

Un interfaz hipertexto para un sistema de gestión de tesauros

Juan Antonio Pastor Sánchez

Tomás Saorín Pérez

Grupo de Investigación en Tecnologías de la Información.

Universidad de Murcia

0.1. Resumen

El objetivo de este trabajo es mostrar las posibilidades de creación de un interfaz básico para sistemas de hiperdocumentos partiendo de Tesauros generados con un sistema Gestor de Construcción de Tesauros. Para ello se han construido varias aplicaciones interrelacionadas entre sí, partiendo de un modelo de datos que ha permitido la creación de un esquema conceptual para la construcción de Tesauros normalizados a nivel semántico-estructural, y se ha desarrollado e implementado un Sistema Gestor de Construcción de Tesauros, para finalmente elaborar un interfaz de navegación de Tesauros en hipertexto.

Los principales resultados han sido un Sistema Informático bajo entorno windows que conjuga la potencia de un Sistema Gestor de Construcción de Tesauros en Bases de Datos Relacionales, con la flexibilidad del interfaz de navegación de los mismos realizado mediante el Sistema de Hipertexto GUIDE. Se presentan dos modelos de interfaz bajo hipertexto para la gestión de Tesauros. Como conclusiones se puede afirmar que el hipertexto se adapta perfectamente a la gestión de lenguajes documentales tipo Tesauro, y que estamos ante una nueva perspectiva en el almacenamiento y recuperación de la información documental, lo cual nos lleva a una nueva orientación en los procesos de organización del conocimiento. (Autor)

Palabras clave: Sistemas Gestores de Tesauros. Hipertexto. Modelo de datos. Diseño de interfaces.

0.2. Abstract

This paper shows the possibilities that exist to create a basic hypertext interface on a Thesaurus Management System. The project has been built in three stages: a) conceptual data model for the thesaurus; b) development of the management tools, that is, the thesaurus management system; and c) the hypertext

interfaces. The system has been developed on Windows and combines the power of relational DBMS and the navigational flexibility of Guide, a hypertext shell. Two models of interfaces are proposed. It can be concluded that hypertext interfaces improve thesaurus management, and open a new path for knowledge organization research. (Author)

Keywords: Thesaurus Management Systems. Hypertext. Data Models. Interface design.

1. Introducción al Esquema Conceptual para la Construcción de Tesoros

Hemos introducido en anteriores trabajos la idea del diseño a nivel abstracto de un tesoro utilizando un Modelo de Datos, en un intento de proveer una herramienta para construir Tesoros normalizados a nivel semántico-conceptual. El producto de estos trabajos fue la creación de un Esquema Conceptual para la Construcción de Tesoros.

Dicho esquema conceptual se realizó mediante el uso del Modelo de datos Entidad-Categoría-Relación que viene a ampliar el Modelo Entidad-Relación que inicialmente propuso Chen.

Este esquema es básico en el diseño, desarrollo e implementación de nuestro Sistema Gestor para la Automatización en la Construcción de Tesoros (el cual abreviaremos como SGAT).

2. Sistema Gestor para la Automatización en la Construcción de Tesoros: SGAT

SGAT es una herramienta informática diseñada con el generador de aplicaciones DBfast 1.7 en entorno Windows, que permite construir Tesoros, creando un glosario de Términos, de Idiomas y de Relaciones entre los mismos, al tiempo que comprueba la integridad de la red semántica del Tesoro en construcción.

Viene a ocupar un hueco en el mundo de las aplicaciones informáticas en el mundo de los lenguajes documentales, pues si bien existen sistemas de recuperación de información que llevan incorporado un módulo gestor de la consulta mediante Tesoros, el proceso de construcción de los mismos se encuentra en fase artesanal o semiautomatizada. El sistema presentado permite automatizar casi todo el proceso de construcción de un Tesoro, controlado de modo interno la coherencia de las relaciones y el mantenimiento de las mismas.

Podemos resumir las características de SGAT en los siguientes puntos:

- Sistema de estructura modular, fácilmente integrable en otras aplicaciones documentales.

- Distinción de los distintos términos del tesauro estableciéndose una tipología léxica: Descriptores, No-Descriptores, Campos Semánticos y Notas Explicativas.
- Creación de las relaciones semánticas entre términos del tesauro y sus inversas.
- Control idiomático.
- Comprobación de la coherencia de las relaciones semánticas.
- Entorno navegacional durante la construcción y consulta de tesauros.
- Listado de tesauros construidos.

En los gráficos 1. y 2. podemos ver un volcado de pantalla que ilustra un ejemplo del uso del SGAT, centrándonos en la construcción de relaciones de un Tesauro provisional sobre Informática y del Sistema Navegacional que permite su consulta.

Observamos como tras establecer el campo semántico (Aplicaciones), en el cual vamos a establecer una relación, y una vez escogido el término origen de la misma (Gestión de Formularios), podremos seleccionar el tipo de relación en cuestión, de entre un menú controlado, en el que solo se presentan aquellas

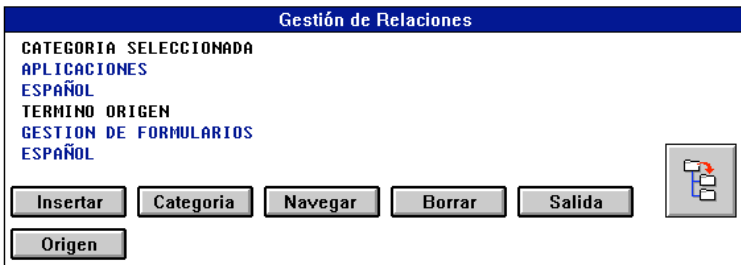
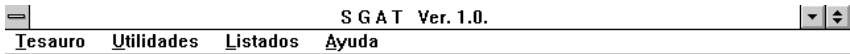


Fig. 1

Tipo de Relación
DOM: Pertenencia
EI: Equiv. Idiomática
NE: Nota Explicativa
UP: Término Equivalente
TG: Término Genérico
TE: Término Específico
TA: Término Asociado
USE: Término Preferente
INV: Inverso

LISTA DE TERMINOS
Página Anterior <<
Página Siguiente >>
EMULACION DE TERMINAL
FINANZAS
GESTION
GESTORES DE BASES DE DATOS
GESTORES DE BASES DE DATOS EN RED
GESTORES DE BASES DE DATOS JERARQUICAS
GESTORES DE BASES DE DATOS RELACIONALES
GRAFICOS
GRAFICOS CIENTIFICOS
GRAFICOS DE GESTION COMERCIAL
GRAFICOS DE PRESENTACION
GRAFICOS GEOGRAFICOS
HERRAMIENTAS DE DESARROLLO
HERRAMIENTAS DE DESARROLLO DE HIPERTEXTO
HERRAMIENTAS DE DESARROLLO DE SISTEMAS EXPERTOS
HIPERTEXTO
HOJAS DE CALCULO
INTERCAMBIO DE FORMANTOS GRAFICOS
LENGUAJES (SISTEMAS EXPERTOS)
MULTIMEDIA

Fig. 2 y 3

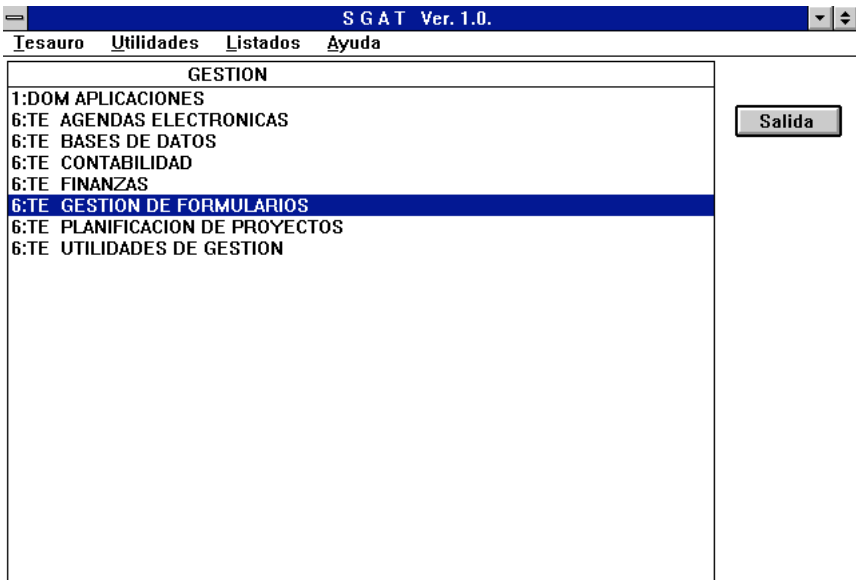


Fig. 4

posibles dada la naturaleza del término.

A continuación, en el gráfico 3. establecemos el término destino (*gestión*) y podremos observar en la consulta del tesaurus en construcción como se ha creado esta relación (gráfico 4.).”Picando” con el ratón sobre cualquiera de estas relaciones, pasaremos a observar una estructura similar pero esta vez la referida a la del término cuya relación hemos seleccionado, creándose pues un entorno de navegación en la consulta del Tesaurus.

3. Interfaz de Navegación del SGAT en Hipertexto.

Sin embargo el entorno SGAT no está diseñado para su uso en la indización de documentos, solamente es una herramienta para construir Tesaurus y sigue siendo necesario un interfaz de navegación y consulta que facilite la indización de documentos dentro de una aplicación documental.

La aplicación documental que hemos escogido para la creación de un interfaz de navegación de SGAT es el sistema de hipertexto *Guide*, ya que ofrece una serie de relaciones de distinta tipología entre objetos ideal para construir un interfaz de navegación y consulta de los Tesaurus generados por SGAT, y para crear

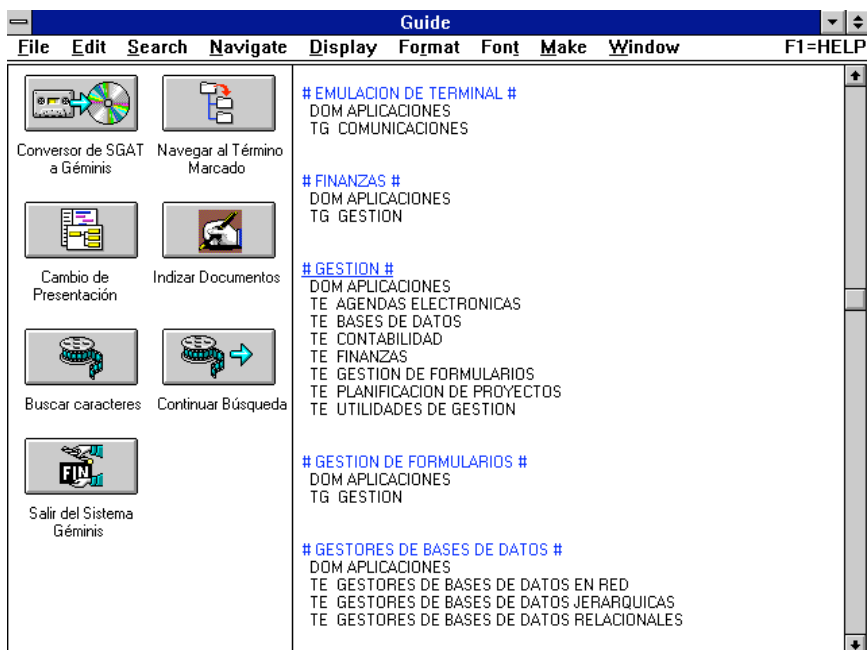


Fig. 5

un entorno de almacenamiento y recuperación de información documental.

GUIDE es un sistema de hipertexto bajo entorno WINDOWS que permite crear documentos con posibilidad de insertar varias páginas (o *frames*) en un mismo documento y almacenar información textual o gráfica que permite ser estructurada en bloques que se relacionan entre sí.

De este modo y usando el sistema *Guide* hemos creado el *interfaz Géminis* para la navegación y consulta de Tesoros construidos con SGAT. La integración entre ambos programas se produce a través de la opción de *convertor de SGAT a Géminis*, que permite convertir automáticamente un listado ASCII de SGAT a una estructura de Hipertexto, creando los objetos y el sistema de relaciones (gráfico 5).

Los dos aspectos más a destacar son la navegación entre los términos y la combinación de presentaciones:

- Mediante la opción de *navegar al termino marcado* nos moveremos al

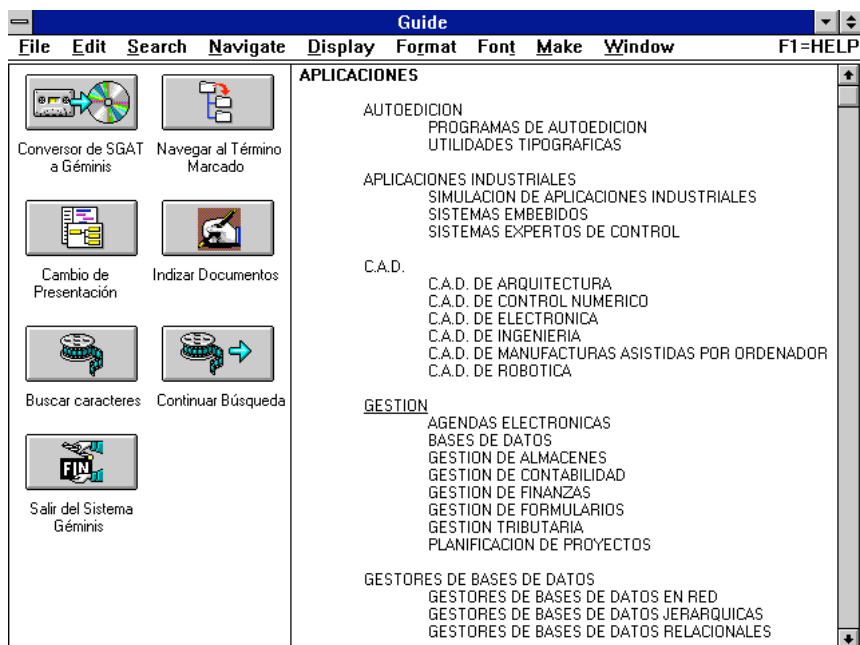


Fig. 6

encabezamiento de termino cuya relación marcamos, es decir si situamos el puntero de escritura sobre *el termino específico contabilidad* y tras “Pinchar” en el botón de la opción de *navegar al termino marcado* nos desplazaremos al encabezamiento *contabilidad* y podremos visualizar los términos con los que se relaciona pudiendo repetir la operación de navegación cuantas veces deseemos.

- Mediante la opción de *cambio de presentación* se vasculará entre una presentación alfabética y otra sistemática conservándose localizado el término al que se ha navegado con anterioridad, permitiendo así una mejor comprensión del contexto de los términos.

Estas son las opciones básicas del sistema, el resto de opciones tales como *buscar caracteres* y *continuar búsqueda* permite la búsqueda de una cadena de caracteres entre el texto del Tesauro.

4. Otro interfaz de navegación para tesauros en hipertexto

Una vez vista la interconexión realizada entre un sistema de construcción de

Tesauros, que controla la coherencia de las relaciones a nivel semántico-estructural (no de contenido semántico) y un sistema de hipertexto que sirve como interfaz para su consulta, presentaremos otro Sistema de Hipertexto para la Gestión de Tesauros, que se plantea desde una perspectiva distinta.

Se ha construido un modelo para la implementación en hipertexto de Tesauros, que ofrece una novedad en cuanto a la presentación: Si en los Tesauros tradicionales en papel encontramos generalmente las presentaciones:

- *Alfabética Completa*: En la que aparecen descriptores y no-descriptores seguidos de una lista de sus relaciones genéricas, específicas, asociativas, de campo semántico, de sustitución o de nota.
- *Jerárquicas*: Pretende presentar una visión general de las relaciones jerárquicas que se establecen entre los términos. Se suele presentar de dos maneras: Sistemáticamente a partir de los campos semánticos y descendiendo progresivamente de nivel de profundidad, presenta la cadena jerárquica completa de todos los descriptores, o bien simplemente Temáticamente indicando los descriptores que pertenecen a un determinado campo semántico.
- *Índices permutados*: Que permiten el acceso a través de las diversas palabras significativas de los descriptores.

El Sistema de Hipertexto que se presenta consigue la fusión de todas ellas generando un único modelo de presentación, al que hemos convenido en llamar *presentación sistemática completa*, como se puede apreciar en los gráficos 7. y 8.

Las características más importantes del Sistema, denominado *interfaz cocktail* son:

1. *Profundización selectiva*: Sólo se profundiza en aquella rama del tesoro que interesa, mediante sucesivas expansiones, manteniéndose las restantes del Tesoro como contexto.
2. *Desplazamientos transversales*: Permite con una simple pulsación sobre un término relacionado, navegar hasta la posición del tesoro en la que se encuentra el término.
3. *Reorganizar*: Después de navegar, expandir y consultar múltiples términos, se puede producir una cierta desorientación. Esta opción del menú de control devuelve el tesoro a su primer nivel de expansión, donde solo se visualizan los campos semánticos y sus *topterms* correspondientes.
4. *Retroceder*: Esta es una de las opciones más potentes en cualquier sistema de navegación. Mediante ella se puede recorrer en sentido inverso el trayecto seguido hasta el momento (almacena hasta 100 movimientos).

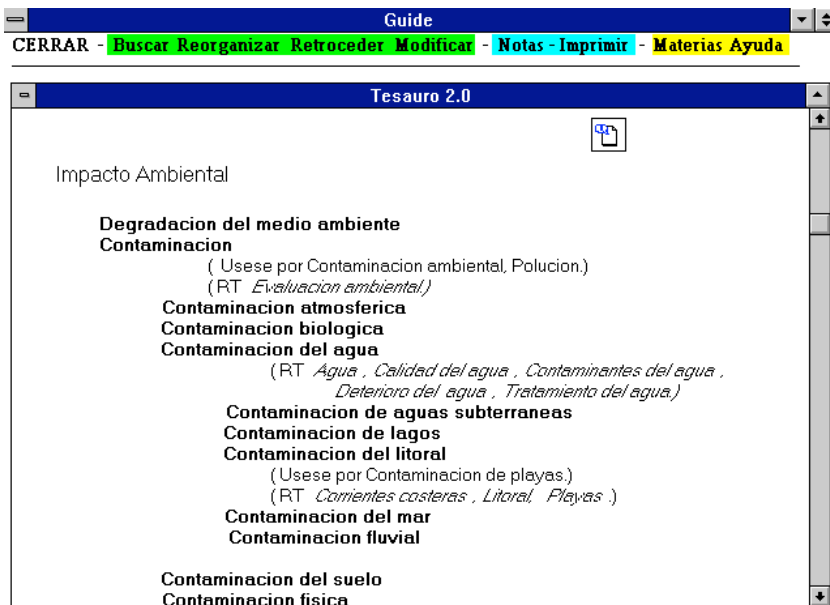
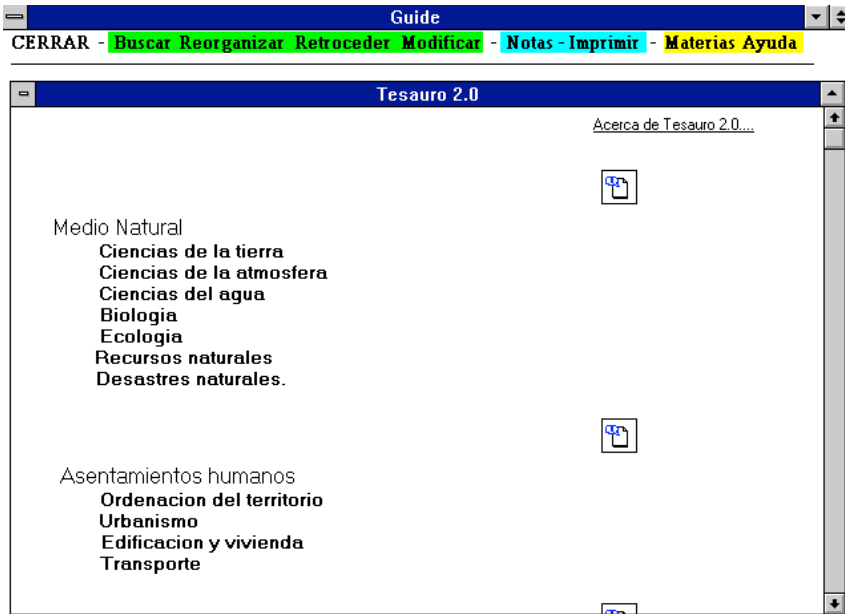


Fig. 7 y 8

5. *Búsqueda en texto libre*: Permite buscar una cadena completa o bien texto truncado, añadiendo además la posibilidad de búsqueda extendida, pues localiza a aquellos descriptores que tienen a uno especificado como término relacionado (gráfico 9)

Uno de los puntos más a destacar, que demuestra la validez del modelo propuesto, es la correspondencia entre la estructura interna de relaciones del sistema de hipertexto usado y la estructura de un Tesauro:

- Término específico: Botón de expansión
- Término genérico: Retroexpansión
- Término asociado: Botón de Referencia
- Nota de alcance: Botón de Nota
- Opciones de contro: Botón de Comando

Las ventajas que se derivan de este sistema de consulta y navegación de tesauros radica en la agilidad de navegación y en el carácter intuitivo del interfaz,

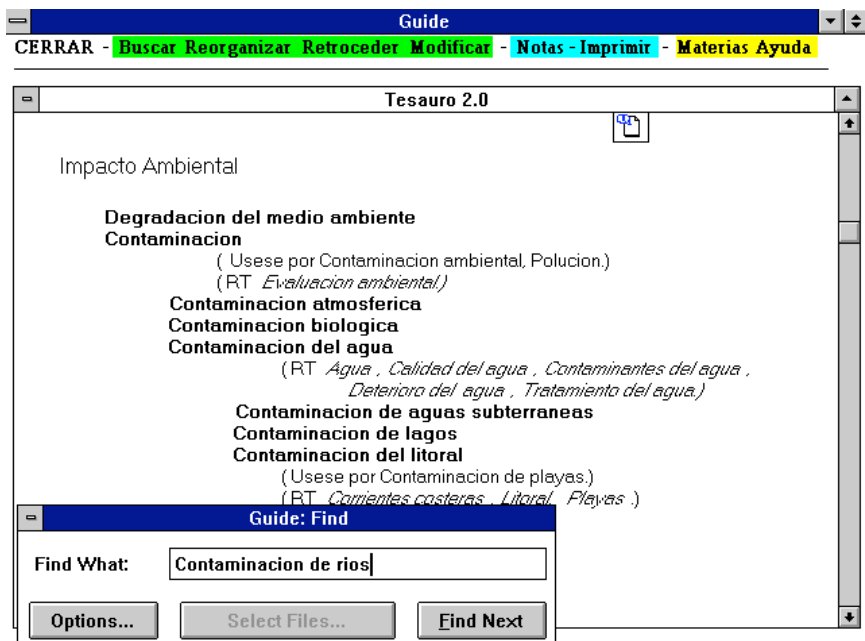


Fig. 9

denotado por:

- Las relaciones de Término Genérico, Término Específico, Use y Campos Semánticos se presentan de modo intuitivo, sin necesidad de una notación específica, mediante un simple sangrado, de un modo visual.
- Evita el cambio de presentación desde la alfabética a la sistemática, con la consiguiente pérdida del contexto. La presentación sistemática que se ofrece integra todas las relaciones del Tesoro.

Sin embargo además de todas estas ventajas, este prototipo de interfaz plantea algunos problemas, como son:

- Sólo se adapta de modo efectivo a Tesoros de estructura monojerárquica y monolingües.
- Tiene problemas en la conexión con un módulo que controle la coherencia de las relaciones, como el SGAT anteriormente presentado.
- El establecimiento de relaciones entre los términos del Tesoro debe realizarse de un modo totalmente manual, siendo extremadamente complicado construir un generador automático de los mismos.

Finalmente decir que el aspecto más relevante de este interfaz de Tesoros es que presenta los términos y las relaciones de modo sintético y no del modo analítico al que nos tienen habituados la mayoría de los Tesoros actuales, permitiendo una mejor comprensión de la estructura global del lenguaje, de modo que el esfuerzo innecesario de manejo sea el mínimo posible.

5. Conclusiones

¿Para que vale todo lo expuesto hasta el momento? Además de la posibilidad de usar los sistemas de navegación de Tesoros para realizar consultas controladas a Bases de Datos se nos presenta una nueva vía, todavía inexplorada.

Sobre hipertexto se ha hablado mucho sin llegar a aclarar del todo las posibles vías de solución de su problema principal: la desorientación del usuario, la pérdida, la confusión, el caos ...

La razón de la desorientación es que las relaciones de hipertexto se establecen de modo arbitrario y casual, por lo que planteamos la necesidad de un esquema conceptual de construcción/navegación en hipertexto, que explique las relaciones y dé contenido a los movimientos realizados o por realizar por parte del usuario.

Proponemos un nuevo paradigma de organización de documentación, que no puede ser un modelo universal, pero si adaptado a cada tipología documental. Para llevar a cabo esta empresa existen dos perspectivas:

- a) La metafórica que asocia una imagen conocida con el modelo de organización de información (un tarjetero, una estantería, etc.).
- b) La que modeliza una estructura conceptual como la propuesta por Antonio Rodríguez de las Heras (arquitectura de bucle abierto) y la propuesta teórica hecha por Isidre Canals basada en el concepto de Tesauro por facetas.

El prototipo implementado por nosotros (en fase de prueba) para la construcción de corpus de documentos completos se basa en la modelización de dos estructuras paralelas y profundamente interrelacionadas:

- a) Normalización de la estructuración de documentos en hipertexto y de la equivalencia entre textos lineales e hipertexto.
- b) Sobre ella un interfaz general más amplio al concepto de documentos en hipertexto basado en una red semántica, cuyo caso específico propuesto aquí sería la del Tesauro; esto es *un interfaz que explica las relaciones y que modeliza la navegación siendo esta igual para todo el sistema de documentos: el hiperdocumento*.

De este modo si en este trabajo hemos planteado el modo de gestionar tesauros mediante un sistema de hipertexto, también se plantea la necesidad de gestionar a los sistemas de hipertexto mediante un tesauro, o una red semántica formalizada.

6. Bibliografía

- Aitchinson, J.; Gilchrist, A (1987). Thesaurus construction. A practical manual. 2ª ed. London : ASLIB, 1987.
- André, J. ; Furuta, R. ; Quint, V. (comp.) (1989). Structured documents. Cambridge : Cambridge University Press, 1989.
- Heras, A. R. de las (1991). Navegar por la información. Madrid : Fundesco.
- Pastor Sánchez, J. A. (1992). Diseño, Desarrollo e Implementación de un Sistema Gestor para la Construcción de Tesauros. Murcia : Universidad, Escuela Universitaria de Biblioteconomía y Documentación. Proyecto fin de carrera, obra inédita.
- Rodríguez Muñoz, J. V. (1992). Construcción del esquema conceptual del tesauro mediante un modelo de datos. Murcia : Universidad, Facultad de Informática. Proyecto fin de carrera, obra inédita.
- Saorín Pérez, T. (1993). Generación de un Sistema Hipertexto orientado a la Gestión de Tesauros. Murcia : Universidad, Escuela Universitaria de Biblioteconomía y Documentación. Proyecto fin de carrera, obra inédita.