



Número 4, juny 2000

Distribució de bases de dades en el web amb *Knosys Internet*

Ernest Abadal Falgueras
Facultat de Biblioteconomia i Documentació
Universitat de Barcelona
abadal@eubd.ub.es

Raúl Martínez Rivas
Universidad Politécnica de El Salvador
mr21@latinmail.com

1 Introducció: distribució de bases de dades al web

No hi ha dubte que actualment el web és el sistema de distribució de bases de dades documentals que més s'està utilitzant i el que té majors perspectives de futur. El motiu és senzill: l'usuari que consulta la base de dades només cal que disposi d'un navegador i, aleshores, pot accedir als registres de forma actualitzada disposant pràcticament de totes les prestacions de consulta que tenen els sistemes de gestió documental;¹ no necessita instal·lar al seu ordinador cap versió client del programa que gestiona la base de dades, sinó que tan sols ha d'indicar la seva petició dins d'un formulari html per tal de rebre les respostes també en aquest format.

Els elements que intervenen en aquest procés són els següents:

- Navegador (p.e. *Netscape* o *Explorer*).
- Servidor httpd (p.e. *Internet information server*, *Apache*, etc.).
- Programa CGI (p.e. *Knosys Internet*, *WWWISIS*, *WebPublisher*).
- Interfície de consulta.
- Base de dades.

D'aquestes peces potser la menys coneguda sigui el protocol CGI (Common Gateway Interface) que actua de sistema de comunicació (una passarel·la)² entre els dos sistemes (la base de dades i el web). El CGI és un estàndard per mitjà del qual un servidor web es pot comunicar amb un programa extern. El protocol CGI estableix una forma d'enviar dades des d'una pàgina web –per mitjà d'un formulari– i de processar-les mitjançant un fitxer executable –programa CGI– que està situat al directori cgi-bin, o equivalent, d'un servidor.

El programa CGI és una aplicació informàtica escrita en llenguatge de programació (Perl, C, C++, etc.) que posteriorment és executada i interpretada per un servidor web per tal de poder contestar peticions d'informació dels usuaris. El programa CGI és capaç de llegir/interpretar les ordres que se li transmeten des d'un formulari html, algunes d'elles introduïdes per l'usuari (p.e. els termes de cerca) i altres corresponents a paràmetres generals (p.e. ubicació dels programes i de la base de dades al servidor, format de visualització, etc.). A continuació, les executa i el resultat el transfereix a l'usuari en format html. Actualment, els programes CGI acoblables a sistemes de gestió documental són bastant costosos.

A part d'això, és necessari, a més, preparar una interfície de consulta adaptada a la base de dades que s'ha de consultar que tingui en compte els camps que s'han definit, els formats de visualització, etc. Aquesta interfície es construeix amb codificació html i també utilitzant els codis de programació de la CGI. Consta bàsicament de tres elements:

- pantalla de consulta;
- pantalla visualització de resultats (l·listat);
- pantalla de visualització del document.

La taula 1 intenta resumir els passos que es produeixen en la consulta a una base de dades per mitjà d'un navegador web, des que l'usuari s'hi connecta fins que rep una l·lista de resultats a la seva petició. D'aquesta manera es pot comprovar quina és l'actuació de cada un dels elements descrits anteriorment.

Taula 1. Funcionament del procés de consulta d'una base de dades

- El navegador obre una connexió amb el servidor httpd. Se li mostra un formulari de consulta de la base de dades en html (interfície de consulta). L'usuari introdueix les condicions de cerca i n'ordena l'execució.
- El servidor executa l'aplicació, inicialitza les variables de l'entorn (els paràmetres introduïts per l'usuari al formulari i altres paràmetres ocults) i envia la informació per mitjà de l'entrada estàndard. Aquestes variables s'emmagatzemen en un registre "virtual", cada una d'elles amb una etiqueta i un valor.
- El programa CGI examina les variables d'entorn i, si és necessari, processa les dades entrades (p.e. efectua la consulta a la base de dades).
- L'aplicació genera els resultats.
- El resultat torna al servidor httpd per mitjà de la sortida estàndard.
- El servidor envia la informació al navegador en format html.

La majoria de sistemes de gestió documental disposen d'aplicacions (programes CGI) que permeten que les bases de dades es puguin consultar a Internet des d'un formulari web. P.e. *Wwwisis*,³ *Knosys Internet*, o *Inmagic WebPublisher* serien les CGI corresponents a *CDS/ISIS*, *Knosys Windows* i *Inmagic DB/Text*, respectivament.. A la taula 2, i només a títol d'exemple, s'inclouen diverses adreces de bases de dades creades amb aquests programes.

Taula 2. Exemples de bases de dades documentals al web

Programa	Bases de dades
<i>FileMaker</i>	<i>BIGPI (Geologia de la Península Ibèrica)</i> < http://www.bib.ub.es/bigpi/bigpi.htm >
<i>Inmagic (WebPublisher)</i>	Coordinadora Documentació Biomèdica < http://www.doc6.es/cdb >
<i>Knosys Internet</i>	Micronet < http://www.micronet.es/knosysinternet > («Nuestros usuarios», ofereix una llista de bases de dades)
<i>Wwwisis</i>	Multimedium < http://www.multimedium.org.ve > SABA-DOC < http://www.ajsabadell.es/cdsx >

Dels programes esmentats ens centrarem a descriure i valorar l'estructura i funcionament de *Knosys Internet*.

2 Knosys Internet

Knosys Internet és un programa CGI que permet distribuir al web bases de dades que han estat creades o que són gestionades amb *Knosys* per a Windows. L'aplicació ha estat creada per Micronet <<http://www.micronet.es>>,⁴ amb l'objectiu de proporcionar un entorn fàcil i senzill per a la publicació de bases de dades al web. *Knosys Internet* es compon, bàsicament, de quatre elements –servidor de consultes, programa CGI, programa de publicació de bases de dades, programa monitor del servidor de consultes– que tot seguit descriurem.⁵ L'administrador de la base de dades ha de treballar essencialment amb el programa de publicació de bases de dades, i amb el monitor, que s'ocupa d'activar i desactivar la petició de consultes a les bases de dades instal·lades, entre altres funcions.

2.1 Servidor de consultes Knosys

El servidor de consultes per a bases de dades *Knosys* s'encarrega de gestionar la base de dades i processa les consultes realitzades pels usuaris. La seva funció és permetre, per mitjà d'una sintaxi pròpia i el protocol de comunicació TCP/IP, que les aplicacions client puguin realitzar consultes sobre bases de dades *Knosys*.

Aquesta és una peça vital del sistema ja que centralitza totes les consultes sobre les bases de dades que gestiona. Sempre ha d'estar activat, d'aquí ve que s'instal·li en un servidor (el de Windows NT, en el cas analitzat). El sistema és un servei multitasca: permet l'accés a diversos usuaris alhora i mantenir les darreres consultes emmagatzemades en memòria per tal de no haver-les de tornar a buscar. (El servidor de consultes *Knosys* es troba en el fitxer `knod.exe` i fa servir la llibreria d'enllaç dinàmic `wkno32.dll`).

2.2 Programa CGI

Aquest programa es troba instal·lat en el fitxer executable `Know3.exe`, i és la seva funció establir la comunicació entre el navegador web de l'usuari i el servidor de consultes *Knosys*. Des de la pàgina de consulta que utilitza l'usuari per fer cerques a la base de dades s'ha d'incloure una crida al programa per tal que, quan s'executi aquesta pàgina (és a dir, quan l'usuari premi el botó «Cercar») s'activi automàticament el programa CGI al servidor, es puguin llegir i interpretar les variables que envia l'usuari (els termes de cerca), i es presentin els resultats al navegador de l'usuari com una pàgina html.

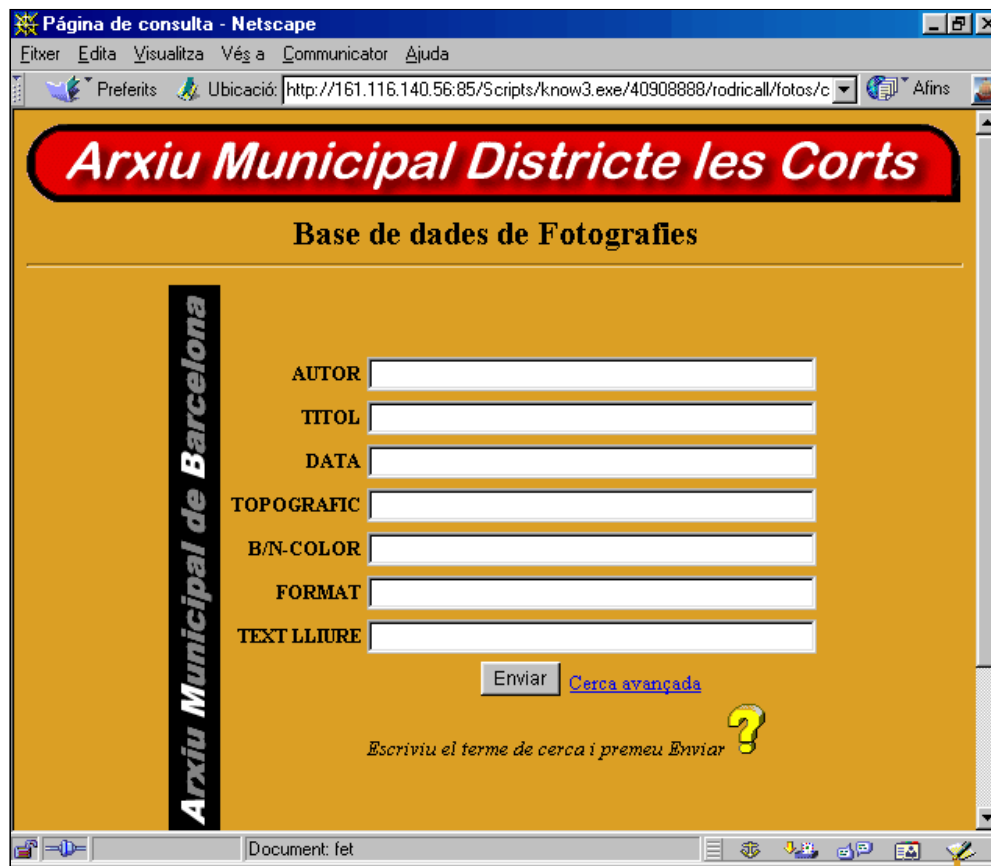
Per altra banda, també permet, per mitjà del fitxer `imagen.exe`, visualitzar les imatges que puguin estar associades als documents.

2.3 Programa de publicació de bases de dades

Es tracta d'un assistent per a la instal·lació a la xarxa de la base de dades pròpia. El programa de publicació de bases de dades ("instabd.exe") és el que s'encarrega de crear automàticament les pàgines actives que possibilitaran les consultes des d'un navegador web sobre una base de dades. La seva funció és, doncs, generar automàticament la interfície de consulta de les bases de dades, que es compon bàsicament de tres elements: pàgina de consulta, pàgina de resultats, i pàgina de visualització del document (a l'apartat 1 ja s'hi ha fet referència).

- **Pàgina de consulta**

La pàgina de consulta és el formulari que indica a l'usuari quines són les seves opcions de cerca. L'assistent de *Knosys Internet* permet especificar diferents models de pantalla: amb requadres de text, amb selectors de camp, o amb llistes desplegable, i també indicar sobre quins camps es pot fer la consulta.



The screenshot shows a Netscape browser window titled "Pàgina de consulta - Netscape". The address bar shows the URL: `http://161.116.140.56:85/Scripts/know3.exe/40908888/rodricall/fotos/c`. The main content area has a red banner with the text "Arxiu Municipal Districte les Corts" and a subtitle "Base de dades de Fotografies". Below this is a search form with the following fields:

- AUTOR
- TITOL
- DATA
- TOPOGRAFIC
- B/N-COLOR
- FORMAT
- TEXT LLIURE

At the bottom of the form is an "Enviar" button and a link for "Cerca avançada". A yellow question mark icon is also present. The browser's status bar at the bottom shows "Document: fet".

Figura 1. Exemple de pàgina de consulta amb requadres de text

També es poden consultar els valors predefinitos per l'usuari en algun dels camps. És el que *Knosys*

anomena «Admitidos». P.e. en el camp llengua podria haver-hi els valors *anglès*, *castellà*, *català*, *francès*, etc. que serien mostrats a l'usuari en un desplegable.

- **Pàgina de visualització de resultats**

Knosys Internet proporciona a l'usuari la possibilitat de consultar el llistat amb els resultats en dos formats: taula (on cada document és una línia i cada columna, un camp) o resum (inclou un registre dessota l'altre, sense estructura de taula). A més, es pot definir el color del fons i el color del text.

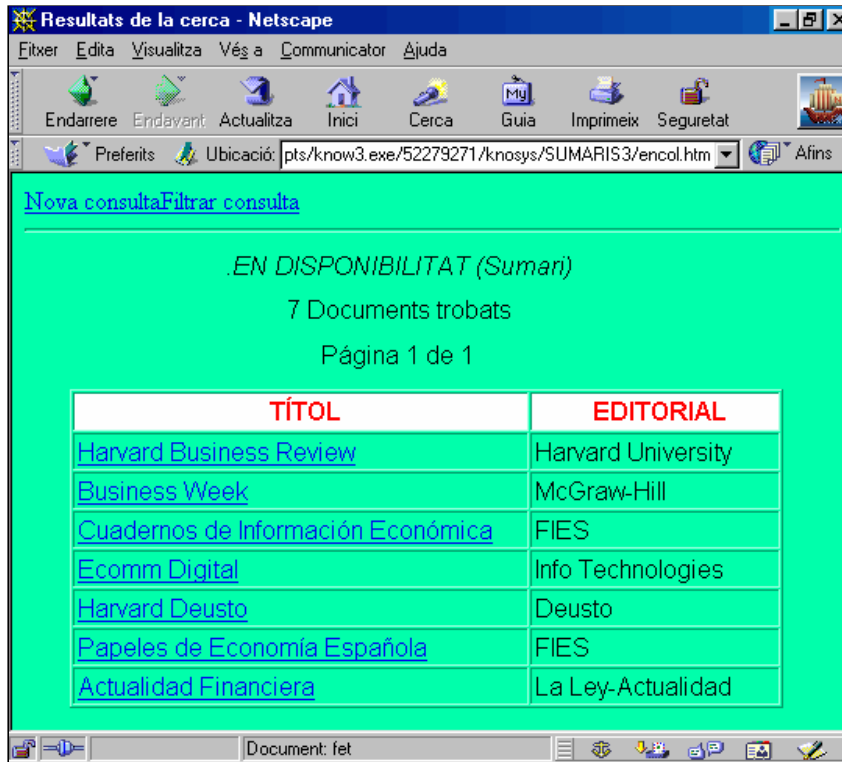


Figura 2. Pàgina de visualització de resultats en format taula

Quan es fa clic sobre les paraules subratllades es passa a la visualització del registre sencer (vegeu Pàgina de document).

- **Pàgina de document**

Inclou el document complet sol·licitat (que, a més del text, també pot contenir imatges). Pel que fa a la configuració d'aquesta pàgina només es pot triar quins són els camps que apareixeran (i el seu ordre) i si hi constarà, o no, el nom del camp. No es poden incloure signes de puntuació ni tampoc altres elements textuais per enllaçar camps.

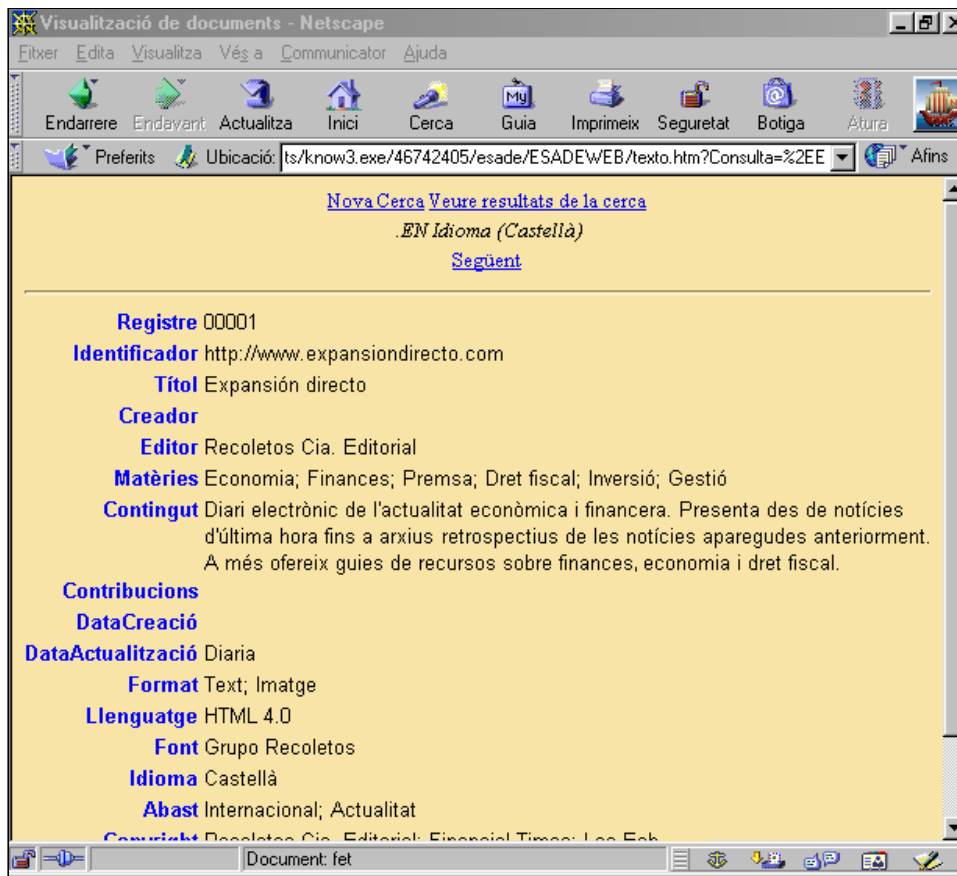


Figura 3. Exemple de pàgina de visualització de documents

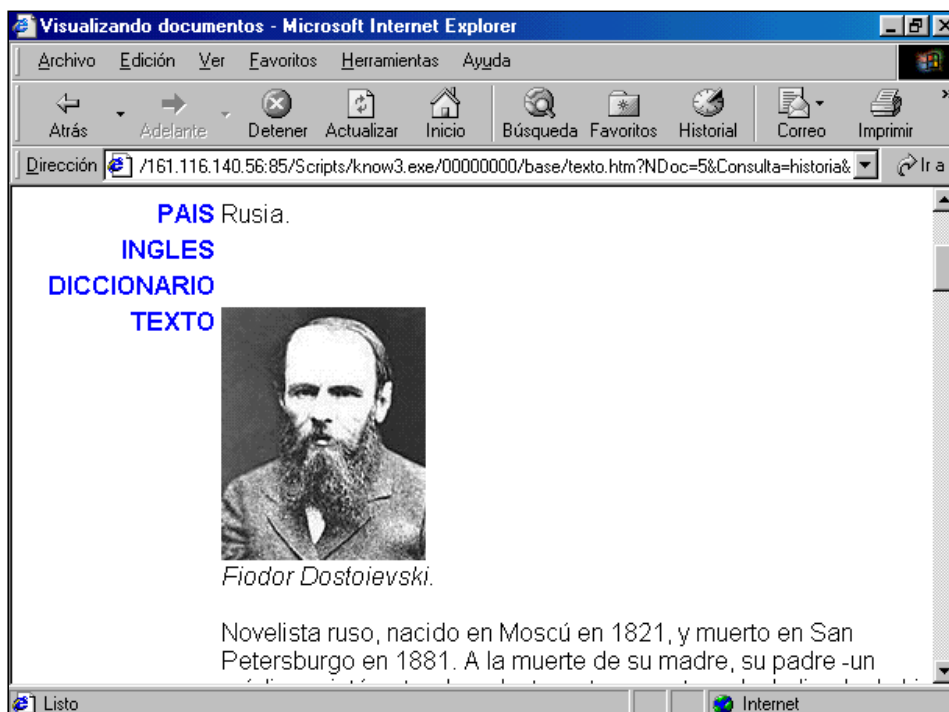


Figura 4. Exemple amb imatges

Les pàgines creades per l'assistent poden ser manipulades per l'administrador de la base de dades utilitzant codis html i també el llenguatge de programació de *Knosys Internet*, que té una estructura semblant als llenguatges Pascal i C++. D'aquesta manera es poden superar les limitacions de disseny que té l'assistent.

2.4 Programa monitor del servidor de consultes

El programa monitor és una pàgina activa (`monitor.htm`) que permet controlar en tot moment l'estat del servidor de consultes, veure informes dels accessos a les bases de dades, i tancar o reiniciar les bases de dades del servidor *Knosys*.

Amb aquest programa es poden atendre un nombre de consultes simultànies que dependrà del tipus de llicència que s'ha adquirit.

El monitor també disposa d'una opció per a generar informes sobre l'activitat del servidor que inclou un registre de totes les consultes que s'han realitzat a les bases de dades (indica l'IP de l'usuari i totes les equacions de cerca que s'han formulat), com també tota l'activitat d'administració del servidor i de les seves bases de dades (activació, desactivació, etc.).

Taula 3. Exemple d'estadístiques d'activitat

```

QRY 26Abr2000 16:46:09 <- [User: 161.116.140.19 Base: PERALADA]: QUERY 12 .EN Tipologia (Universitat)
QRY 26Abr2000 16:46:13 <- [User: 161.116.140.19 Base: PERALADA]: QUERY 12 .EN Tipologia (Manual)
QRY 26Abr2000 16:46:18 <- [User: 161.116.140.19 Base: PERALADA]: QUERY 12 .EN Tipologia (Llibre de notes)
QRY 26Abr2000 16:46:22 <- [User: 161.116.140.19 Base: PERALADA]: QUERY 12 .EN Tipologia (Encants)
QRY 26Abr2000 16:46:24 <- [User: 161.116.140.19 Base: PERALADA]: QUERY 12 .EN Tipologia (Capbreu)
QRY 26Abr2000 19:06:44 <- [User: 161.116.140.14 Base: AARTICLE]: QUERY 31 .EN Autor (Deleuze Gilles)
QRY 26Abr2000 19:06:47 <- [User: 161.116.140.14 Base: AARTICLE]: QUERY 31 .EN Autor (Fortuny Bonet Francesc)
QRY 26Abr2000 19:06:50 <- [User: 161.116.140.14 Base: AARTICLE]: QUERY 31 .EN Autor (García Calvo Agustín)
QRY 26Abr2000 19:10:50 <- [User: 161.116.140.11 Base: FOTOCELH]: QUERY 20 paper .Y .EN Col·lecció (celh)
QRY 26Abr2000 19:11:25 <- [User: 161.116.140.11 Base: FOTOCELH]: QUERY 20 .EN Col·lecció (Foto Arxiu Josep Bonastre)
QRY 26Abr2000 18:37:29 <- [User: 161.116.140.26 Base: SACU]: QUERY 22 .EN Periodicitat (Mensual)
QRY 26Abr2000 18:37:36 <- [User: 161.116.140.26 Base: SACU]: QUERY 22 .EN Idioma (Anglès)
QRY 26Abr2000 18:37:40 <- [User: 161.116.140.26 Base: SACU]: QUERY 22 .EN Idioma (Català)
QRY 26Abr2000 18:42:09 <- [User: 161.116.140.23 Base: MRRJ]: QUERY 29 .EN Autor/es (piquero)
QRY 26Abr2000 18:19:37 <- [User: 161.116.140.17 Base: TDUC1]: QUERY 2 .EN Universitat (Universitat de Barcelona)

ADM 26Abr2000 11:19:57 <- [IP = 161.116.140.56]: ADMINUSER 161.116.140.56
ADM 26Abr2000 11:19:57 <- [User: 161.116.140.56]: REINICIABASES
ADM 26Abr2000 18:06:02 <- [IP = 161.116.140.56]: ADMINUSER 161.116.140.56
ADM 26Abr2000 18:06:02 <- [User: 161.116.140.56]: ABRECERRADAS

SYS 26Abr2000 14:40:19 Cerrando base de datos "ejemplo5"
SYS 26Abr2000 14:40:19 Cerrando base de datos "CONMAR"
SYS 26Abr2000 14:40:19 Cerrando base de datos "CDMCA"
SYS 26Abr2000 18:06:02 Abriendo base de datos "d:\alumnos\39698821\AARTICLE.DOK"
SYS 26Abr2000 18:06:02 Tipo Servidor: Oro
SYS 26Abr2000 19:03:10 Abriendo base de datos "d:\alumnos\53121024\BESCOLA.DOK"
SYS 26Abr2000 19:03:14 Abriendo base de datos "d:\alumnos\52788980\RONYONS.DOK"

```

2.5 Valoració

Es tracta d'un programa amb prou prestacions com per satisfer les necessitats de distribució de bases de dades de centres petits i mitjans. Tot i això, l'aplicació encara s'està desenvolupant en molts aspectes (moltes vegades gràcies a les peticions de diversos clients que han anat sol·licitant modificacions i noves prestacions, que es van incloure a les versions actualitzades del programa).

El producte destaca per la seva senzillesa en la instal·lació i utilització i, en aquest sentit, segueix el model ja instaurat per *Knosys Windows*, d'estructura i funcionament molt simples.

Pel que fa al cost, cal assenyalar que disposa d'una interessant política de preus perquè permet ajustar el cost de la llicència al nombre de registres i al nombre de connexions simultànies. Això permet que centres petits i mitjans puguin obtenir una llicència a un preu més assequible, ja que poden ajustar el preu en funció del valor que tenen en el seu cas ambdues variables.

Pel que fa a les mancances del programa, coincidim bastant amb la valoració que van fer José V. Rodríguez i Tomás Saorín⁶ en un article que descriu una experiència de distribució d'una base de dades documental amb *Knosys Internet*, tot i que potser posen excessivament l'accent en els aspectes més negatius. Passem a comentar alguns dels problemes:

- **L'assistent per a publicar la base de dades al web**

Tot i que és molt clar el suport que facilita aquesta eina (en tres minuts es pot tenir disponible una base de dades per a la seva consulta al web), no es pot oblidar que la configuració que permet és molt bàsica, ja que no es poden dissenyar pàgines de consulta que facin servir més d'un tipus d'interacció (requadres de text, botons de ràdio, desplegable), no es pot donar l'opció de poder escollir l'operador que s'ha d'aplicar per a les operacions entre diferents requadres de text (s'activa la intersecció, per defecte), tampoc es poden configurar la majoria de missatges que apareixen a la pàgina, ni els botons, etc. Tots aquests elements s'han de retocar a mà per l'administrador des del fitxer de configuració de les pàgines de consulta, llistat i document. Per altra banda, els fitxers que es generen (i que contenen les especificacions de disseny de la interfície) són molt llargs i és difícil moure-s'hi per a retocar-los.

- **Ordenació de resultats**

La llista de resultats que obté l'usuari com a resposta a la seva petició no es pot ordenar per cap criteri: ni per ponderació del resultat (el sistema no disposa d'eines per a fer el càlcul ponderat i mostrar en primer lloc els registres que són més pertinents segons unes regles internes del programa), ni per camp (a l'usuari, a vegades, li pot interessar que el llistat s'ordeni per l'autor, pel títol o per la data).

- **Consultes a diverses bases de dades**

Tal i com ja passa a *Knosys* per a Windows (i en la majoria de programes de la seva tipologia), no es pot demanar que el programa realitzi consultes a més d'una base de dades.

Aquests dos darrers inconvenients són resolts en la majoria de sistemes d'indexació o motors de cerca: *Excalibur*, *Search'97*, *Oracle InterMedia*, tot i que, en aquest cas, es tracta d'un tipus de programa amb més prestacions i, per tant, amb un cost sensiblement superior.

Vistos els avantatges i els inconvenients, i per tal de concloure, cal destacar que es tracta d'un producte que pot ser d'interès especial per a les necessitats de centres petits i mitjans, que constitueixen el tipus de client al qual el programa li permet satisfer la majoria de les seves peticions.

Notes

¹ Es pot trobar una descripció i anàlisi de les característiques, estructura i tipologia dels sistemes de recuperació de la informació a: Codina, Lluís; Abadal, Ernest. «Gestió documental amb microordinadors: característiques, estructura i tecnologia dels sistemes de gestió documental», *Item*, 11(1992), p. 72-100.

² Es parla de passarel·la (*gateway*) perquè fa referència a la funció de relació entre el servidor web i les aplicacions externes. El protocol CGI permet comunicar un programa que està instal·lat en el servidor amb un servidor httpd, per a obtenir documents html *dinàmics* (és a dir, que es generen en un moment determinat, ja que varien segons quina ha estat la petició de l'usuari).

³ Es pot consultar el comentari que se'n va fer a: Abadal, Ernest; Cuadrado, Maribel; Gascón, Jesús; Omella, Ester. «Disseny i creació de bases de dades bibliogràfiques amb CDS/ISIS: l'experiència de SABA-DOC». *BiD*, núm 3 (desembre 1999). <<http://www.ub.es/biblio/bid/03abadal.htm>>

⁴ Marcos, Mari-Carmen. «Micronet, empresa española de producción, mantenimiento y distribución de información», *El profesional de la información*, vol 9, nº 3 (marzo de 2000), p. 20-24.

⁵ *Manual del usuario: Knosys Internet*. Madrid: Micronet, [1998?]. 125 p.

⁶ Rodríguez Muñoz, José V.; Saorín, Tomàs. «Modelado documental de servicios de información en web», *El profesional de la información*, vol 7, nº 9 (septiembre 1998), p. 10-18.