

Extracto del libro "Automatización de bibliotecas con SIABUC", por favor no haga copias de este material ni lo distribuya sin autorización expresa de los autores.

Si está interesado en adquirir un ejemplar de este libro puede obtener más información en http://siabuc.ucol.mx/libro_abs8.htm

Automatización de bibliotecas con SIABUC

**José Román Herrera Morales
Juan Luis Campos Salcedo
Evangelina Serrano Barreda
Luz María Pérez Santa Ana
Jorge Rafael Gutiérrez Pulido**



UNIVERSIDAD DE COLIMA

2004

Automatización de bibliotecas con SIABUC / J. Román Herrera

Morales... [et al.] -- México : Universidad de Colima, c2004.

416 p.: il., fot., tablas, gráficas; 25 cm.

Incluye bibliografía

ISBN 970-692-214-5.

1. Bibliotecas-Automatización 2. Información-Tecnología 3. Biblioteconomía
-Innovaciones técnicas 4. Bibliotecas-Procesamiento de datos

I. Campos Salcedo, Juan Luis II. Serrano Barreda, Evangelina

III. Pérez Santa Ana, Luz María IV. Gutiérrez Pulido, J. Rafael

LC Z678.9 .A4 A97

Primera Edición

? Universidad de Colima, 2004.

Dirección General de Servicios Bibliotecarios

Av. Gonzalo de Sandoval, 444

Colima, Colima, México

<http://siabuc.ucol.mx/>

email: siabuc@ucol.mx

ISBN 970-692-214-5

Impreso en México · Printed in Mexico

Diseño de Portada: Edgar Adalid Sánchez Ángulo

Consultor editorial: Alfredo Ameneiro Castro

Corrección de estilo: Patricia Sánchez Sandoval

Pero quizás en nuestras bibliotecas no tengamos proyectada la digitalización de todo el acervo por obvias razones, (limitaciones en recursos financieros, humanos y materiales, y por supuesto, la carencia de autorización de los autores de las obras para publicar su trabajo en internet a través de una biblioteca digital), pero si podríamos utilizar estas facilidades que brinda SIABUC y el formato MARC para extender los servicios que podemos darles a los usuarios, uno de estos puede ser, digitalizar únicamente las portadas y contraportadas de los libros, capturar un buen resumen en la ficha y si es posible digitalizar también la tabla de contenido de los libros. Sin duda si pudiéramos proporcionar esta información adicional a la de una simple ficha, sería de gran utilidad para los usuarios de nuestros sistemas de búsqueda en nuestras bibliotecas.

Ya sea que se decidan por implementar una biblioteca digital con recursos completamente digitalizados para visualizar documentales en video, leer el texto completo de los libros o revistas, o si simplemente se quiere brindar información adicional a las fichas referenciales mediante la inclusión de las portadas y tablas de contenido, el SIABUC es la herramienta apropiada que le permitirá realizar todas estas tareas sin necesidad de dismantelar la biblioteca ni comprometer el presupuesto financiero de varios años.

16.1.1 Características y servicios de la biblioteca digital con SIABUC8

En general las bibliotecas digitales pueden ser de diversos tipos y pueden ofrecer varios servicios como los ya mencionados en el apartado 14.1.1 del capítulo 14, pero para el caso de SIABUC, comentaremos cuales son los servicios básicos con los que podrá contar al implementar su biblioteca digital:

Interfaz amigable y fácil de usar para la captura de las fichas de cada uno de los recursos digitalizados que conforman el acervo de la biblioteca digital

Utilización del mismo tipo de metadatos para la descripción de los recursos de información: el conjunto de etiquetas MARC

Personalización completa de la interfaz de consultas y de la visualización de los resultados, con la facilidad de acceder a los recursos digitales mediante simples ligas a los URLs.

Otros servicios que también se pueden incluir son las alertas o reportes de novedades bibliográficas que se generan en formato HTML desde el módulo de préstamos (véase sección 7.3 del Capítulo 7)

16.1.2 Requerimientos para implementar la biblioteca digital con SIABUC8

Recursos humanos

Se requiere de personal técnico para realizar las labores de digitalización y de especialistas para la administración de la red y configuración de servidores. Por supuesto es indispensable el personal bibliotecario que registrará los metadatos de cada uno de los recursos de información.

Información a digitalizar

La riqueza de una biblioteca digital es su contenido, para lo cual se tiene que identificar y seleccionar previamente cual es el material que va a ser digitalizado. Puede tratarse de alguna colección especial que esté libre de derechos, o bien, que se cuente con el permiso de él, o los autores.

Infraestructura tecnológica

Para la digitalización de información se requerirán de computadoras, cámaras digitales o escáneres y software de procesamiento de textos o imágenes y de OCR (Optical Character Recognition).

Para el almacenamiento de los archivos digitales que conforman el acervo de la Biblioteca Digital, se requiere contar con una computadora servidor con suficiente capacidad de disco duro y velocidad de procesamiento, también se requerirá de una buena configuración de red y acceso a internet. Debe tomarse en cuenta que el texto completo y las imágenes digitalizadas ocupan un espacio considerablemente mayor en disco duro que las simples fichas.

16.1.3 Metodología para implementar la biblioteca digital con SIABUC8

1. Selección del material a digitalizar. Identificar cuales son los textos, material multimedia y/o ligas a recursos electrónicos que conformarán el acervo de la biblioteca digital.
2. Digitalización de la información. Para cada tipo de información se deberán aplicar diferentes técnicas de digitalización y almacenarse en un servidor de web utilizando una nomenclatura predefinida para facilitar su acceso y recuperación.
3. Catalogación. Se utilizará el módulo de análisis para la captura de las fichas o “metadatos”, usando algún formato personalizado e incluyendo la etiqueta 856 con la que se podrán relacionar una o múltiples ligas (URLs) a recursos digitalizados para cada ficha.
4. Indexado y actualización de la base de datos. Habrá que actualizar tanto los archivos de los recursos digitales como las bases de datos en el servidor Web en forma periódica.
5. Configuración del servidor de web. Prácticamente el único requisito es que el servidor web funcione en plataformas Windows y pueda ejecutar aplicaciones tipo CGI. También se tendrá que verificar lo relativo a permisos de escritura y lectura a las carpetas donde se alojan el programa Web_S8.exe y las bases de datos (Similar a como se montan los catálogos en web, véase el capítulo 14 para más detalles).
6. Implementación de la interfaz web para la consulta del acervo de la biblioteca digital. Esta interfaz se genera y personaliza de la misma forma en que se montan los catálogos en web con SIABUC8
7. Personalización del formato de salida. Esta parte se hace adaptando el archivo de configuraciones CFG, definiendo formatos de salida propios e incluyendo la etiqueta 856 para que se genere una liga HTML, la cual nos dará acceso al recurso digital directamente desde el navegador de internet o aplicación asociada según el tipo de archivo.
8. Mantenimiento y actualización periódica del sitio web de la Biblioteca Digital. Esto incluye mantener los archivos de índices actualizados así como la verificación de los URLs para acceder a los recursos digitales, sobre todo si se tienen referencias a recursos electrónicos donde la información se hospeda fuera de los servidores de la institución; todo esto sin olvidar los correspondientes respaldos de toda la información.

16.2 Manos a la obra

16.2.1 Tipos de recursos digitales que se pueden asociar a fichas referenciales

Mencionamos que se tienen que identificar que tipo de información se va a digitalizar y que productos digitales vamos a obtener de la misma; por ejemplo, de un libro, podríamos tener disponibles su portada, su resumen y su texto completo.

SIABUC8 está diseñado para asociar varios tipos de recursos digitales a una misma ficha, básicamente son 5:

- ?? Las portadas del material
- ?? Las tablas de contenido de la publicación
- ?? Cualquier imagen relacionada con el recurso
- ?? El documento digitalizado del recurso en cualquier formato que sea reconocible a través de un navegador (DOC, PDF, HTML, TXT, etc.)
- ?? Cualquier referencia a algún sitio de Internet (cualquier URL)

16.2.2 La digitalización de los materiales o recursos de información

Una vez identificado el recurso o documento a digitalizar se tendrán que aplicar diversas técnicas y evaluar que tipo de formato es el más conveniente para el tipo de información del que se trate. De manera muy general y sin que sea una recomendación que aplique en todos los casos, podríamos sugerir los siguientes formatos para representar los diferentes tipos de documentos:

Para documentos de texto completo, utilizar formatos **PDF** y **HTML**, en el caso de los documentos PDF (Portable Document Format), éstos se han convertido en casi un estándar *de facto*, para la publicación de documentos con gran riqueza en estilo y contenido, con PDF se puede presentar un documento digital que conserva las proporciones y apariencia que tiene el documento original en papel, además brinda facilidades de impresión y protección contra modificaciones del documento. También, y a diferencia del HTML se conserva como un único archivo la información que contiene todo el documento como texto, gráficas e imágenes, lo cual lo hace más fácil de administrar y transportar a través de la red.

Para el caso de documentos de audio, se ha vuelto muy popular el formato **MP3** (sí, el de las canciones), resultando que los archivos digitalizados en este formato conserven gran nitidez y calidad y además son muy ligeros.

Para archivos de imagen en movimiento o propiamente videos, los formatos reinantes actualmente son el **AVI** (audio y video interleaved), **RM** (Real Media) donde se pueden lograr archivos de video digital con la calidad similar a la de un DVD y con un tamaño de archivo muy compacto si se compara con los formatos nativos de video.

Para el caso de las imágenes las más comunes son **JPG** (como para fotos a color), el formato **GIF** y más recientemente el formato **PNG**.

No nos adentraremos mucho en este tema sobre técnicas de digitalización y formatos de archivos, para enfocarnos propiamente a los detalles de la implementación de la biblioteca digital. En caso de necesitar más información puede visitar el tutorial en línea “Llevando la Teoría a la Práctica” de la Biblioteca de la Universidad de Cornell de Estados Unidos en <http://www.library.cornell.edu/preservation/tutorial-spanish/index.html>

16.2.3 Recomendaciones para asignar el nombre a los archivos de los recursos digitales

Es conveniente mencionar algunas recomendaciones para asignar el nombre a cada uno de los archivos resultantes del proceso de digitalización, de tal forma que sean fácilmente identificables y asociables a la ficha que describe su contenido. Se sugiere la siguiente nomenclatura:

Utilizar nombres de archivo con 10 caracteres más su extensión del tipo de archivo, desglosándose de la siguiente manera:

6 dígitos para el número de la ficha, complementados al inicio con ceros.

4 caracteres alfanuméricos para indicar el tipo de recurso, tamaño y / o consecutivo:

Los 3 primeros caracteres definen el tipo de recurso, pueden ser:

- ?? PTA para portadas
- ?? TOC para tablas de contenido
- ?? IMG para imágenes
- ?? DOC para documentos digitales
- ?? URL para sitios Web

Para el caso 4to. carácter, se usará:

- ?? “g”, si la imagen de una portada es grande: 500 píxeles de ancho y alto proporcional
- ?? “m”, si la imagen de una portada es una vista miniatura: 100 píxeles de ancho y alto proporcional
- ?? “1”, número consecutivo del 1 al 9 para recursos que no sean portadas, o
- ?? “_”, un guión bajo si no aplicara ninguno de los casos anteriores.

Por último, la extensión correspondiente al formato del archivo (.jpg, .gif, .pdf, etc.)

Ejemplos:

Si se trata de una imagen que representa la vista en miniatura de la portada de un libro capturado en la ficha 1 de SIABUC y digitalizada en formato JPG, el nombre del archivo será: 000001PTAm.jpg

Si se tiene un archivo PDF que representa el capítulo 1 del libro registrado en la ficha 412, el nombre que se asignaría a este archivo sería: 000412DOC1.pdf

Si fuera la tabla del contenido digitalizada en PDF del mismo libro, sería: 000412TOC_.pdf

Si fuera la imagen de la portada del mismo libro en tamaño grande y formato PNG, sería: 000412PTAg.png

Si fuera la segunda gráfica, una foto, un diagrama o un esquema de ese documento de la misma ficha 412 digitalizado en formato GIF, el nombre del archivo sería: 000412IMG2.gif

Los archivos digitalizados y apropiadamente nombrados se deberán de ubicar en carpetas públicas a partir de las cuales se puedan acceder desde un servidor de web, o de una red de área local.

Por ejemplo, la imagen 0013520PTAm.jpg se almacenó en la carpeta “bdigital\docs” del servidor de web, por lo tanto para recuperar esta información a través de la biblioteca digital, su URL será:

<http://miservidor.com/bdigital/docs/0013520PTAm.jpg>

Y esta deberá ser, precisamente, la información que se registrará en la etiqueta 856 de la ficha correspondiente, con algunas consideraciones que se tratarán más adelante.

16.2.4 La captura de los metadatos

Para el proceso de captura de los metadatos de la biblioteca digital o dicho de manera coloquial, para el proceso de catalogación de las fichas de la biblioteca digital, simplemente se utiliza el módulo de análisis de SIABUC, donde a partir de la actualización 3.28 incluye 4 formatos de captura predeterminados (libros, videos, material de audio y tesis) y en cualquiera de estos formatos ya se incluye de manera predeterminada la etiqueta 856.

Es esta etiqueta 856 la que de manera “mágica” hace posible que un sistema de administración de información bibliográfica referencial como lo es SIABUC, pueda utilizarse para permitir el acceso a información de texto completo o de recursos multimedia. La etiqueta 856⁵ del formato MARC llamada “Electronic Location and Access” se utiliza para registrar el camino o ruta de acceso para poder recuperar recursos de información de forma electrónica (en formato digital) que están disponibles para cada una de las fichas o títulos que se están capturando en Análisis y que formarán nuestro acervo digital.

A continuación describiremos el proceso de captura de la ficha de un libro, la cual además de registrarse los campos tradicionales como título, autor, mención de edición, temas, ISBN, etc, se le incluirá la información para acceder al recurso digital que representa la versión en texto completo de ese libro.

Primero, se puede definir un formato personalizado de captura para las fichas de la Biblioteca digital, el cual debe incluir de manera obligatoria la etiqueta 856. Esto se hace desde el “Menú de Utilerías” del módulo de Análisis en la sección de “Formatos”.

⁵ Etiqueta 856. <http://www.loc.gov/marc/bibliographic/ecbdhold.html>
<http://www.oclc.org/bibformats/en/8XX/856.shtm>
Guía para utilización del campo 856, Traducción en español
http://cuib.laborales.unam.mx/~felipe/cat_Internet/campo856.htm

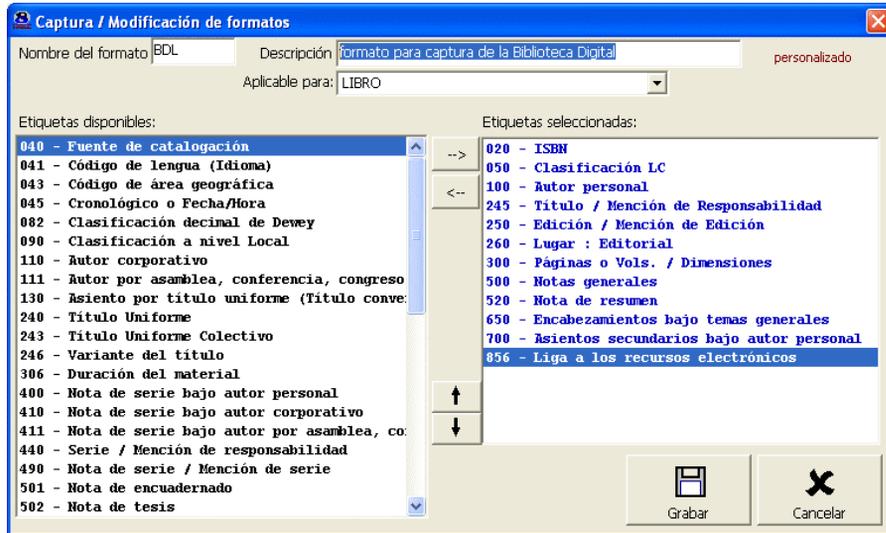


Figura 16.1 Definición de un formato personalizado para la captura de fichas de la biblioteca digital

Una vez definido este formato de captura se debe “activar el formato” para que podamos utilizarlo en la captura y modificación de fichas.

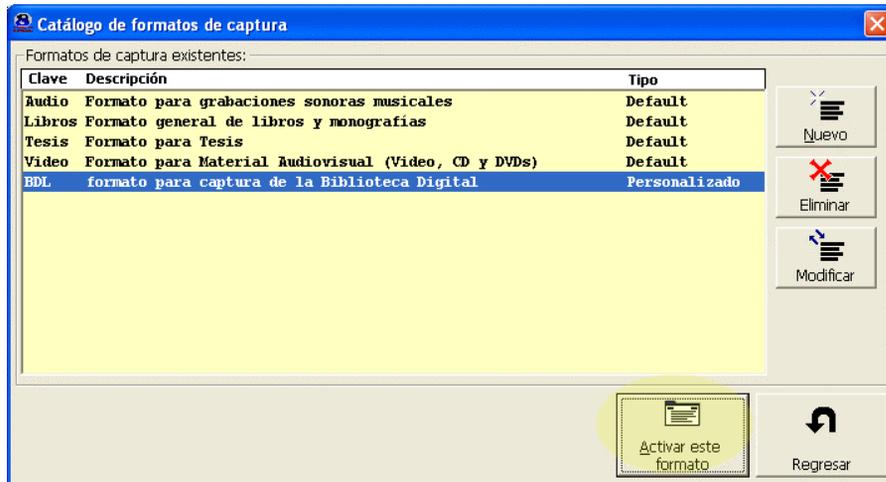


Figura 16.2 Activación del formato de captura de la biblioteca digital

Se procede a la captura de la ficha introduciendo los datos correspondientes para cada etiqueta MARC y la liga del documento digital en la etiqueta 856.

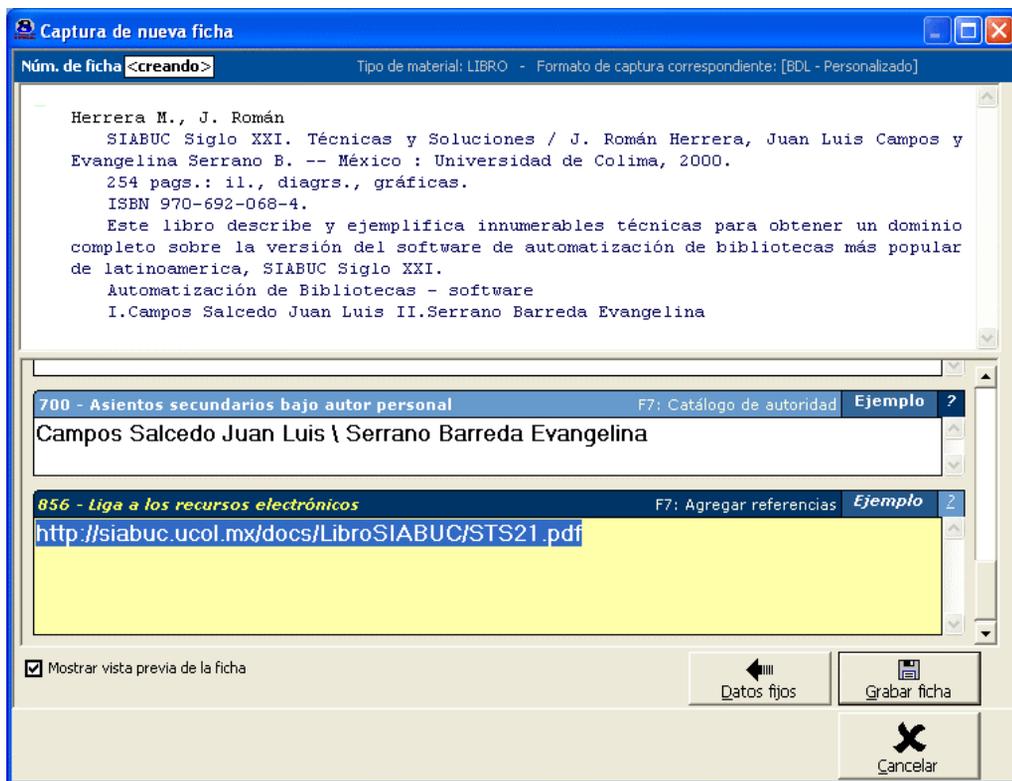


Figura 16.3 Ejemplo de captura de una ficha incluyendo un URL en la etiqueta 856.

Se guarda la ficha se capturan los datos de los ejemplares disponibles, aunque en este caso, como se trata de recursos digitales, pudiera ser que no se tengan físicamente en papel los ejemplares, pero para efectos del indizado de SIABUC, siempre se requiere que cada ficha tenga al menos un número de adquisición pues de lo contrario esa ficha no sería incluida como parte de las fichas a indizar. Posteriormente, se puede explorar la ficha desde la opción de consultas del mismo módulo de análisis, para tener un previo de lo que podría ser la forma de mostrar los resultados en la interfaz web de la biblioteca digital

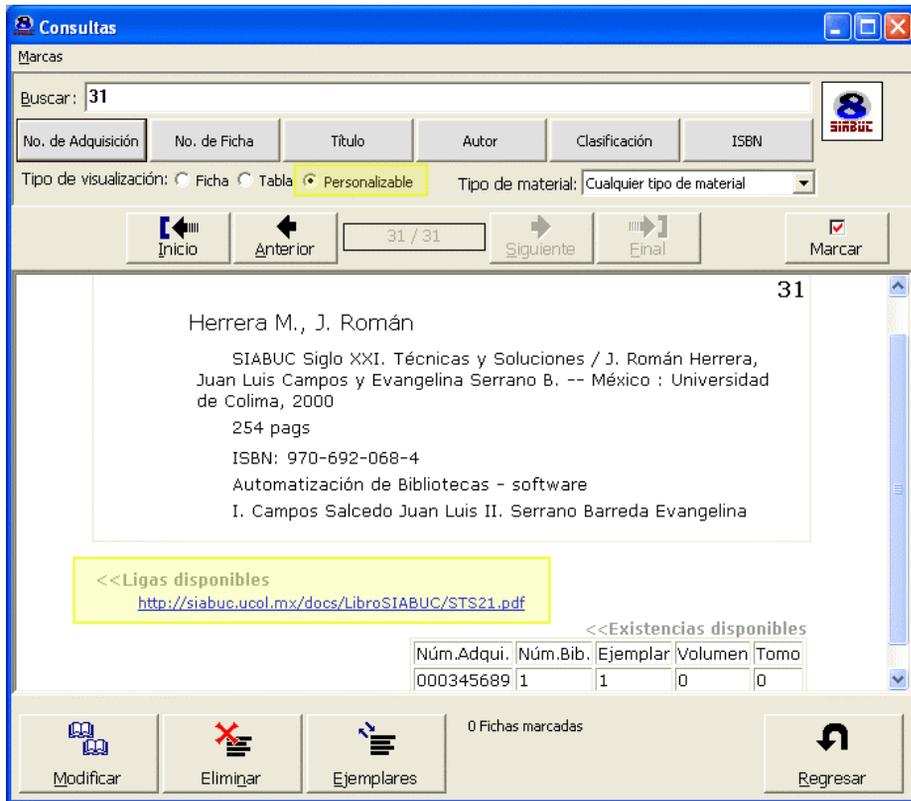


Figura 16.4 Interfaz de consultas de análisis donde se despliegan las ligas capturadas usando el formato de visualización personalizado.

Para poder desplegar la información de los recursos digitales en la ficha, es necesario seleccionar el tipo “personalizable” en los formatos de visualización de consultas, también hay que notar que las ligas disponibles aparecerán en la parte inferior de la ficha y desde ahí se puede abrir ese documento haciendo click sobre el mismo, esto funciona exactamente igual que en cualquier página de internet.

En la descripción del ejemplo anterior, solamente se utilizó un único recurso asociado a la ficha, mismo que normalmente sería el texto completo en un solo archivo. Abirá ocasiones en que, por cuestiones de organización o de velocidad de transmisión, se dividida un libro completo en varios archivos, de tal forma que se puedan recuperar individualmente por secciones o por capítulos, etc. A continuación describiremos el mismo ejemplo del libro de SIABUC Siglo XXI pero en el cual se ha digitalizado y generado varios archivos, incluyéndose las portadas, contraportadas, tablas de contenido y los diferentes capítulos del libro, para ello se aplicarán las recomendaciones descritas en la sección 16.2.3 para la asignación de los nombres a los archivos generados; y además se describirá una herramienta auxiliar que se incluye dentro de el módulo de análisis: “*El asistente para agregar referencias de recursos digitales?*”.

Para utilizar este asistente, debemos estar en la opción de “Captura / Modificación de Fichas” y bastará con que pulsemos la tecla “F7”, estando sobre la posición de captura de la etiqueta 856 (tal como si estuviéramos llamando al catalogo de autoridad).

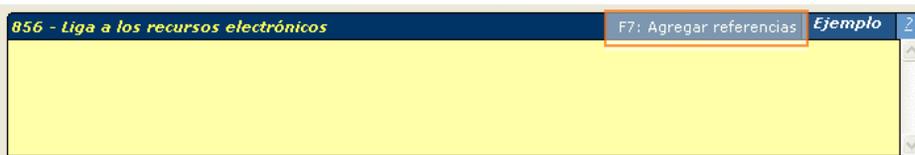


Figura 16.5 Opción para abrir el asistente para captura de referencias en la etiqueta 856

Con el asistente podemos asociar múltiples recursos digitales y agregar cierta información clave que utiliza internamente el SIABUC para identificar el tipo de recurso digital del que se trata, además de que también podemos previsualizar cada una de estas ligas para verificar que estén bien escritas y que son accesibles a través de internet.

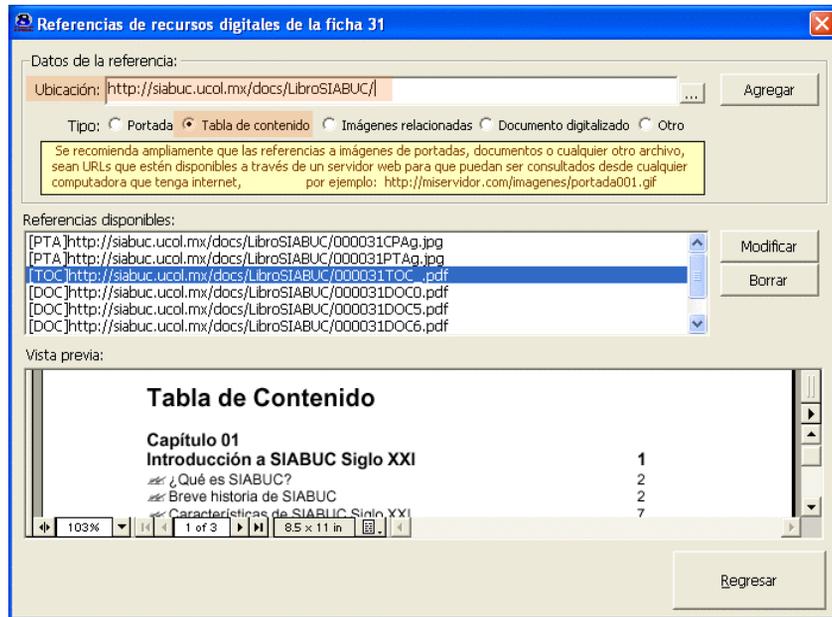


Figura 16.6 Interfaz del asistente de Referencias de Recursos Digitales

Al utilizar el asistente para la captura de “Referencias de recursos digitales”, SIABUC agregará automáticamente cierta información adicional a cada URL, anteponiéndole un prefijo de texto que indica el tipo de recurso, más una descripción del mismo, por ejemplo:

?? [IMG{descripción de la imagen o figura}]

?? [DOC{descripción del documento}]

?? para el caso de portadas y tablas de contenido la descripción del recurso se omite quedando [PTA{}}] ó [TOC{}}], tal y como se muestra en la siguiente figura:

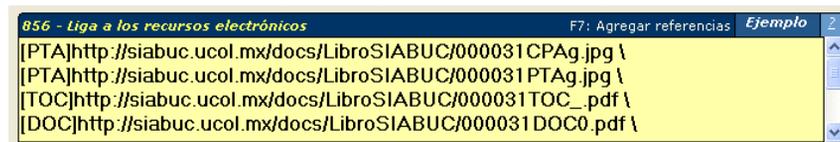


Figura 16.7 Ejemplo de referencias capturadas en la etiqueta 856 en las que se incluyen prefijos del tipo de documento

Una vez capturadas todas las referencias a recursos digitales, podemos visualizar la ficha catalográfica en pantalla utilizando nuevamente el formato de visualización “personalizable” .

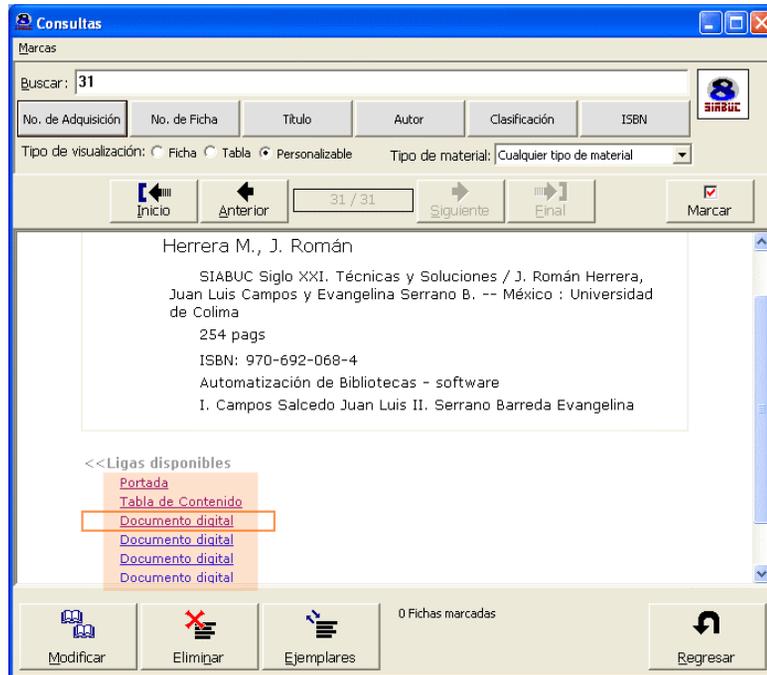


Figura 16.8 Ejemplo de visualización de una ficha en Análisis, que incluye varias referencias digitales y de diferente tipo.

Desde la misma interfaz de consulta del módulo de análisis podemos explorar las ligas disponibles y visualizar los documentos referenciados, en este caso, para cada recurso digital que se desee explorar se abre en una nueva ventana del explorador de internet que tenga predeterminado.

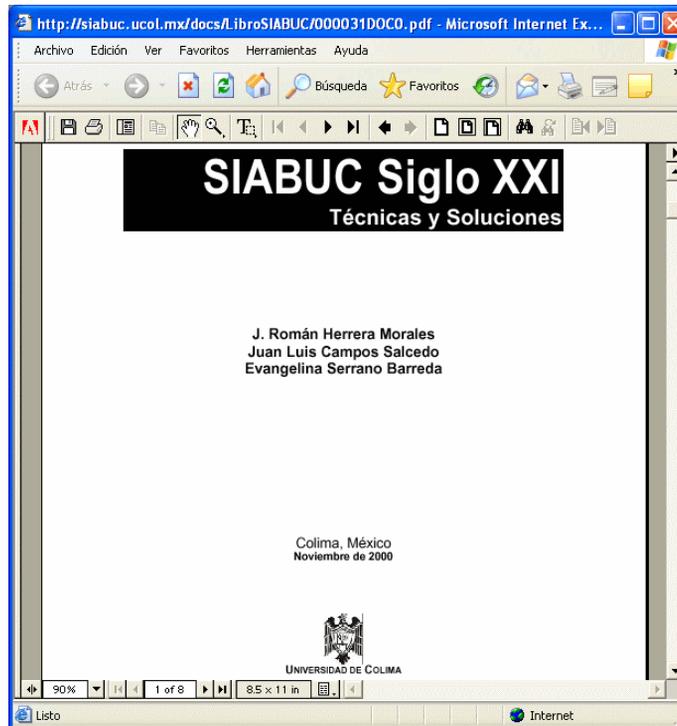


Figura 16.9 Visualización de un documento PDF abierto en una ventana independiente del navegador.

Cabe mencionar que un comportamiento y apariencia similar tendrá el módulo de consultas para los usuarios, siempre y cuando, se utilice la opción de formato “personalizable” de visualización.

16.2.5 La creación de la interfaz de la biblioteca digital

Es importante recalcar que la implementación de la interfaz de la biblioteca digital, el montar los catálogos en web y tareas relacionadas, son actividades que requieren del apoyo de personal con perfil informático, ó con cierto nivel de experiencia en el manejo de servidores y páginas web.

Una vez que ya se tienen digitalizados cierto número de documentos que formarán el acervo de la biblioteca digital, y que ya se tiene capturadas sus fichas o metadatos, el último paso es montar la interfaz de la biblioteca digital.

A partir de la actualización 3.3 de SIABUC8 publicada en noviembre de 2004, se distribuye un paquete con un conjunto de archivos llamado *WebPaqBDL.zip* que contiene una interfaz HTML prefabricada lista para montar la Biblioteca Digital, bastará con descompactar estos archivos y copiarlos en el servidor de web y hacer unas simples adecuaciones a los archivos de configuración para ya tener funcionando nuestra propia biblioteca digital.

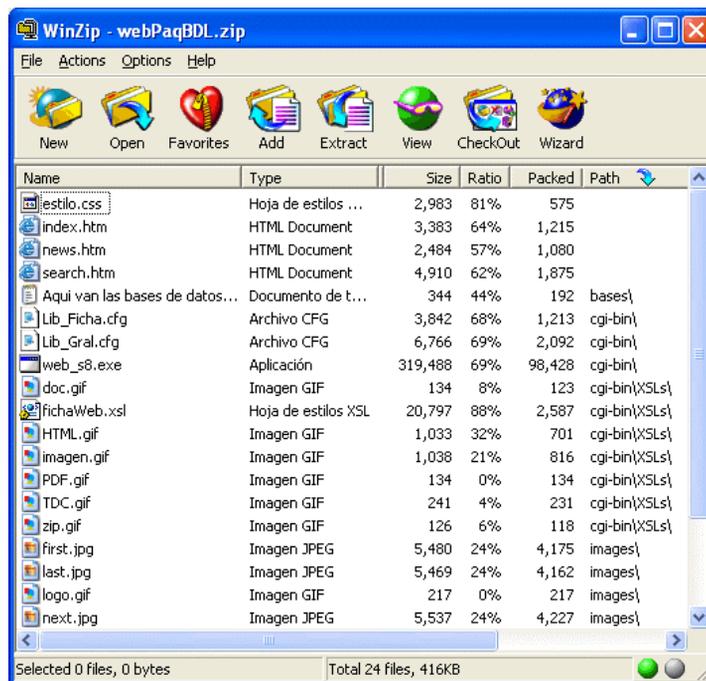


Figura 16.10 Contenido del paquete webPaqBDL.zip