

Algunas Sugerencias sobre la Evolución de las Bases de Datos de la Unión Europea



por Alfonso Moreira (jmoreira@uv.es)
y Carmen Gímez (gimenezc@uv.es)
Centre de Documentació Europea
Universitat de València



Introducción

El desarrollo de las tecnologías de la información y la aparición de un protocolo mundial de telecomunicaciones (TCP/IP) con la consiguiente popularización de la red Internet, ha golpeado de frente al sector de la información on line de tipo profesional. Dejando a un lado el caso del CD-ROM -el beta de los noventa según algunos comentaristas-, la revolución que estos avances están propiciando, junto con la incertidumbre que ello conlleva, amenaza de forma directa la misma base de los negocios de los distribuidores de información on line.

La situación actual resulta paradójica. Por un lado, el acceso de los usuarios a sus servicios resulta más sencillo que nunca -desde un punto de vista técnico- lo que ha conducido a la apari-

ción de un nuevo cliente que ya no tiene necesariamente que ser el especialista que utiliza complicados lenguajes de interrogación. Por otro lado, el abaratamiento de precios de máquinas y programas permite la progresiva entrada de nuevos actores en el negocio de la información on line. Es previsible por tanto, que una mayor competencia repercute en los costes de manera favorable para el usuario; pero también es indudable que la clarificación de los servicios ofrecidos a través de la red Internet, acabará beneficiando a aquellos que sean capaces de ofrecer contenidos precisos y de calidad: desde este punto de vista, parece claro que los mejor situados en la competición que se avecina son los distribuidores de bases de datos tradicionales.

Por último señalar la posibilidad de que la globalización del mercado haga innecesaria la existencia de los host tradicionales como vías de paso hacia

las bases de datos, al existir la posibilidad de establecer conexión directa con sus productores.

Ante este panorama el desconcierto es general tanto en el seno de los operadores privados como entre los públicos. Los dos grandes distribuidores de bases de datos de la Unión Europea -ECHO y Eurobases- no constituyen un caso aislado y sufren las mismas amenazas y tienen las mismas oportunidades que los demás distribuidores on line.

Nuestra intención en este artículo es la de realizar una serie de sugerencias sobre la transformación del servicio telemático de la UE que pueden conducir, siempre desde nuestro punto de vista, a una mejora del mismo. La duplicidad editorial de la Unión Europea -bases de datos electrónicas y publicaciones en papel- entendida como dos ofertas informativas distintas corresponde a una época pasada; en la actualidad resulta aconsejable la integración de ambos servicios en una serie de nuevos productos accesibles electrónicamente mediante la red Internet. Ello no significa, en un primer momento, la eliminación de determinados productos on line o impresos, sino la multiplicación de la oferta informativa, dejando en manos del usuario la elección final.

En la base de esta transformación tenemos el hecho de que todas las publicaciones de la UE se encuentran ya formando parte de ficheros electrónicos. Su posterior edición, bien electrónica o impresa, tiene su origen en un fichero informático. Nuestra propuesta es la de abrir toda esa información textual a la consulta pública a través de un servidor W3, donde se integrarían los nuevos productos compartiendo características de base con los productos impresos -más que con las bases de datos-, si bien las peculiaridades del medio permiten añadirles una serie de funciones imposibles en la edición impresa, como buscadores, actualización inmediata, etc.

Aunque la Comisión Europea dispone en la actualidad de un servidor W3 central (<http://www.europa.int>) y varios satélites del mismo (ISPO, CORDIS y ECHO, así como los de las oficinas de Representación en los Estados miembros comentados en este mismo número en el artículo de E. Aleixandre), las deficiencias localizadas en dicho servicio nos servirán de base para la realización de una crítica constructiva.

Aunque cuenta con aciertos notables (el servidor de documentos sobre la Conferencia Intergubernamental, por ejemplo), el servidor de la Comisión no resulta un buen ejemplo de claridad informativa. Con una estructura deficiente, obtusa en cuanto a contenidos, carente de productos definidos y con preocupantes vacíos informativos (al parecer existen diversas direcciones generales que no están en esta guerra), el servidor Europa produce un especie de confusión generalizada en el usuario que creemos tiene fácil solución.

Así pues, repasaremos en principio las características actuales de los host comunitarios señalando su composición, acceso, lenguajes de interrogación, etc, para pasar después a realizar algunas sugerencias concretas sobre la necesaria transformación del servicio. Desde nuestra perspectiva de documentalistas en un Centro de Documentación Europea creemos poder aportar nuestro conocimiento de las necesidades reales del usuario final, que no son siempre tenidas en cuenta por los responsables de los servicios informativos de las grandes instituciones.

Los Distribuidores on line de la UE

Si bien existen varios distribuidores de información electrónica de la UE, no es el objeto de este artículo repasar en profundidad estos servicios on line y el contenido de las bases de datos por ellos distribuidos¹. No comentamos pues las bases de datos de Eurostat -Cronos, Comext, Regio, , etc-, ni repasamos las que contienen información proveniente de las actividades del Parlamento europeo - Epoque-. Centraremos nuestros comentarios sobre los dos grandes Host: Eurobases y Echo, nombrando algunas de las características de sus bases de datos más significativas.

EUROBASES, sobre lenguas y lenguajes

EUROBASES es uno de los dos grandes distribuidores de la UE que ofrece acceso a una gran variedad de bases de datos comunitarias a través de una dirección X.25 o haciendo Telnet a la dirección IP *eurobases.lu*. Estas bases de datos

son: *Celex*, –recoge el derecho comunitario en sus diferentes versiones lingüísticas–, *Scad* –base de datos bibliográfica–, *Rapid* –servicio de prensa de la UE–, *Info 92* –sobre la realización del mercado interior–, *Eclas* –Fichero de la biblioteca de la Comisión–, *Sesame* –proyectos de investigación y desarrollo de la energía–, *Abel*. –Diario oficial, sólo sumario–, *Oil* –Boletín petrolero semanal–, *Eurocron* –indicadores macroeconómicos, regionales y resultados de las explotaciones agrarias–, *Idea* –directorio de los representantes de todas las instituciones–.

Sus productos estrella se basan en la utilización del complicado, por engorroso, lenguaje de interrogación MISTRAL, cuyos comandos, siendo los mismos, se escriben de forma diferente según la versión lingüística en la que trabajamos: en el caso de *Celex* por ejemplo, es muy normal el tener que cambiar de idioma de interrogación bien por exigencias del usuario, bien porque algunos documentos no se encuentran disponibles a texto completo en determinados idiomas como sucede con el castellano. Quizá la razón que explica la aparición de nuevas bases de datos que continúan utilizando Mistral sea el hecho de que se trata de un lenguaje que ofrece buenas posibilidades para la interrogación de bases de datos documentales, al disponer de la posibilidad de indexación posicional, lo que representa una ventaja en el caso de grandes bases de datos documentales o con documentos de gran extensión; sin embargo es un lenguaje poco conocido, incluso por los profesionales, y un buen rendimiento se obtiene sólo después de muchas horas de conexión. Desde nuestro punto de vista, y existiendo en el mercado software potente, competitivo y de uso más amigable, sólo una actitud muy conservadora justifica esta tendencia.

El nivel de uso de las bases de datos de la UE, si consideramos la demanda actual de información, es muy bajo; y pensamos que ello radica fundamentalmente en la dificultad de acceso a las bases de datos, pero también en la complejidad de su uso que no propicia la confianza en el medio por parte del usuario.

Las bases de datos de EUROBASES. Algunos ejemplos

CELEX (Communitatis Europaeae Lex) fue creada en 1969 para uso interno aunque no fue hasta 1981 que se dio acceso general al público. *Celex* es una de las bases de datos más importantes de las producidas por la UE, por su relevancia política al contener el Derecho comunitario en su sentido más amplio, incluidos actos preparatorios, preguntas interpuestas en el Parlamento y jurisprudencia, y por los medios puestos a su disposición que hacen que gran parte de los documentos de *Celex* contengan el texto integral en las diferentes versiones lingüísticas. *Celex* se subdivide actualmente en 8 sectores, aunque está pensada desde hace años (con la promesa de la introducción de la jurisprudencia nacional en aplicación del derecho comunitario y la doctrina jurídica del derecho comunitario) la ampliación a 10; lo que unido al lenguaje Mistral, el utilizado para su organización e interrogación hacen de ella una bases de datos extremadamente compleja.

El problema de *Celex* reside en que la obtención de resultados de muy alto nivel sólo es posible para aquellos usuarios que sean buenos conocedores de su estructura y del lenguaje Mistral, lo cual no es una situación “idónea” para permitir una utilización más masiva de dicho medio. Una alternativa a la conexión on line ha sido para muchos usuarios la utilización del CD Rom que incorpora un software de búsqueda más sencillo, aunque cubre solamente los sectores 3 y 4 –full text– y tiene una actualización trimestral, que se convierte en semestral en el caso de aquellas lenguas que siendo oficiales son consideradas de “segunda división”. No parece entonces que sea ésta la alternativa más idónea para la consulta de los contenidos de *Celex*.

SCAD es la principal base de datos bibliográfica de la UE que se refiere a todos aquellos documentos publicados en el mundo y que tienen que ver directa o indirectamente con la UE. Accesible públicamente desde 1987 y con referencias que cubren desde 1983 hasta hoy día, la base se actualiza semanalmente y los nuevos registros se publican también en versión papel bajo el nombre de SCAD boletín.

Los documentos de Scad se hallan clasificados en cuatro sectores distintos: documentos que se refieren a los principales actos comunitarios –sector A–, publicaciones oficiales –sector B–, artículos de revistas y periódicos –sector C– y dictámenes de los agentes sociales –sector D–.

Los registros de Scad están formados por 20 campos, la mayoría de los cuales pueden ser consultados directamente por tener su propio léxico –fichero índice–. Scad utiliza el mismo software de recuperación que Celex, y aunque su consulta es de menor complejidad debido tanto a la mayor similitud entre sus registros como al hecho de que sólo contenga referencias bibliográficas y ninguno de sus campos sea a texto completo, no creemos que el acceso actual sea el adecuado al momento presente: mejorar el producto ofrecido es necesario y factible.

ECHO, nuevos aires on-line

ECHO (European Community Host Organization) es el más reciente de los distribuidores de información electrónica de la Comisión Europea. Su creación data de 1980 con el objetivo de estimular el desarrollo y promoción de los servicios electrónicos de información en la Unión Europea. La puesta en marcha de ECHO significó para los servicios electrónicos de la Comisión la aparición de un nuevo lenguaje de interrogación –el CCL: Common Command Language– más simple y de fácil uso que los existentes hasta dicho momento y la estandarización del acceso a las nuevas bases creadas a partir de entonces en un punto único (con anterioridad una dirección X25, en la actualidad la dirección IP *echo.lu*). También se debe reseñar la introducción del multilingüismo durante el proceso de interrogación. Sin embargo, las bases de datos más útiles seguirían albergadas en Eurobases utilizando lenguajes de interrogación tan arcaicos como el ya comentado Mistral (incluso en 1995 han aparecido nuevas bases de datos consultables utilizando esta vieja gloria de los lenguajes de interrogación).

Además de actuar de distribuidor de numerosas bases de datos producidas por los servicios de información de las instituciones comunitarias, ECHO tiene otras funciones:

+ Servidor de difusión y demostración sobre las

ventajas de los servicios electrónicos de información a empresas y centros de investigación,
+ Servicio de atención a usuarios mediante los seminarios de sensibilización, la distribución de documentación sobre las bases de datos, disketes de formación, etc,

+ Apoyo al mercado de la información mediante el acceso en línea a las bases de datos I'M Guide (repertorio de servicios electrónicos de información en Europa) y a Experts Guide (repertorio de personas y organizaciones especializadas en el mercado de los servicios de información),

+ Puesta en marcha de proyectos innovadores relacionados con el acceso mediante lenguaje natural, la interrogación multilingüe, el intercambio electrónico de datos (EDI), etc.

Las bases de datos de ECHO

En la actualidad, ECHO alberga una serie de bases de datos que se pueden agrupar en cuatro grandes dominios:

1. Ayuda al usuario . CCL-Train (formación en línea acerca del CCL), I'M Guide (informaciones prácticas sobre bases de datos existentes, forma de acceso y su contenido), News on line con Echo facts for users (boletín informativo sobre el propio host y el mercado de la información) y con I'M News Archive (versión on line del boletín de alerta sobre el mundo de la información), XIII Magazine (versión electrónica de la revista publicada por la DG XIII), DUNDIS (Repertorio de sistemas de información y bases de datos de las Naciones Unidas) y UNESIB (referencias bibliográficas de documentos publicados por la UNESCO).

2. Investigación y desarrollo tecnológico . CORDIS (RTD-Programmes, RTD-Projects, RTD-Publications, RTD-ComDocuments, RTD-Results, RTD-Acronyms, RTD-News, RTD-Partners y RTD-InfoPoints), BIOREP (proyectos comunitarios en el ámbito de la biotecnología), DOMIS (fuentes informativas sobre materiales), EUREKA (proyectos EUREKA) y EURISTOTE (repertorio de 15.000 tesis o estudios desde los años 50).

3. Ser vicios lingüísticos. EURODICAUTOM (base de datos de términos técnicos y científicos) y THESAURI (recopilación de tesauros).

4. Industria y economía. TED (licitaciones públicas de más de ochenta países), JUSLETTER (actos jurídicos legislativos en Europa), MISEP (textos comunitarios y nacionales sobre la política de empleo en Europa) y EMIRE (introducción a la leyes nacionales sobre empleo y prácticas en la industria).

La manera más sencilla de contactar actualmente con ECHO es accediendo a su servidor 3W en la dirección <http://www.echo.lu> o acciando un Telnet a la dirección *Echo.lu* con el password anonymous.

Algunas sugerencias

La posible transformación del panorama descrito en los apartados anteriores puede ser analizado desde dos puntos de vista no excluyentes.

En primer lugar el acceso a las bases de datos existentes en la actualidad es susceptible de mejora mediante la actualización del software que lo soporta. Grandes paquetes gestores de bases de datos documentales cuentan ya con una versión W3 -el caso de BASIS-Web- e incluso nos consta (Eurobases Bulletin, nº 1 -february 1996) que algunos pasos se están dando en Eurobases de cara a la modernización del acceso. Estos nuevos gestores garantizan la continuidad de la carga por métodos tradicionales mientras que los usuarios se benefician de las ventajas de la consulta mediante pantallas-formulario html.

La aparición de Mistral Web permitiría -¡ por fin !- la jubilación de la versión comando del mismo lenguaje de interrogación y la puesta en marcha de un verdadero servidor de documentos que cubriría, según nuestra experiencia, el 90 % de las necesidades informativas de investigadores, alumnos, empresas y público en general. CELEX W3 permitiría la descarga automática (vía ftp) de directivas, reglamentos, decisiones y sentencias en el idioma deseado por el usuario. De esta forma, el servicio se tarificaría en relación con los documentos descargados: se paga por lo que verdaderamente se necesita. Si bien no permite la precisión de búsqueda deseada por los expertos en la jerga mistral, una adecuada presentación que permita la búsqueda por número CELEX,

Diario Oficial en el que se ha publicado el documento y por temas, siguiendo la clasificación de Repertorio de Legislación Comunitaria Vigente, sería suficiente en la mayoría de los casos, especialmente si se proporciona la posibilidad de utilizar el campo fecha para posteriores actualizaciones. Técnicamente este tipo de búsquedas es factible a través de wais u otros programas similares que entienden preguntas en lenguaje natural. Para aquellos insatisfechos siempre les quedaría la posibilidad de la consulta on line.

Proyectos en los que se encuentra actualmente Eurobases, como la cobertura total de los actos legislativos en los once idiomas oficiales o la necesaria accesibilidad de la legislación consolidada, son totalmente integrables en el nuevo servidor. Además se debería de aprovechar para proceder a la carga del texto completo de los documentos preparatorios -COM, PE DOC y CES- que la base de datos actual recoge en el sector 5 de su estructura. El servidor de documentos COM incluiría a la totalidad de libros blancos, libros verdes, proyectos de directivas y otras propuestas de la Comisión demandadas de manera sistemática por los usuarios de los centros informativos de la Unión Europea en los quince países miembros.

Aunque con menor urgencia, debido a su menor importancia, las demás bases de datos referenciadas en párrafos anteriores, son también susceptibles de transformarse en servidores de información accesibles a cualquiera que cuente con una cuenta de conexión a Internet. Rapid, la base de datos de noticias oficiales de la Comisión, se encuentra, de hecho ya casi integrada en el servidor Europa de la Comisión Europea aunque no de manera satisfactoria. Por poner un ejemplo no sería extremadamente difícil conseguir que la consulta de la base de datos del Servicio de Documentación de la Comisión (SCAD) pudiese realizarse mediante un sencillo formulario al que podríamos señalarle, por ejemplo, que nuestros intereses son los de conseguir bibliografía sobre la política agrícola comunitaria, en castellanos e inglés, de revistas internacionales publicadas los dos últimos años.

En relación con este punto, señalar la excepcional acogida que experiencias piloto de este tipo - servidor CORDIS en Irlanda

<http://slarti.ucd.ie/inttelec/cordis.html> - han tenido entre la comunidad de usuarios y profesionales del medio. Resulta sorprendente comprobar como unos usuarios avezados en las dificultades del comando -nosotros mismos- preferían realizar sus búsquedas rellenando un sencillo formulario en el servidor irlandés, en lugar de conectarnos on line, vía telnet, con el host correspondiente -ECHO en este caso-. Por cierto, el servidor experimental no duró más de un año, la falta de financiación del servicio les obligó a cancelarlo. Era simplemente demasiado bueno para aquellos tiempos.

Una segunda mirada, complementaria con la anterior, se basaría en el concepto señalado más arriba acerca de la cercanía existente entre los productos informativos en su versión papel y en su versión W3. Desde nuestro punto de vista, la carencia actual de contenidos de la red, reside en parte, y es normal que así sea en una primera etapa de indefinición e incertidumbre, en la preponderancia en las publicaciones electrónicas de la forma sobre el fondo. El esfuerzo titánico del sector de los informáticos por poner en marcha algunos servidores nos invita a pensar en una revista o un periódico en el que el responsable de la maqueta decidiera sobre los contenidos, lo cual como es lógico conduce a la aparición de un producto que si bien se intenta adaptar a las características del nuevo medio, no aprovecha en absoluto sus posibilidades. El resultado es un producto inútil que lejos de interesar al público, le decepciona. Trabajos de este tipo los podemos encontrar a lo largo de toda la red. Existen, claro está, notables excepciones pero ejemplos como la publicación EUR-OP News (<http://europa.eu.int/en/comm/opoce/wel4.html>) en la que una publicación extremadamente útil en su versión impresa ha sido convenientemente masacrada en su versión electrónica, nos reafirma en nuestras convicciones.

Se trata pues, de definir convenientemente los productos que se ofrecen y, cuando hablamos de información comunitaria el producto estrella es uno y múltiple: el Diario Oficial de la Unión Europea (un solo diario y tres series diferentes). La posibilidad de poner en marcha una versión electrónica del DOCE parece no presentar excesivos problemas técnicos: Una versión diaria del mismo, en la que los usuarios puedan descargar

a su ordenador aquellas directivas, reglamentos o convocatorias que sean de su interés, parece solamente chocar contra los intereses económicos de la propia Oficina de Publicaciones de la Comunidad Europea (OPOCE) en primer lugar, y de sus distribuidores en los países miembros en segundo término.

El DOCE es un bestseller en quince países y resulta impropio arruinar un negocio que funciona. La suscripción a la tres series del mismo -legislación, comunicaciones y suplemento- cuesta en la actualidad alrededor de 200.000 ptas, aunque la mayor parte de ellas se diluyen en coste del papel, envíos e intermediarios. -Además ocupa una cantidad considerable de espacio en la estantería-. La puesta en marcha de la versión electrónica del Diario Oficial con unas tarifas basadas en la descarga de documentos, siendo más económica para el usuario, produciría recursos suficientes para el mantenimiento de los servicios informativos de la Comisión Europea, a la vez que acercaría las actividades de la UE a la gran masa de ciudadanos. Además, es ecológico.

Sin extendernos en las características del servicio, solamente señalar la extremada sencillez de consulta de la versión electrónica del Diario Oficial. Nos conectamos, repasamos el índice del mismo y, si procede, descargamos en nuestro ordenador el documento pertinente. Resultaría similar al funcionamiento actual de TED (base de datos de licitaciones) que también realiza sus cargas pecuniarias en relación con el número de documentos mostrados por el servidor.

Otras de las publicaciones susceptibles de reconversión dentro de las posibilidades de las nuevas tecnologías de la información es el Boletín de la Unión Europea. Este Boletín de periodicidad mensual, con contenidos de utilidad notable, resulta escasamente aprovechable debido a la tardanza en su elaboración, edición y envío. La aparición de una versión electrónica del mismo dentro de unos plazos razonables -el día 10 de cada mes- dotaría a Centros de Documentación, Euroventanillas e interesados en general de una herramienta apropiada para el seguimiento de la actividad de las diferentes instituciones y políticas comunitarias. Las mejoras del mismo en su versión electrónica serían obvias: links con los

documentos allí señalados, posibilidad de búsquedas, consulta de referencias anteriores, etc.

Por último, solo señalar que publicaciones como Panorama de la Industria (anual de análisis sobre los sectores industriales en la UE), Informe Anual de Actividades, El Informe sobre política agrícola, Boletín Semanal de actividades del Tribunal de Justicia de la UE, Europa en Cifras o los Anuarios temáticos de la Oficina Estadística de la UE, son igualmente apropiadas para la reconversión, por tratarse de publicaciones que contienen un tipo de información generalista que resulta extremadamente útil para la resolución de un gran porcentaje de consultas informativas sobre la Unión Europea.

Conclusión

Los distribuidores on line oficiales de la Unión Europea tienen ante sí la responsabilidad de actuar de acuerdo al signo de los tiempos. Con una serie de decisiones adecuadas, las amenazas se tornarán en mejoras en el servicio. Hemos tratado de elaborar una propuesta basada en una serie de puntos concretos desde el punto de vista del usuario. También somos conscientes de no haber tocado los aspectos económicos de la transformación del servicio de información, pero creemos que es tiempo de actuar para mejorar de manera cualitativa y cuantitativa el servicio. A nuestro modo de ver resulta una oportunidad de llenar el vacío informativo que siempre se ha achacado a la Unión Europea. Esperemos que se aproveche.

Notas

1 Comisión Europea
European Union database directory
Luxemburgo: OPOCE, 1995
ISBN: 92-826-9580-8

