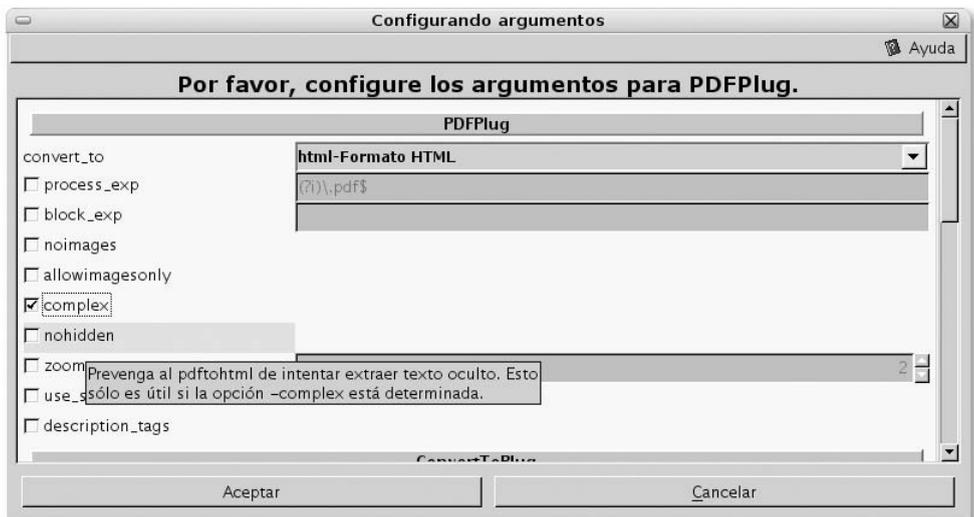


Figura 10.26. Configuración de PDFPlug.

Esto significa que un documento en formato PDF se ha procesado en primer lugar por PDFPlug, que lo pasa a formato HTML, y un intermediario lo reenvía a HTMLPlug, que tras su proceso propio lo pasa, a su vez, al BasPlug. Situando el cursor encima de cada etiqueta de parámetro, GLI le muestra un texto de ayuda sobre el parámetro. La documentación sobre los plugins y sus parámetros no es tan detallada como sería deseable. Para activar un parámetro, pulse sobre la caja de selección situada delante del mismo, e introduzca el argumento en el campo situado a la derecha de la etiqueta de parámetro.

7. En ocasiones, a pesar de haber asignado los plugins adecuados, pueden producirse problemas de rechazo de documentos, que no son incluidos en la colección (esta información la obtendrá cuando ejecute las opciones de creación de la colección). Esto puede ocurrir debido a dos causas principales:
 - a) El formato del documento tiene alguna característica que produce un error en el plugin que lo importa. Hay documentados problemas con ficheros PDF, que, por razones no aclaradas, no son admitidos por PDFPlug. Otro problema lo pueden plantear ficheros a los que se han asignado claves de seguridad en sus aplicaciones de origen.
 - b) El orden en que son cargados los plugins en el fichero de configuración también puede producir problemas. Ésta es la razón de disponer de botones **Mover arriba** y **Mover abajo**, que sirven para variar el orden de los plugins. Por ejemplo, es muy recomendable situar GAPlug y HTMLPlug al inicio de la lista.

El conjunto de plugins que incorpora Greenstone es capaz de procesar documentos textuales, presentaciones, bases de datos, imágenes y bibliografías. La documentación sobre los plugins, su funcionamiento y configuración puede encontrarse en la *Greenstone Developer's Guide*. En la Tabla 10.2. se resumen los principales plugins disponibles.

Tabla 10.2. Plugins disponibles en Greenstone

Plugin	Documentos que procesa	Indicaciones
BasPlug	Todos	Plugin fundamental, procesa todos los documentos para integrarlos en Greenstone, identifica idiomas y metadatos básicos de forma automática.
ArcPlug	Todos	Fundamental, comunica los documentos importados con el proceso de creación de la colección.
RecPlug	Todos	Fundamental, explora la estructura de la colección si hay directorios y subdirectorios, incorpora los metadatos a la colección.
GAPlug	Todos, .xml	Fundamental, procesa los documentos ya en formato Greenstone, tras la importación.

Tabla 10.2. Plugins disponibles en Greenstone (*continuación*)

Plugin	Documentos que procesa	Indicaciones
TextPlug	Texto plano, .txt	Opcional.
WordPlug	Microsoft Word, .doc	Opcional, extrae como metadato título la primera línea del documento, conveniente completar con HTMLPlug.
PDFPlug	Adobe Acrobat, .pdf	Opcional, extrae como metadato título la primera línea del documento, conveniente completar con HTMLPlug. En ocasiones rechaza documentos sin motivos aparentes.
PSPlug	PostScript, .ps, .eps	Opcional, puede necesitar un visor de PostScript, se recomienda GhostScript, también software libre.
ExcelPlug	Microsoft Excel, .xls	Opcional, extrae como metadato título la primera línea del documento, conveniente completar con HTMLPlug.
HTMLPlug	Páginas web, .htm, .html, .asp, .php, cgi,.shtml	Opcional. Páginas web en diferentes formatos, extrae los metadatos básicos de las etiquetas HTML. Mantiene activos los enlaces a otras páginas web. Recomendable incluirlo como complemento a otros plugins.
ImagePlug	Ficheros gráficos, .jpg, .jpeg, .gif, .tif, .tiff, .png, .bmp	Opcional, para crear colecciones de imágenes. Según la documentación, sólo funciona en Linux/UNIX, y necesita la instalación previa de Imagemagick. En realidad, con la misma configuración puede ejecutarse en sistemas Windows.
EMAILPlug	Correo electrónico, ficheros que acaben en dígitos, o en dígitos.email	Opcional, para mensajes de correo electrónico, extrae como metadatos las cabeceras de los mensajes. Puede dar problemas con mensajes en formato MIME.
LaTextPlug	LaTEXT	Opcional, para ficheros creados con TEXT y con LaTEXT.
MARCPlug	Ficheros en formato MARC, Machine Readable Record	Opcional, para registros bibliográficos de sistemas de informatización de bibliotecas.
PPTPlug	Microsoft Powerpoint, .ppt	Opcional, conveniente completar con HTMLPlug.
ISISPlug	Bases de datos documentales en formato CDS/ISIS, .mst, .fdt, .xrt	Opcional.
MP3Plug	Ficheros de sonido MP3, .mp3	Opcional, para crear colecciones sonoras.

Tabla 10.2. Plugins disponibles en Greenstone (*continuación*)

Plugin	Documentos que procesa	Indicaciones
OAIPlug	Ficheros de la Open Archives Initiative, OAI	Opcional, sirve para crear colecciones de registros en el marco del proyecto internacional Open Archives Initiative (http://www.oai.org), permite la conexión a servidores OAI remotos.
ZIPPlug	Ficheros compactados, .tar, .zip, .gzip, .bzip, .gz, .tgz	Procesa el contenido de ficheros compactados, pero necesita de programas externos para ello.

También se encuentran disponibles plugins para otros formatos de fichero, que no se han recogido en la Tabla 10.2 al tener poco uso en el ámbito español, como por ejemplo BibTextPlug, ReferPlug y ProCitePlug (para bibliografías en formatos BibText, Refer y Procite, respectivamente), SRCPlug (ficheros de código fuente), o GBPlug (textos electrónicos del Proyecto Gutenberg).

10.8.4. TIPOS E ÍNDICES DE BÚSQUEDA

La interfaz de búsqueda a texto completo que ofrece Greenstone, mediante el botón **Buscar**, responde al esquema clásico, al estilo de los motores de búsqueda de Internet. Sin embargo, en colecciones en las que los documentos ofrezcan una estructura muy estricta, como mensajes de correo o bases de datos, es posible configurar una interfaz de búsqueda por campos, con un mayor nivel de delimitación.

1. Para ello, pulse sobre la entrada **Tipos de búsqueda**, en el panel izquierdo.
2. En el panel derecho, marque la casilla **Permitir búsquedas avanzadas**. Se activa el espacio inferior, que muestra la opción **form**. Debe seleccionar el motor de indicación que va a utilizar. Desde la versión 2.62, se puede elegir entre *mgpp* y *Lucene*.
3. En el menú desplegable **Tipos de búsqueda**, seleccione **plain**, y pulse sobre el botón **Añadir tipo de búsqueda**. En el cuadro **Tipos de búsqueda**, la primera de ellas será la disponible por definición, y la segunda estará accesible a través de la opción **Preferencias** de la colección (Figura 10.27). Si desea cambiar la configura-

**Figura 10.27.** Diálogo de definición de tipos de búsqueda.

ción posteriormente, bastará con desmarcar la casilla **Permitir búsquedas avanzadas**, y Greenstone volverá a mostrar la búsqueda simple.

4. Si accede a la colección a través del servidor Greenstone, podrá observar la nueva interfaz de búsqueda a texto completo, basada en campos y operadores booleanos (Figura 10.28).

Figura 10.28. Interfaz de búsqueda por campos, aplicando los índices creados en el apartado siguiente.

La definición de la interfaz avanzada debe completarse con la selección de los índices textuales sobre los que llevar a cabo las búsquedas de cadenas de texto. La creación de índices específicos para determinados campos y/o metadatos (por ejemplo, un índice de autores, o un índice de resúmenes), permite efectuar la búsqueda contra índices más pequeños (y más rápidos), y de contenidos especializados (por ejemplo, un índice sólo del contenido de resúmenes dará resultados más específicos que el índice de todo el texto de los documentos).

Para definir índices específicos es necesario seguir el siguiente procedimiento:

1. Pulse sobre la entrada **Índices de búsqueda**, en el panel izquierdo. En el campo **Índices asignados** puede aparecer alguno previo, o bien aparecer vacío.
2. Para añadir un índice, introduzca un **Nombre de índice** en el campo correspondiente. Es recomendable que sea significativo (por ejemplo, autores, resúmenes...), ya que será el nombre que se muestre en la interfaz de búsqueda.
3. Despliegue el menú **Fuente del índice**. Se muestra un listado de los posibles índices. Es recomendable disponer de un índice **text**, ya que asegura que el texto completo de todos los documentos será indizado, y se podrá buscar contra él. Pulse sobre el botón **Añadir índice**. El índice creado se mostrará en el campo **Índices asignados**.
4. Se pueden crear índices sobre los metadatos que fueron introducidos en la pestaña **Enriquecer**. Para ello, en **Fuente del índice** busque aquellos que comienzan

por dc. Por ejemplo, para crear un índice de palabras clave, introduzca esta expresión en el campo **Nombre del índice**. Después, seleccione del menú desplegable **Fuente del índice** la entrada dc.Claves. Por último, pulse sobre el botón **Añadir índice**.

5. Añada índices similares para títulos y para autores. Una vez terminado, puede usar los botones **Mover arriba** y **Mover abajo** para modificar el orden de presentación en el menú desplegable de la interfaz de búsqueda (Figura 10.29).



Figura 10.29. Asignación de índices para la colección de ejemplo, mostrando el menú de selección desplegado.

6. Si dese eliminar algún índice, pulse sobre la entrada del mismo en el campo **Índices asignados**, y pulse el botón **Eliminar índice**.

En el menú desplegable se muestran más opciones. Las que comienzan por ex. son metadatos extraídos automáticamente por Greenstone de los documentos. Los correspondientes a Dublin Core facilitados por el creador de la colección comienzan por dc. **Allfields** añadiría automáticamente todos los índices posibles.

Si se está procesando documentos complejos, con múltiples parte, como secciones o subsecciones, es interesante revisar las opciones que se muestran al pulsar la pestaña **Administrar niveles**. Si Greenstone ha detectado estas secciones en los documentos que forman la colección, o se han creado en los documentos XML correspondientes, pueden definirse índices a nivel de documentos completo, de secciones y de párrafos.

Si no se ha activado la opción **Permitir búsquedas avanzadas**, explicada en párrafos anteriores, observe que el panel de **Índices de búsqueda** cambia su presentación (Figura 10.30). En cualquier caso, el procedimiento para crear los índices es similar. Si cambia el tipo de búsqueda, deberá volver a crear los índices desde el comienzo. En todos los casos siempre los índices continuarán disponibles, para usarlos desde la interfaz de búsqueda sencilla (Figura 10.31).

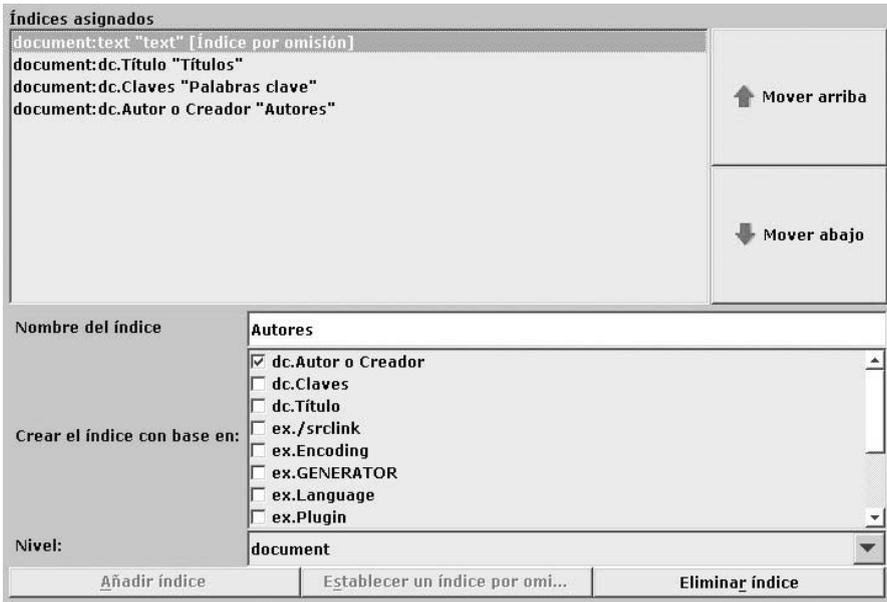


Figura 10.30. Panel de creación de índices con la Búsqueda avanzada desactivada.

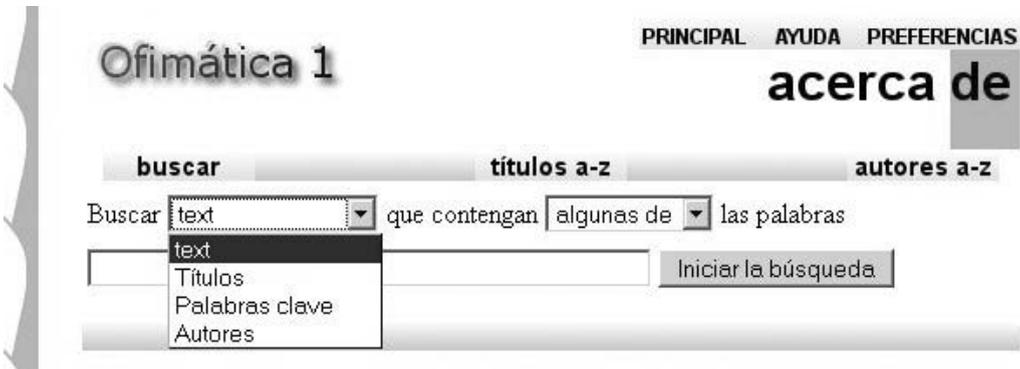


Figura 10.31. Índices disponibles en la interfaz de búsqueda sencilla.

10.8.5. FILTROS Y BÚSQUEDAS CRUZADAS

Greenstone facilita la creación de filtros, que sirven para fijar una limitación previa a las búsquedas. Por ejemplo, se puede fijar un filtro que limite las búsquedas a un subconjunto de documentos, que cumplan una condición previa (un autor, una palabra clave...). Para decidir si es necesario, o no, establecer un filtro sobre los documentos es necesario hacer un análisis previo. El procedimiento de crear un filtro es sencillo, ya que basta con pulsar sobre la opción **Particionar índices** en el panel izquierdo, y rellenar los campos de **Nombre del filtro de la subcolección, Atributo del documento...** (que contiene el listado de meta-

datos disponibles), **Expresión regular...** (que debe ser una expresión regular en lenguaje Perl), decidir si **Incluir** o **No incluir**, e introducir los **Indicadores que se usarán**. Para activar los filtros creados, es necesario pulsar sobre la pestaña **Asignar particiones**, e introducir un **Nombre de la partición**, seleccionar una **Partición asignada a la colección**, y pulsar sobre el botón **Añadir partición**. Por último, puede completarse el filtro, usando como criterio el idioma, pulsando sobre la pestaña **Asignar idiomas**, y seleccionando el idioma en el menú **desplegable Idioma que se va a añadir**. En la colección de ejemplo se han definido dos filtros. El resultado de esta definición de filtros es la aparición, en la interfaz de búsqueda, de un menú desplegable que da la posibilidad de seleccionar los filtros (Figura 10.32).

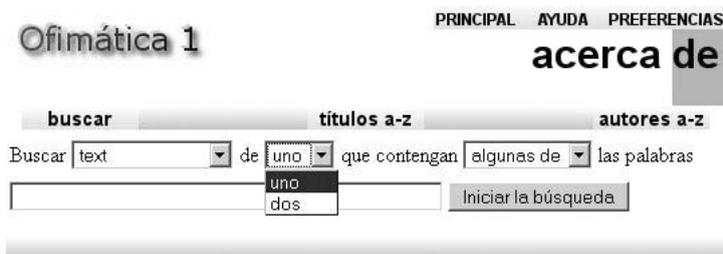


Figura 10.32. La interfaz de búsqueda con los filtros activados.

La segunda posibilidad que ofrecen los índices es la capacidad de hacer búsquedas contra varias colecciones, de manera simultánea. Pulse sobre la entrada **Búsqueda cruzada a través de varias colecciones**. Se muestra un listado de las colecciones disponibles en el servidor. Para seleccionar las colecciones sobre las que ejecutar la búsqueda, marque sobre las casillas de verificación situadas a la izquierda de cada colección. Una vez reconstruida la colección, la pantalla web inicial de la misma le informa de las colecciones sobre las cuales va a buscar. En caso de que el usuario desee buscar sólo sobre alguna de ellas, la página **Preferencias** le permitirá seleccionar, mediante casillas de verificación, las necesarias (Figura 10.33). Debe recordar que esta búsqueda cruzada o simultánea funciona correctamente cuan-

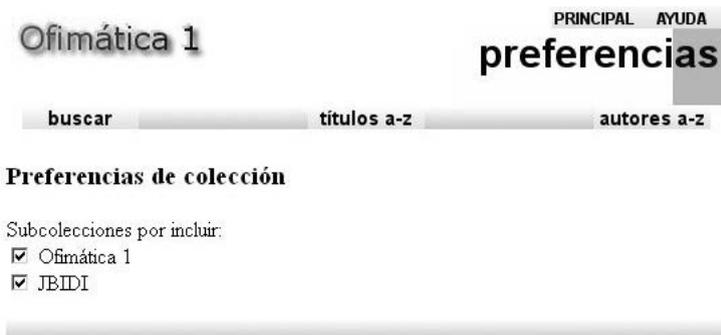


Figura 10.33. Preferencias para la búsqueda simultánea en varias colecciones.

do en la colección existen índices comunes, contruidos sobre metadatos comunes (por ejemplo, en ambas colecciones un índice de autores, construido sobre el metadato dc.Autor).

10.8.6. CLASIFICADORES DE EXPLORACIÓN

La capacidad de búsqueda a texto completo es fundamental para el acceso a la información contenida en colecciones de documentos. Sin embargo, cada vez es más necesario disponer de herramientas que faciliten al usuario acceder a la información mediante técnicas de exploración (*browsing*). Greenstone le ofrece la posibilidad de definir y utilizar estructuras de organización y exploración de la colección de los documentos, atendiendo a las características de contenido y a los metadatos de los documentos. Estas estructuras reciben el nombre de clasificadores, y son accesibles a través de las opciones que la página web de la colección muestra en su barra de botones, situados a la derecha del botón **Buscar**, que se muestra primero (véase la Figura 10.32). Se pueden definir diferentes clasificadores para las colecciones, en virtud de aquellas estructuras que sean más adecuadas a los conocimientos y capacidades del usuario. El tipo y las características de los clasificadores se definen rápidamente mediante GLI, y, al igual que el resto de las opciones tratadas en este apartado, se incorporan como órdenes en el fichero de configuración de la colección.

1. Pulse sobre la entrada **Clasificadores de exploración**, en el panel derecho de GLI.
2. Borre los exploradores creados previamente por Greenstone, seleccionándolos y pulsando sobre el botón **Eliminar clasificador**.
3. Para crear un nuevo clasificador, en el cuadro **Controles de edición**, despliegue el menú **Seleccione el clasificador que desea añadir**, seleccione **AZList**, y pulse sobre el botón **Añadir clasificador** (Figura 10.34).



Figura 10.34 Selección del tipo de clasificador.

4. Se abre entonces la ventana de configuración del clasificador. En el parámetro **metadata**, seleccione el valor dc.Título del menú desplegable, y pulse **Aceptar** (Figura 10.35). Se ha creado un clasificador que mostrará un listado de todos los documentos existentes en la colección, que contendrá el título de los documentos según el metadato dc.Título, ordenado alfabéticamente.

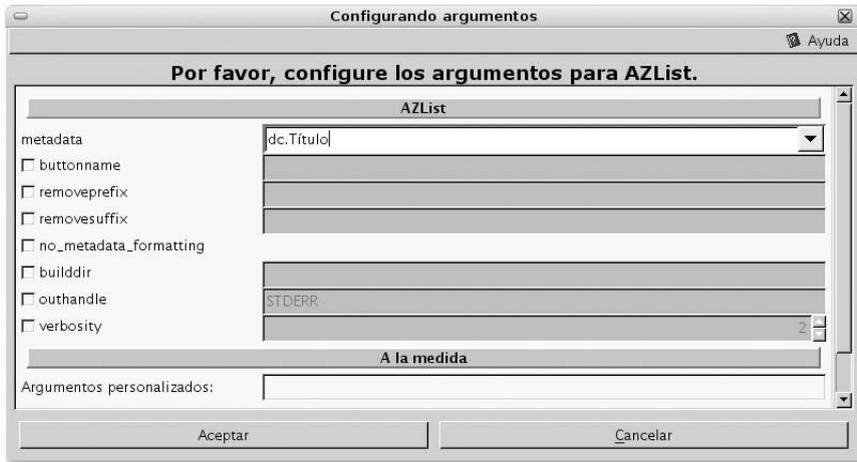


Figura 10.35. Ventana de definición de parámetros de AZList.

- Para crear un segundo clasificador, siga lo indicado en el punto 3, pero ahora seleccione un clasificador de tipo **AZCompactList**. Pulse **Añadir clasificador**. La ventana de opciones de configuración muestra más parámetros (Figura 10.36), entre los cuales son interesantes los siguientes:

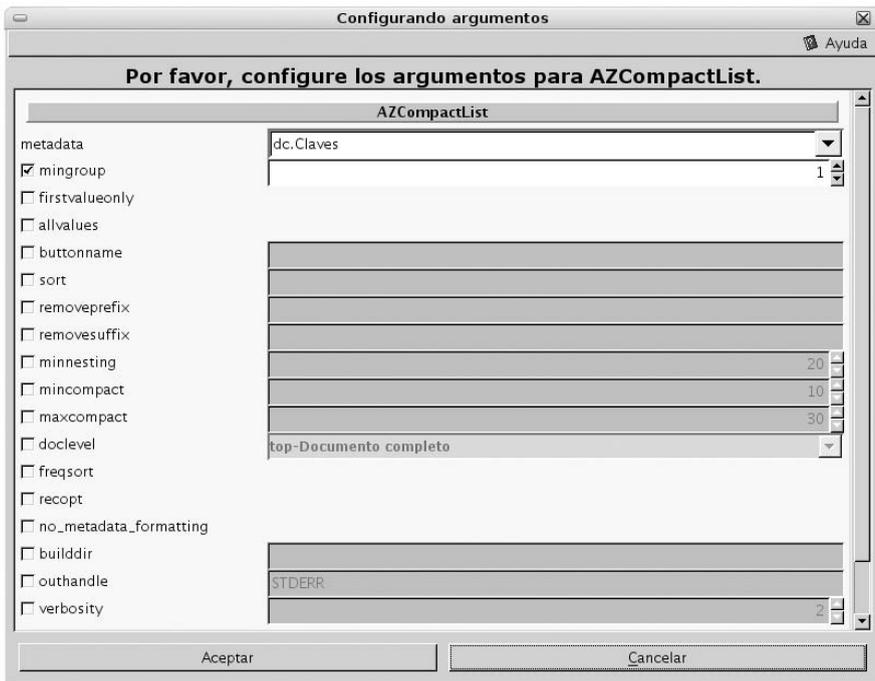


Figura 10.36. Opciones de configuración de AZCompactList.

- a) **Metadata:** ya utilizado, sirve para definir el metadato que se va a usar como criterio para organizar los documentos dentro del clasificador.
 - b) **Firstvalueonly/Allvalues:** determina si se usará sólo el primer metadato, o todos, para organizar los documentos dentro del clasificador. Estos parámetros son interesantes en colecciones con metadatos repetidos, tipo dc.Clave o dc.Autor.
 - c) **Mingroup:** número mínimo de documentos que, con el mismo metadato, formarán un agrupamiento jerárquico desplegable (por ejemplo, coincidencia de autor o de palabra clave). **Minnesting** determinará la creación, no de un grupo, sino de una lista anidada.
 - d) **Mincompact/Maxcompact:** número mínimo y máximo de documentos que se mostrarán por página. Greenstone añadirá los iconos de navegación necesarios.
 - e) **Freqsort:** determina si la ordenación dentro del clasificador se hace por la frecuencia de aparición de metadatos similares, y no por ordenación alfabética.
6. Pruebe a variar alguna de ellas. Marque la casilla de verificación situada a la izquierda de cada parámetro, y cambie el valor asignado, si es necesario. Para terminar, pulse sobre el botón **Aceptar**. Vuelva a construir la colección, y acceda a la misma a través de su cliente web. Observe el resultado de los clasificadores de título (Figura 10.37) y de autor (Figura 10.38). Observe también que Greenstone asigna directamente los botones y los iconos necesarios.

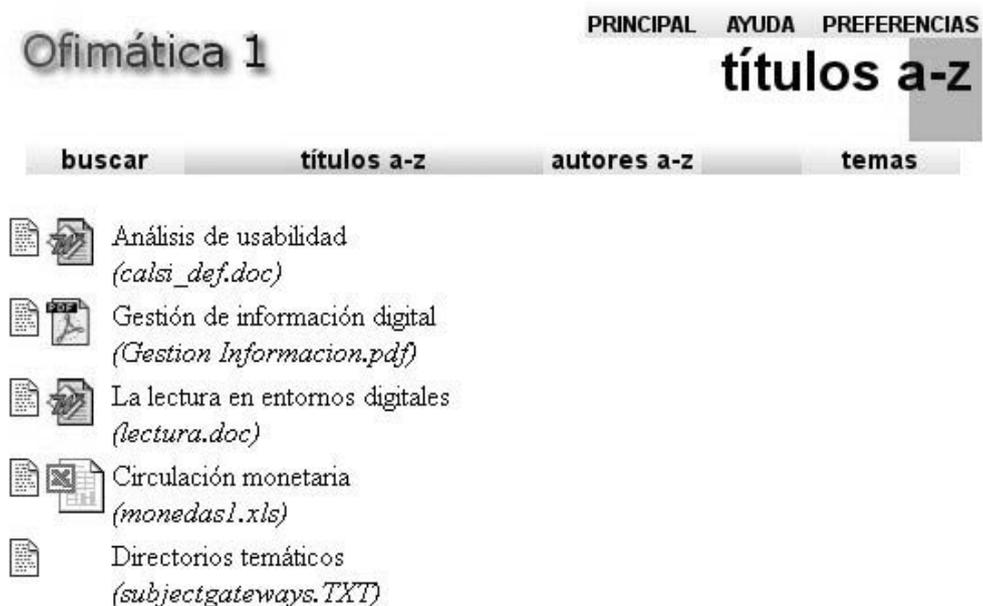


Figura 10.37. Documentos mostrados bajo el clasificador AZList para Títulos.

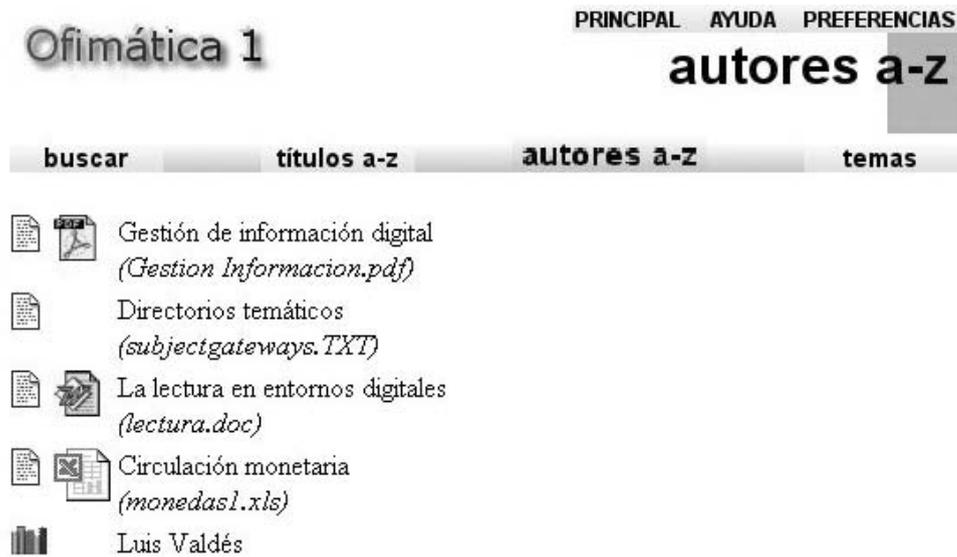


Figura 10.38. Documentos mostrados bajo el clasificador AZCompactList para Autores. Se anidan los documentos en los que se repite el autor.

- En el ejemplo de las figuras, se ha creado un tercer clasificador **AZCompactList**, que usa el truco de Mingroup 1, para que los documentos se muestren desde el comienzo agrupados por palabras clave (Figura 10.39). La pulsación en el icono correspondiente despliega el listado de documentos que tienen esa palabra clave (Figura 10.40).

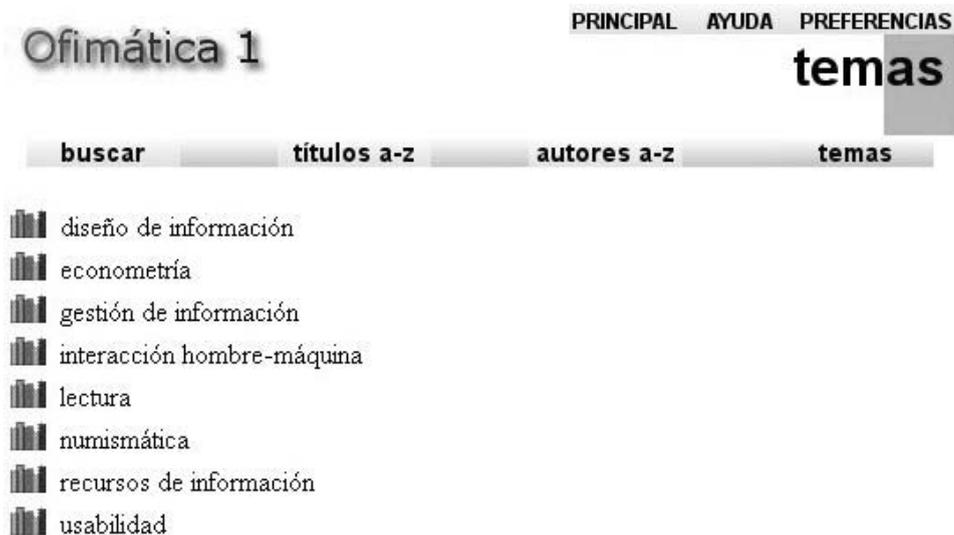


Figura 10.39. Clasificador AZCompactList, con documentos anidados en categorías.

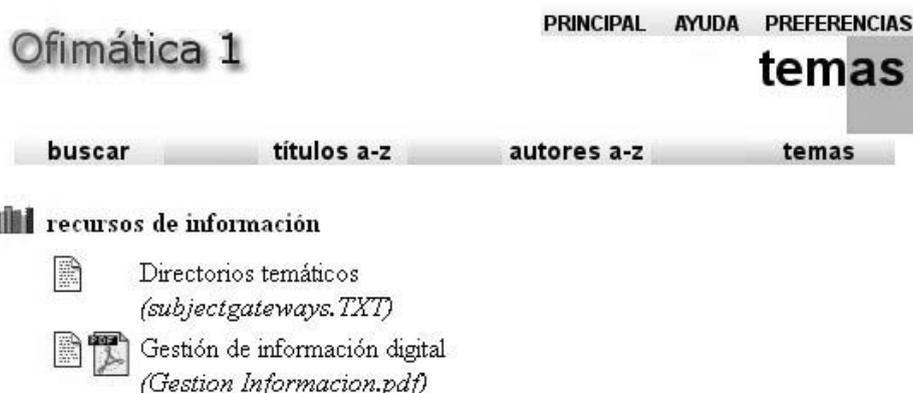


Figura 10.40. Despliegue de los documentos dentro de una categoría.

Sobre el ejemplo, es necesario considerar varias cuestiones. En primer lugar, los clasificadores AZ crean listas alfabéticas, y cuando el número de documentos supera el límite mínimo, muestra directamente un menú alfabético (A-B-C...-Z), desde el cual acceder a los documentos. En este ejemplo no aparece ya que el número de documentos incluido no alcanza ese umbral. En segundo lugar, la presentación del listado de documentos en cada clasificador no parece ayudar a su comprensión, ni resalta el metadato o metadatos relacionados con el mismo. El formateo de las listas de Greenstone, y de los documentos incluidos, que permite la personalización de la presentación, se tratará en el siguiente apartado. Por último, como se puede apreciar en el menú desplegable, no son los únicos clasificadores que pueden utilizarse para la exploración de una colección. En la Tabla 10.3 se resumen los clasificadores disponibles.

Tabla 10.3. Clasificadores disponibles en Greenstone

Clasificador	Indicaciones
List	El más sencillo, muestra una lista alfabética de documentos.
AZList	Muestra una lista alfabética de documentos, ordenados y agrupados alfabéticamente según un metadato dado. Añade menú de navegación propio.
AZCompactList	Muestra una lista alfabética de documentos, ordenados y agrupados alfabéticamente según un metadato dado. Permite mayor control de elementos por página y navegación interna. Añade menú de navegación propio.
DateList	Muestra una lista alfabética de documentos, ordenados y agrupados por fecha, atendiendo al metadato disponible. Puede organizar los documentos por años y por años y meses. Añade menú de navegación propio.
Hierarchy	Organiza los documentos de forma jerárquica, usando como criterio una clasificación externa, que debe incluirse en un fichero de texto. Un metadato se contrasta con el fichero que contiene la clasificación. Puede anidar los documentos en varios niveles.

Tabla 10.3. Clasificadores disponibles en Greenstone (*continuación*)

Clasificador	Indicaciones
HTML	Introduce una página web con un URL.
Phind	Analiza las frases de los documentos para crear un índice, mediante el cual localizar los términos en su contexto anterior y posterior de frase. Permite ejecutar búsquedas.
Collage	Clasificador para colecciones de imágenes, facilita la creación de una página tipo hoja de contactos.
xxxSectionxxx	Clasificadores que organizan los documentos atendiendo a las características de las secciones de los mismos. Para ello es necesario haber estructurado previamente los documentos.

10.8.7. FORMATOS DE PRESENTACIÓN

El paso final para la creación de la colección de documentos es la definición de los formatos de presentación de los listados de documentos, y de los documentos en sí mismos. Greenstone toma en consideración listas verticales, o VList, listas horizontales, o HList, listas por fechas, o DateList, y documentos. Se pueden definir formatos de presentación generales, o diferentes para cada clasificador, y para los resultados de las búsquedas. Si no se especifica a qué clasificador corresponde el formato, se aplicará a todas las listas. Los formatos se definen mediante una cadena de caracteres en la que se combinan funciones, componentes, variables y código HTML. Debe observar que las funciones mediante las cuales pueden formatearse los resultados de un clasificador dado comienzan por CL1...n. Esto quiere decir que CL1 corresponderá al primer clasificador de la lista disponible de clasificadores, CL2 al segundo, etc. Si no utiliza estas funciones CLx, el formato de las listas será igual para todos los clasificadores. La definición de formatos de presentación también se lleva a cabo en la pestaña **Diseño**.

1. Pulse sobre **Formatos** en el menú del panel izquierdo. Se muestran en el panel derecho las opciones de definición y configuración de los formatos. Borre las entradas de formatos ya existentes, y que Greenstone añada automáticamente al crear una colección, seleccionando la entrada correspondiente en **Comandos de formato asignados actualmente**, y pulsando el botón **Eliminar formato**.
2. Para definir un formato para los resultados de las búsquedas, en el diálogo **Controles de edición**, despliegue el menú **Elegir función**, y seleccione **Search**. En el menú desplegable **Componente afectado**, seleccione **Vlist**. De forma automática, en el cuadro **Cadena de formato HTML** se ha añadido un código que combina HTML con código de variables de Greenstone, que puede editarse y ajustarse según las necesidades.
3. Revise el código del formato. Podrá apreciar que está formado por etiquetas de código HTML, como <td>, variables de Greenstone, entre corchetes como [link], y condiciones, entre llaves, como {Or/If}. El formato predefinido establece una es-

estructura de tabla sencilla, en la cual se insertan, por ejemplo, un icono del documento, que actúa como enlace al mismo, mediante [link][icon][link]; el contenido del metadato título del documento, [dls.Title],[dc.Title],[ex.Title],Untitled, que entre llaves, y precedido del condicional {Or}, muestra el que esté disponible de todos ellos. Borre la primera línea de la cadena de formato. Pulse sobre el botón **Añadir** para asignar el formato. El resultado de este formato se muestra en la Figura 10.41. Observe que eliminando la primera línea del formato, se ha limitado el acceso del usuario sólo al documento original, evitando que pueda acceder al documento XML resultado del procesamiento, y que aparecía bajo el icono del documento de texto.

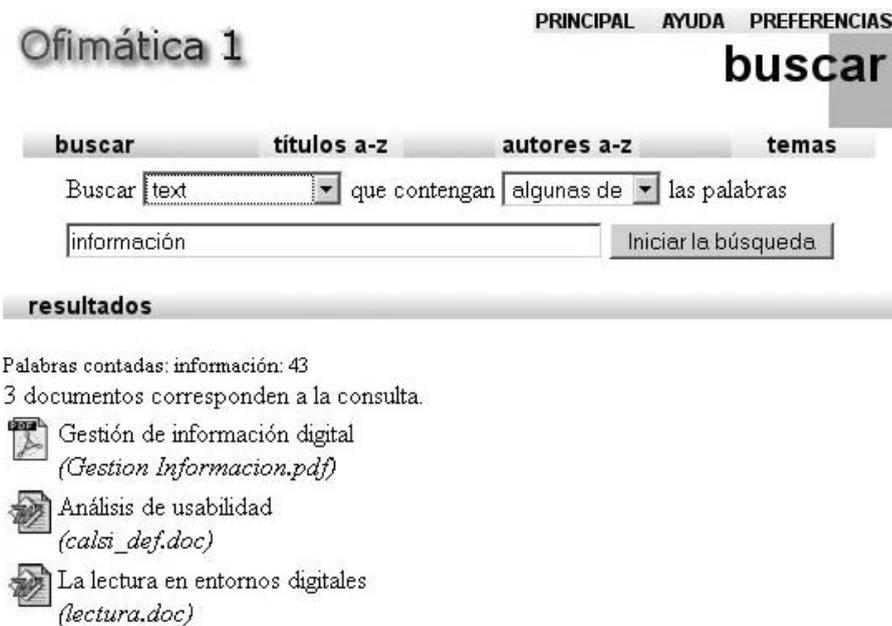


Figura 10.41. Presentación del formato SearchVList.

4. Para añadir un formato para listas verticales, que se aplique a cualquier clasificador, vaya de nuevo a **Controles de edición**, deje vacío el parámetro **Elegir función**, y seleccione **VList** del menú desplegable **Componente afectado**. Para modificar el código, borre la última línea desde {If} hasta }.
5. Para completar la información que ofrecerá el listado, se va añadir el metadato autor para cada documento. Sitúe el cursor en la última línea del código, y copie el código <td valign=top>. Acuda el menú desplegable **Variables**, y seleccione la variable [dc.Creator]. Pulse sobre el botón **Insertar**. Rodee la variable con el código HTML para itálica, <i></i>. Para terminar, cierre el código con </td> (Figura 10.42). Pulse sobre **Añadir formato**. Ahora, acuda a la colección para comprobar la visualización del listado (Figura 10.43).



Figura 10.42. Código del formato VList.

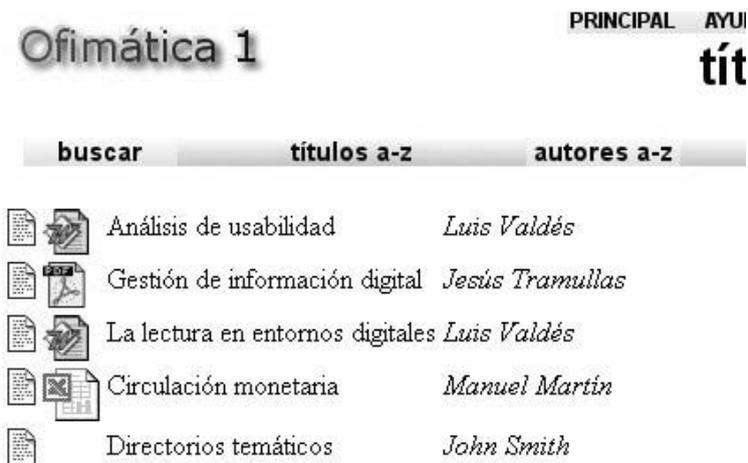


Figura 10.43. Presentación del formato VList.

6. Se pueden elaborar los formatos de presentación tanto como se desee. Por ejemplo, se va a presentar una presentación mejorada para el clasificador Temas. Pulse sobre **Elegir función**, y seleccione **CL1: AZCompactList...** Seleccione **VList** como **Componente afectado**. Ahora, va a proceder a mejorar el código que se muestra en **Cadena de formato HTML**. Para ello, introduzca el código de la Figura 10.44, escribiendo las etiquetas HTML necesarias, e incorporando las **Variables** de Greenstone del menú desplegable correspondiente.
7. Pulse **Añadir formato**. Para visualizar el resultado, vaya a la página web de la colección, y seleccione el clasificador **Temas** (Figura 10.45). El enlace al documento en XML se ha colocado en el título del documento, sustituyendo al icono. Se mantiene el icono de enlace al documento original. El autor se ha situado en cursiva

Controles de edición

Elegir función ▼

Componente afectado ▼

Cadena de formato HTML

```
<td valign=top>[ex.srclink][ex.srcicon][ex./srclink]</td>
<td valign=top bgcolor="#FFFFCC"><b>[link][dc.Title][/link]</b><br>
<i>[dc.Creator]</i></td>
```

Variables ▼

Figura 10.44. Código del formato para CLIVList.



Figura 10.45. Presentación del clasificador Temas con el formato asignado CLIVList.

debajo del título del documento, y todo ello se ha colocado en una celda con un fondo amarillo suave, para resaltar los datos del documento.

La configuración de los formatos de salida y de visualización de documentos ofrece opciones más avanzadas, para las cuales es necesario conocer con detalle los parámetros y los metadatos disponibles para cada colección. Para ajustar la presentación de los documentos individuales, el administrador deberá practicar con **DocumentHeading**, **DocumentContent** y **DocumentText**, que admiten similares funciones y atributos que las

listas de documentos. Por otra parte, estos formatos de presentación son independientes de las posibles modificaciones que se puedan hacer a la interfaz general de Greenstone. La modificación de la interfaz depende de la modificación de los ficheros de macro que controlan la configuración y los componentes de la aplicación. Como se puede imaginar, la explicación detallada del lenguaje de macros de Greenstone supera ampliamente los objetivos de este capítulo. Un tutorial detallado ha sido incluido entre las referencias de la Tabla 10.1, y también puede acudirse a la *Greenstone Developer's Guide*.