

Bases de datos y bibliotecología: como deshacer la innecesaria incomunicación

Eduardo Villanueva Mansilla

Pontificia Universidad Católica del Perú,
Departamento de Humanidades; Lima,
Perú E-mail: evillan@pucp.edu.pe

RESUMEN

El artículo presenta aspectos del diseño y uso de las bases de datos que nos hacen reflexionar sobre aspectos básicos que hacen que éstas realmente sean una herramienta útil para la recuperación de la información. Lleva a la preocupación que todo diseñador de bases de datos llega a tener para lograr las interacciones de entidades y los modelos relacionales y distribuidos; y la normalización en el ingreso y la salida de la información, así como el éxito de los productos de las últimas tecnologías en beneficio de la interacción amigable del usuario con la base de datos, como pueden ser los WWW y los CGI. Cabe aclarar que no profundiza sobre estos aspectos, sino invita a su cuestionamiento retomando la permanente discusión entre estructuras y manejo de bases de datos.

ABSTRACT

The article presents aspects of databases design and use that make us reflect on the basic aspects that really make them an useful tool for information retrieval. The paper is concerned with the involvement that every database designer must have for achieving the interactions among entities and relational and distributed models; and the standardization in the input and output of information as well as the success of the last technologies products for the sake of the user's friendly interaction with the database, such as the WWW and the CGI. It's worth mention that the paper do not make deep commentaries on this aspects, but rather it invites to their questioning on the basis of the permanent discussion between database structures and management.

Uno de los asuntos que más problemas suele traer a los bibliotecarios en su desempeño profesional es el relacionado con las bases de datos. El uso de las aplicaciones para crear y mantener bases de datos, ya sean éstas comerciales o no, se usen a través de sistemas de consulta remota o no, es cosa de todos los días. Sin embargo, muchas veces no es posible comprender plenamente todo lo que significa este uso.

El diseño de bases de datos avanzadas es el dominio de la teoría de bases de datos, área que ha sido propia de los ingenieros de sistemas y profesionales de los sistemas de información. Con el término diseño de bases de datos no nos referimos al simple traslado de las estructuras de des-

cripción y análisis de documentos que son parte de lo que, al menos en el Perú, se conoce como procesos técnicos, sino las que se utilizan en los sistemas comerciales tanto en bibliotecas como fuera de ellas. La creación de catálogos en línea, mediante aplicaciones administradoras de bases de datos, es parte importante de nuestro trabajo cotidiano, pero es sólo un aspecto, menor por cierto, de la teoría que mencionamos.¹

LOS MODELOS DE DISEÑO DE BASES DE DATOS

La teoría de bases de datos tiene años de desarrollo. Si bien se puede hablar de varias formas de administración o mantenimiento de grandes colecciones de da-

tos en computadora desde fines de los años cincuenta (anteriormente la administración se limitaba al procesamiento de los datos, más que a la consulta estructurada de los mismos), ya a fines de los años sesenta había una variedad de productos comerciales para la administración en línea, es decir, a través de terminales que respondían directamente una consulta de un usuario dado. Los modelos originales de bases de datos, basados en estructuras jerárquicas arborescentes, siguen siendo usados en, por ejemplo, sistemas de reservas de aerolíneas.

El gran salto en la administración de datos se produce con los sistemas relacionales, basados en el modelo relacional propuesto por Edgar Codd, de la IBM,

1 Los libros sobre bases de datos de ingeniería suelen ser bastante gruesos, y están dirigidos a un mercado muy específico, por lo que no son de lectura sencilla. En todo caso, el manual de Henry Korth y Abraham Silberschatz *Fundamentos de bases de datos*. 2da. ed., México: McGraw Hill 1993 es bastante bueno.

dado a conocer a través de un famoso artículo de 1970.² En lo fundamental, este modelo comparte muchos elementos con uno propuesto posteriormente, el modelo entidad-relación,³ por lo que una presentación de este último servirá como base para entender las consecuencias prácticas de ambos.

Entendemos el modelo entidad-relación como una propuesta de ver los datos como objetos del mundo real, diferenciables entre sí por sus características básicas. Una *entidad* es el caso de un objeto dado, susceptible de ser descrito por la colección de características que tienen (llamadas *atributos*), y diferenciable de otros objetos. Así, la persona llamada Eduardo Villanueva es vista como una entidad en donde sus características son, entre otras, el nombre (eduardo) y el apellido (villanueva). Tenemos que una serie de entidades que comparten elementos característicos comunes se conoce como un *conjunto de entidades*. Estos conjuntos son por definición de número ilimitado, pero de atributos limitados o limitables.

Los usuarios de una biblioteca, que comparten las características comunes de sus datos personales, sus identificadores como códigos o números de usuarios, y sus privilegios y limitaciones para usar la biblioteca, son entendidos como el conjunto de entidades *usuario*. Los documentos que pueden ser utilizados por las entidades *usuario* son a su vez el conjunto de entidades *documento*. Cuando un

usuario pide prestado un libro, establece una asociación entre él (como representante del conjunto de entidades *usuario*) y un libro o título determinado (representante del conjunto de entidades *documento*); a este tipo de asociación circunstancial entre instancias de dos o más conjuntos de entidades se le llama *relación*. De la misma manera que las entidades, una serie de relaciones con atributos comunes se conoce como *conjunto de relaciones*.

Formalmente, podemos expresar todo esto como:⁴

Una relación matemática de ≥ 2 conjuntos de entidades (posiblemente no distintos). Si E_1, E_2, \dots, E_n son conjuntos de entidades, entonces un conjunto de relaciones R es un subconjunto de:

$$\{(e_1, e_2, \dots, e_n) | e_1 : \alpha \beta E_1, e_2 : \alpha \beta E_2, \dots, e_n : \alpha \beta E_n\}$$

El modelo relacional de base de datos, actualmente el más difundido de todos los aspectos de la teoría y por tanto presente casi en cualquier sistema de cómputo medianamente desarrollado, parte de una abstracción similar, pero entiende a las entidades simplemente como datos, agrupados en *tablas*. Estas tablas tienen un número de columnas con nombres únicos, en los cuales se presentan los atributos de cada entidad. Las tablas que presentamos a continuación muestran como se establecen las relaciones, mediante el recurso de registrar en una de las columnas elementos que son compartidos entre distintos conjuntos de datos.

El vínculo entre las tablas se establece a través del dato compartido entre ambas. El diseño de estas tablas debe considerar las posibilidades de relación entre ellas y cómo se deberá realizar el vínculo. Como puede intuirse al ver gráficamente el modelo relacional, una base de datos que siga estos principios es excelente para manipular conjuntos muy grandes de registros compuestos de pequeños bloques, pero no es exactamente ideal para manejar grandes cantidades de registros con enormes bloques dentro de cada uno (es decir, para manejar campos muy grandes).

Otro aspecto de la teoría de bases de datos tiene que ver con la recuperación de información. Más allá del sistema de creación de índices, propio del modelo relacional, los distintos desarrollos en recuperación no booleana, que hasta hace poco tiempo era el dominio de especialistas muy avanzados o de sistemas muy complejos, están logrando un impacto cada vez mayor en el ámbito de Internet, tanto a través de la muy popular World Wide Web, como de la menos difundida, pero no por ello menos interesante WAIS.

Indirectamente esta combinación de experiencias, la de los catálogos automatizados y la de las bases de datos textuales, nos generan a los profesionales de la información la sensación de estar en dos realidades informáticas distintas; esta percepción está, sin embargo, bastante alejada de la verdad, puesto que la organización de datos en conjuntos buscables es un tema bastante desarrollado en la literatura matemática y de ingeniería. La separación práctica se da en el hecho concreto de que las habilidades desarrolladas por la bibliotecología no suelen permitirnos un acercamiento cómodo a las bases de datos que se alejan del modelo catálogo.

No pretendemos en este artículo mayor detalle en la teoría misma de las bases de datos, sino en las ventajas y desventajas de conocer estas teorías, y en la repercusión que las actuales tendencias en diseño de bases de datos están trayendo al mundo de la información no mediada por bibliotecarios, que es la gran mayoría de información disponible por Internet y

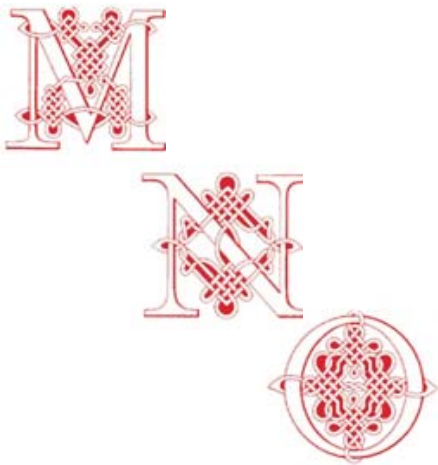
Nombre	Apellido	Código de alumno	Nivel de privilegio
Eduardo	Villanueva	983.222.3.22	A
Liliana	Bandini	984.2.322.33	B
Patricia	Naka	839.3.225.77	A
Guillermo	Vásquez	904.322.5.44	C

Nivel de privilegio	Cantidad máxima de libros	Cantidad máxima de días por libro
A	4	5
B	3	5
C	3	4

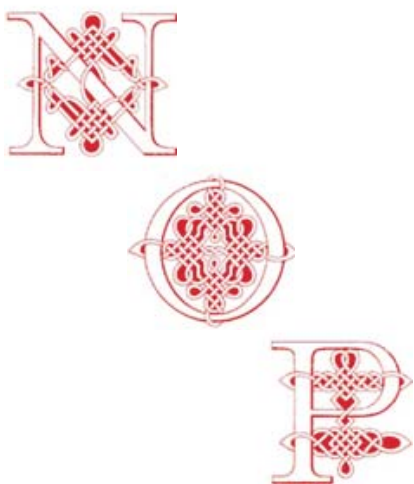
2 E.F.Codd "A relational model for large shared data banks" / . en: *Communications of the ACM*, 13(6), junio 1970, pp.377-387

3 P.P. Chen. "The entity-relationship approach to logical data base design" en: *ACM transactions on database systems*, 1(1), enero 1976. pp.9-36.

4 Cfr. Korth y Silberschatz, en *Op. Cit.* p. 27.



“Mientras la biblioteca toma como centro la representación exacta de entidades únicas (los libros y todas las variantes de los documentos), la base de datos habitual fuera de la biblioteca expresa las distintas iteraciones de entidades casi platónicas”



proveedores comerciales de bases de datos.

Hay pues que revisar las dos grandes diferencias entre las bases de datos como las conocemos y las bases de datos como las entienden los informáticos: la primera es la oposición entre catálogo estático/conjunto transaccional dinámico. La segunda es cuando la información ya no sólo es una reunión de representaciones: estamos hablando de miles de bases de datos que contienen en sí mismas tanto la información documental, como la información para recuperar esos documentos. Comúnmente se llama a estos sistemas texto completo.

PRIMERA OPOSICIÓN: LA CLÁSICA BASE DE DATOS CONTRA MODELOS RELACIONALES Y DISTRIBUIDOS

La representación contra la información

Una biblioteca es un gran repositorio de información, con un inmenso mecanismo intermediador: el catálogo. Por razones diversas, esos “elefantes blancos”⁵ han sido el gran centro de la atención del bibliotecario no durante años, sino durante siglos. Y con justicia, puesto que son una ayuda indispensable aun en la biblioteca de completa estantería abierta.

El catálogo es una representación de una biblioteca, o más propiamente, la reunión de una colección de representaciones, los registros; mismos que tradicionalmente han tomado la forma de fichas, cuando están en una computadora se les conoce como registros. Estas representaciones tratan de acercarnos a los contenidos de los documentos, describiéndolos tanto como ordenándolos de las diversas formas posibles para el usuario potencial o ideal. Las distintas maneras de organizar estas representaciones, sea en catálogos de entrada por ítem o por término⁶ no cambian la esencial verdad del catálogo.

Por cierto, ésa no es la única información que la biblioteca guarda o necesita, pero sí es la principal. La segunda en impor-

tancia es la información sobre los usuarios de la biblioteca, quiénes pueden utilizar los documentos en ella almacenados, y cómo pueden hacerlo.

Este tipo de información es quizá el mejor ejemplo de las *entidades* que interesan a un ingeniero en sistemas. El término entidad no es gratuito, puesto que a diferencia de un catálogo de biblioteca, una base de datos habitualmente es una colección de datos que sólo existe en el sistema, sin la intención de reflejar entes realmente existentes. El detalle es que una base de datos es vista como la descomposición en elementos descriptivos de una entidad, es decir, de un objeto tipo que estará presente en varias cantidades (y con ello, en variantes diversas) en la base de datos. Nadie pretende reducir a un alumno universitario a la reunión de sus datos, o a un cliente bancario a sus estados de cuenta, o a unas ocho o diez papas a algo tan ilusorio como un kilo de papas; en todos los casos, son abstracciones producidas en el hecho mismo de reunir y manipular los datos que entran en una transacción.

Mientras la biblioteca toma como centro la representación exacta de entidades únicas (los libros y todas las variantes de los documentos), la base de datos habitual fuera de la biblioteca expresa las distintas iteraciones (repetición de acciones análogas) de entidades casi platónicas, como puede ser un usuario, un pasajero de avión, un cliente o el ya mencionado kilo de papas. Las iteraciones existen en tanto son creadas a partir o en vistas de una transacción, es decir, una operación en donde esta variante específica de la entidad toma contacto con otra entidad para ser modificada o para crear o desaparecer otras entidades. No son casos más o menos intangibles, sino más bien móviles: *mobilis in mobile*, parafraseando a Verne.⁷

Por ello, la necesidad de entender los principios básicos de la organización de esa otra información no se puede separar tanto la aplicación de la tecnología frente a la aplicación de los principios técnicos

5 Cfr. Maurice Line y M. Grose. “On the construction and care of white elephants: some fundamental questions concerning the catalogue”.— En: *Lines of thought : selected papers of Maurice Line*. Londres: Bingley, 1988. pp. 4-11.

6 Ver el ya clásico artículo de Dagobert Soergel “A general model for indexing languages: the basis for compatibility and integration”, en *Subject retrieval in the seventies: new directions*. Editado por Hans Wellisch. Westport: Greenwood, 1972. Aquí se establece que el catálogo automatizado y el catálogo manual son modelos distintos de organización, basados en principios de entrada de registro distinto.

7 *Vingt-mille lieues sous les mers*, cap. VIII.

de la propia profesión; una realidad contradictoria, pero que siempre parece estar a punto de impedirnos la comunicación con los informáticos.

Resumiendo: la base de datos habitual es un conjunto de iteraciones de una entidad que se relaciona con iteraciones de otras entidades en transacciones, es decir, en interacciones ocasionales. Por ello es necesario el lenguaje de programación de una base de datos; ya que nos permite determinar de qué manera se van a registrar las interacciones entre las *tablas*, que es el nombre que habitualmente reciben las reuniones de iteraciones de una entidad. Lo que solemos llamar base de datos, es para el ingeniero una simple tabla dentro de un sistema de administración de bases de datos, o DBMS.

Para el ingeniero en sistemas, el catálogo/base de datos es una tabla más que debe ser manipulada tanto en sí misma como en relación con las demás tablas que hay en el sistema. No tiene sentido detenerse mayormente en una tabla (el catálogo), sino que la preocupación es más bien cómo combinar la tabla *x* con la tabla *y*, determinando las correctas relaciones y condiciones de transacción. Nada más alejado de la actitud habitual de los bibliotecarios, para quienes *esa* tabla es el centro de su trabajo y actividad.

Una gran ayuda: los ISBDx

El catálogo de una biblioteca es pues la reunión de descripciones básicamente intangibles de una sola entidad, que existe con varias encarnaciones. Estas encarnaciones están estudiadas, documentadas y casi diseccionadas en los distintos documentos que sirven como estándares para el trabajo de catalogación, clasificación e indización de documentos. Fundamentalmente, los ISBDx (Descripción Internacional Bibliográfica Estándar) cumplen la función de ahorrarnos reinventar la rueda a cada momento.⁸

En los ISBDx, y en sus encarnaciones específicas, como las AACR2 podemos ver la detallada descripción de la *entidad documento* tal y como las entiende la profesión bibliotecaria. Desde el momento que no se ha terminado con la preparación de los ISBD específicos para muchos tipos de material, y que los manuales ya publicados están en permanente revisión por los especialistas en el tema, no podemos asumir que todo lo que se puede hacer en el tema de la descripción de la entidad documento esté dicho.

Tampoco debemos dejar de lado el tema de la aplicación de los ISBD en circunstancias y ambientes específicos, ni tampoco el tema del intercambio de los datos descritos a través de los ISBD; todo esto en conjunto forma parte del siempre relevante tema del **Control Bibliográfico Universal**, sobre el cual hay mucho que hacer y que decir.⁹

Lo que no podemos obviar es que, como centro de la atención del bibliotecario, el tema de la descripción es de menor importancia, puesto que ya está bastante trabajado, y que el mayor esfuerzo debe recaer en la manera como garantizar una recuperación de información realmente útil para el usuario final.

Recuperación y bases de datos

Las bases de datos surgen como una alternativa de organización para cantidades masivas de información. Sobra decirlo, pero cualquier gran conjunto de información ha sido recolectado para ser recuperado. Las formas de recuperación de un catálogo automatizado también difieren de las formas de recuperación de las bases de datos no documentales.

Mientras que la prioridad de los bibliotecarios es la determinación tanto de puntos de acceso, sean estos derivados o asignados,¹⁰ así como el control de la terminología por usarse a partir de dichos puntos de acceso, el trabajo del diseñador de un sistema de base de datos es más bien la deter-

minación tanto de los índices relevantes para la recuperación en cada tabla, como de las combinaciones de índices de tablas distintas para realizar cruces relevantes de información; en sistemas de múltiples tablas, estamos hablando de un trabajo muy complicado. El sistema no documental de bases de datos no tiene tantos conflictos con la terminología, mientras que el sistema documental necesariamente tiene que preocuparse de manera fundamental por la terminología.

Aquí podemos pues destacar un primer aspecto de diferenciación entre la tradición bibliotecaria y la tradición de los sistemas de información, en el tratamiento de las bases de datos: el control terminológico es y seguirá siendo un factor importantísimo para cualquier base de datos que trabaje con documentos, donde el valor de la información yace no en las interacciones de los datos sino en los documentos representados o accesibles a través del sistema. A pesar de cualquier atajo que se quiera tomar, es imprescindible contar con intervención humana en la elaboración ya sea de las listas de términos o bien en la misma asignación de éstos.¹¹ Este aspecto del trabajo de información está tomando más importancia en la medida en que las colecciones de documentos computarizados aumentan cada día, como pasa en Internet.

SEGUNDA OPOSICIÓN: CATÁLOGO CONTRA TEXTO COMPLETO

El impacto de Internet

El ciberespacio no es nada nuevo ya. La enormidad de información disponible en miles de servidores en todo el mundo están desde hace tiempo en boca de todos, y ciertamente la creación de mecanismos para acercar los acervos de las bibliotecas a la gran red mundial es una prioridad para la gran mayoría de los bibliotecarios del mundo. La idea de biblioteca digital o

8 Cfr. Dorothy Anderson, *Standard practices for bibliographic description*.

9 Sobre el concepto de UBC y su desarrollo, en la actualidad, una excelente fuente es la página Web de IFLA; especialmente en <http://www.nlc-bnc.ca/ifla/VI/3/ubcim.htm>. También para la situación del Control Bibliográfico Universal en América Latina, ver *Seminar on Universal bibliographic control: papers*. Río de Janeiro: IFLA UBC, 1993.

10 Estamos siguiendo la terminología de A.J. Foskett, propuesta en su clásico trabajo *The subject approach to information*. 4ta.ed., Londres: Bingley, 1985, especialmente en los capítulos 1 y 2.

11 Un interesante repaso sobre el potencial de intervención computarizada en la elaboración de índices de bases de datos documentales se presenta en el artículo de revisión crítica de tendencias "Methodologies for subject analysis in bibliographic databases" / Jessica Milstead, 1 *Information Processing and Management* 28(3), pp.407-432.

virtual es inclusive tema de conferencias y congresos.¹²

Si tuviéramos que definir la biblioteca virtual, podríamos decir que es la unión entre el documento y su representación, el catálogo y la colección unidos, liberados de la asociación necesaria entre ubicación en un lugar concreto y acceso en un lugar concreto. La biblioteca virtual sería un acceso por búsqueda de información no a un catálogo, sino a los documentos mismos, y a través de un sistema que nos permita buscar desde cualquier punto del planeta información almacenada en todo el planeta.

¿Suena conocido? Lo cierto es que tenemos en Internet dos grandes etapas: la primera es de acceso a datos o aplicaciones (FTP, USENET, Telnet). La segunda es la que ha hecho que Internet sea lo que es ahora, acceso a información creada por y para la red misma: Gopher y WAIS, pero sobre todo la World Wide Web. La información de la que se dispone está contenida en computadoras, y la recuperación de esa información debe ser a través de la computadora. Nos hemos liberado del lugar de consulta y del lugar de acceso, y la recuperación es de documentos, no de representaciones.

Si bien la WWW no es precisamente muy buena herramienta de recuperación, así es usada por la mayoría de cibernautas en la actualidad. La gran difusión de ciertos revisadores de la WWW, como el Netscape, ha permitido que se popularicen los CGI, es decir, los programas que sirven de interfaz entre una página de la WWW y una base de datos; basta que el revisador sea capaz de manejar formas, es decir, páginas en donde se llenan espacios, para que a través de un CGI se pueda interrogar amigablemente una base de datos de mucha complejidad.¹³ La WWW no es una herramienta de recuperación, sino de organización hipertextual, pero este atajo permite eludir la necesidad de

ofrecer a usuarios inexpertos acceso a un sistema que probablemente requiera mucha familiaridad. Un primer avance a la biblioteca virtual, un nuevo dilema para los bibliotecarios. ¿Podremos ofrecer al mundo algo de utilidad cuando no haya que hacer catálogos?

Las limitaciones de la recuperación y ciertos cambios previsibles

Como explica Michael Koenig, casi podríamos decir que el modelo más conocido de interrogación de bases de datos, el basado en la lógica booleana, tiene un predominio más fundado en la casualidad que en la calidad.¹⁴ Desde la década de 1970 muchos especialistas han tratado de proponer y de imponer modelos más potentes o al menos más flexibles, como por ejemplo el que usaba LEADERMART, basado en el principio de asignar pesos a los términos significativos presentes en resúmenes, de manera que al hacerse una búsqueda en lenguaje natural, el sistema recuperara los artículos en donde los términos propuestos aparecieran en mayor cantidad. Esfuerzos unificadores han habido muchos, y propuestas para reformular la recuperación de información desprendiéndola de sus limitaciones booleanas hay muchas.

Mientras el acceso a servicios de información tanto hechos localmente (catálogo de biblioteca) como comerciales (bases de datos de acceso remoto) era excepcional, debido al costo de implementación o de suscripción, los usuarios estuvieron dispuestos a aprender interfaces complicadas o a utilizar los servicios mediados por los bibliotecarios. La popularización de las microcomputadoras, el abaratamiento del acceso a la información gracias, entre otros factores, al CD-ROM y finalmente la generalización de las interfaces gráficas de usuarios, con su promesa no siempre cumplida de facilidad de uso, han hecho más exigente

al usuario común, que espera poder acceder directamente a la información, a través de aplicaciones tan sencillas de usar como su procesador de textos o su hoja de cálculo para Windows o Macintosh. Frente a ello, una interfaz no gráfica tiene que simplificarse de manera integral o ceder su sitio a una gráfica, y la lógica booleana tiene que ser escondida detrás de botones y similares.¹⁵

Sin embargo, es a través de Internet que las limitaciones tanto del modelo de búsqueda libre como el de conectores booleanos están siendo consideradas en busca de alguna mejoría, tanto en las cuestiones de fondo como en la apariencia que ofrecen para la manipulación del usuario. De hecho, la experiencia más significativa de recuperación de información en Internet es endógena de la red misma: los buscadores de información de Internet son herramientas para ubicar tanto páginas de World Wide Web como artículos de grupos USENET, o inclusive archivos de FTP o Gopher. Los buscadores de información usan técnicas muy sencillas de recuperación de información y por ello ofrecen resultados de calidad bastante limitada, aunque como descargo podemos decir que los documentos fuente tampoco son muy consistentes con ningún modelo habitual de indización, y la ya mencionada poca disposición desde los mismos orígenes que tiene la WWW a ser realmente una fuente organizada de información. Otro elemento que limita la capacidad de recuperación de estos buscadores es la dependencia de los *metatags*, es decir de indicadores puestos en la misma página por el autor o autores, que indican los contenidos relevantes para la búsqueda; como puede imaginarse, no todos los autores tienen las habilidades para proponer metatags adecuados para todas las circunstancias.

Por ello sólo se puede pensar en los buscadores como paso previo a una industria

12 Cfr. la 61ª Conferencia General de IFLA, realizada en Estambul en 1995, que trató precisamente sobre la Biblioteca Virtual.

13 "The main idea is that you have a web page that has a link to what looks like another web page, but is really a program (in our case a C program) that dynamically figures out what HTML to display..." en: *CCS 15 CGI Information*, en: <http://hamp.hampshire.edu/~agCCS/cs/cgi-stuff.html>

14 Michael E.D. Koenig. "How close we came" / en: *Information Processing and Management*, 28(3), 1992. Koenig también hace un interesante énfasis en lo casual que ha sido el desarrollo de los sistemas de interrogación de bases de datos tal como los conocemos.

15 Por ejemplo, Silver Platter ofrece dos versiones de su SPIRS (Silver Platter Information Retrieval System), una no gráfica que demanda bien poco de la computadora en la que reside el CD-ROM, y una gráfica, bastante sencilla de usar, pero que requiere computadoras más complejas. Inclusive el prototipo desarrollado en Italia de una versión Windows del micro ISIS utiliza botones para seleccionar todas las opciones booleanas y complementarias en la formulación de búsqueda.

más sólida de recuperación de información que, en el futuro, actúe de una manera más planificada.¹⁶ Por ahora, los buscadores no son exactamente muy confiables, aunque cada vez mejoren más y ofrezcan un enorme potencial de ayuda para la navegación.¹⁷

Es imposible aún determinar cuál será el nuevo estándar. Lo cierto es que la presión hacia el manejo de documentos antes que de representaciones, y con ella la reformulación de la biblioteca hacia su encarnación virtual o digital nos propone una alternativa bastante más compleja, y más interesante. El desafío en el camino es que la necesaria coexistencia de bases de datos representativas y bases de datos documentales llevará a la necesaria integración entre ellas o por lo menos a la creación de puentes que impidan que el usuario viva disociado entre realidades casi contradictorias.¹⁸

El trabajo de información es especialmente cuidadoso con el tratamiento terminológico, y cualquier desarrollo útil del ciberespacio pasará necesariamente por la intervención de especialistas en terminología para hacer bases de datos más consistentes y fáciles de buscar, cosa que el proceso de búsqueda de tesoros en medio del fango permita que sea más rápido y constructivo.

UN POCO MÁS DE CIENCIA DE LA INFORMACIÓN: REFLEXIONES PARA REPLANTEAR LA FORMACIÓN PROFESIONAL

El desarrollo de bibliotecas virtuales parece ser, ahora sí, más que una simple inflación de expectativas. La inmensa va-

riedad de servicios gratuitos y comerciales de acceso a información en Internet y en otros medios, como puede ser el CD-ROM, nos lleva a pensar en que las instituciones con necesidades de información documental tendrán cada vez más alternativas de donde escoger. Tanto el medio impreso, que necesariamente requiere del catálogo, como el medio electrónico, que integra al catálogo en sí mismo, tiende a complementarse y probablemente a crear cierto traslape en determinadas áreas tanto de conocimiento como de edición, como el caso ya mencionado de las revistas electrónicas. Y aunque es todavía una preocupación a mediano plazo, si las tendencias continúan siendo vigentes, algún tipo de interacción habrá que facilitarle al usuario.

De hecho, la gran mayoría de aplicaciones comerciales para automatización de bibliotecas tiende a ofrecer entre sus módulos opcionales pasarelas, entre el sistema local y sistemas no locales, estén disponibles en CD-ROM o a través de Internet.¹⁹ Estas funcionalidades son aún caras y opcionales, pero sabremos que la demanda es mayoritaria cuando se conviertan en una función estándar, como ahora lo es el catálogo de acceso público o el módulo de autoridades.

De todo esto, podemos concluir que la relación con las bases de datos tenderá a hacerse cada vez más compleja: no bastará una simple comprensión de cómo crear formatos o administrar respaldos (*backups*) diarios, sino que se hará imprescindible entender las diferencias entre una gran diversidad de propuestas de acceso, que tendrán entre sí también una

gran diferencia de estructura y de capacidad de uso para el bibliotecario y el usuario final.

Una biblioteca tendrá que integrar, de manera rápida y sin conflictos, variedades de información. La imagen para el usuario final puede ser sencilla, una interminable colección de páginas hipertextuales, a la manera de World Wide Web.²⁰ Pero lo que no escapa es que mientras la interfaz se simplifica, las estructuras internas no lo hacen, y comprender cómo interrelacionar una enorme cantidad de bases de datos dispares puede ser tarea tan titánica que se deje de lado. El abandono de la intención de entender las cosas puede ser equivalente a dejarlas en manos de otros, con prioridades y premisas diferentes a las nuestras.

Por todo esto, el camino parece inevitable: coexistencia pacífica (*entente cordiale*) entre mundos radicalmente distintos. El administrador de los puentes será el bibliotecario, tanto como ha sido el administrador de la intermediación entre el conocimiento y los medios de difusión. Cada innovación replantea el modelo clásico de trabajo, y requiere no tanto manejo técnico, sino comprensión de la complejidad y capacidad para administrarla.

No hay cambios en la función final, pero sí en sus varias encarnaciones. Como diría Dickens, *the best of times, the worst of times*.²¹ Antes que nada, adaptarnos sin perder nuestra mayor virtud: pensar en el usuario, antes que en la tecnología.

16 Las grandes empresas de información documental, como Silver Platter y EBSCO, ya se están planteando estos problemas. Una visita a sus websites puede ser ilustrativa: Silver Platter es <http://www.silverplatter.com/> y EBSCO es: <http://www.epnet.com>. También las editoriales están trabajando estos temas, como una visita a Elsevier puede indicar (<http://www.elsevier.com/>).

17 Cfr. el artículo de Nicholas G. Tomaiuolo y Joan G. Packer "Quantitative analysis of five WWW search engines" en: *Computers in Libraries*, volume 16, issue 6, June 1996. También puede revisarse parcialmente en: <http://neal.ctstateu.edu:2001/htdocs/websearch.html>

18 Una interesante discusión sobre estos temas puede verse en la ponencia de Graham Black "Integrating the digital with the non digital: a librarian's perspective" Manuscrito.

19 Cfr. Barry, Jeff, José-Marie Griffiths y Peiling Wang. "Automated system marketplace: jockeying for supremacy in a networked world". En: *Library journal*, abril 1, 1996. pp. 40-51.

20 Existen varios indicios de que los líderes en la creación de software están considerando el crear nuevas interfaces de usuario a la manera de la interfaz de la WWW: recientemente, Bill Gates, dueño de Microsoft, ha anunciado que la próxima gran propuesta de interfaz de su compañía será una interfaz WWW, que hará irrelevante la diferencia entre Internet, una red interna y la computadora del usuario. Cfr. *PC/Computing*, abril 1996, p.87.

21 A tale of two cities, p.1.