

# Aplicación del proceso de *KDD* en el contexto de *bibliomining*: El caso Elogim

---

## Resumen

En la actualidad, organizaciones de toda índole, incluidas las bibliotecas y unidades de información, demandan herramientas informáticas que faciliten el desarrollo de actividades tanto técnicas como administrativas. En este artículo se presenta ELOGIM, una aplicación informática desarrollada en la Biblioteca Fundadores, de la Universidad CES Medellín, como una experiencia exitosa en la aplicación del modelo de descubrimiento del conocimiento en bases de datos – *KDD* – en el contexto de *bibliomining*, que elabora informes estadísticos sobre el uso de los recursos electrónicos y provee información determinante para la toma de decisiones y la gestión en la unidad de información, razón de que sea elemento clave en propiciar la gestión del conocimiento en la comunidad académica universitaria.

**Palabras clave::** gestión del conocimiento; descubrimiento del conocimiento en bases de datos, bibliominería, gestión de unidades de información, gestión de información, minería de datos.

**Cómo citar este artículo:** QUIROZ GIL, Nora Ledis y VALENCIA, Carlos Andrés. Aplicación del proceso de *KDD* en el contexto de *bibliomining*: el caso Elogim. Revista Interamericana de Bibliotecología. 2012. Vol 35, n° 1, pp. 97-108.

KDD application process in the context  
of bibliomining: Elogim case

## Abstract

Nowadays, organizations of all kinds including libraries and information units require tools to facilitate the development of both technical and administrative activities. This article presents ELOGIM, a tool developed at the Fundadores Library of CES University (Medellín, Colombia), as a successful experiment in the application of the knowledge discovery in databases through *bibliomining* context, to prepare statistical reports on the use of electronic resources and provides information to assist decision-making and management within the information unit, which becomes a key element to promote knowledge management in the university academic community.

**Key words:** knowledge management, knowledge discovery databases, *bibliomining*, management of information units, management information, data mining.

**Nora Ledis Quiroz Gil**  
Jefe de Biblioteca Fundadores,  
Universidad CES. Medellín,  
Colombia.  
nquiroz@ces.edu.co

**Carlos Andrés Valencia**  
Analista de Biblioteca. Universidad  
CES. Medellín, Colombia.  
cvalencia@ces.edu.co

## Introducción

El crecimiento constante e incontenible del volumen de la información obliga a las organizaciones a definir y aplicar métodos que permitan percolar el conocimiento de las grandes bodegas de datos (*data warehouse*) generados en cualquier ejercicio industrial o científico que involucre tecnologías (como la astronomía, en la clasificación de las estrellas y galaxias; la administración y el mercadeo, en la definición de perfiles de clientes y programas de fidelización; las finanzas, en las transacciones en bolsa y la detección de fraude; de igual manera en los deportes, en los juegos de azar, en la bibliotecología y, en general, en cualquier actividad humana donde estén presentes los datos), y a plantear su aplicación desde perspectivas globales y no parciales (Riquelme, Ruiz, & Gilbert, 2006). El estudio a conciencia de la información obtenida con base en los datos, al proporcionar información útil sobre tendencias o directrices, proyectivas y prospectivas que den argumento a las decisiones organizacionales, da pie a pensar que la evolución de las organizaciones es posible.

Las bibliotecas y unidades de información no son ajenas a la irrefrenable marea de datos pero tampoco han descubierto cómo generar conocimiento a partir de ellos, potencializarlos y aprovecharlos para optimizar presupuestos, renovar o crear nuevos servicios y definir indicadores de gestión que sustenten decisiones estratégicas.

Con el término “descubrimiento de conocimiento en bases de datos”, Knowledge Discovery Databases o KDD, fue designado el proceso que pretende obtener conocimiento a partir de datos almacenados en bodegas (*data warehouse*). En dicho proceso se incluyen la preparación de los datos, el análisis estadístico, el algoritmo usado para la minería de datos y la evaluación e interpretación de los mismos, obteniendo como resultado el descubrimiento de conocimiento. Según Fayyad *et al.* (2002), el KDD es “el proceso no trivial de identificar patrones válidos, novedosos, potencialmente útiles y comprensibles a partir de los datos.”

Uno de los elementos del proceso KDD tiene que ver con la minería de datos, término usado por Witten *et*

*al.* (2011), como el proceso del cual se obtiene nuevo conocimiento, útil y comprensible a partir de grandes volúmenes de datos almacenados en distintos formatos. Para Vieira *et al.* (2009), es necesario, por el conocimiento adquirido sobre los clientes, ser capaz de interpretar sus objetivos, expectativas y deseos. Esto se consigue con la *minería de datos* o con la *minería de datos dirigida al cliente* (*customer centric data mining*), una colección de técnicas y métodos que facilitan la adquisición y retención de la parte del mercado que cabe en una empresa (*market share*). Las metas de atención y reducción de costos también valen para las organizaciones no lucrativas, gubernamentales, o no.

En el ámbito de las bibliotecas se acuña el término “*bibliomining*” (Riquelme, Ruiz, & Gilbert, 2006), para hacer referencia a la técnica de minería de datos que procesa la información generada en ellas. El *bibliomining* tiene una estrecha relación con la bibliometría y la documentación (Banerjee, 1998), (Michail, 1999), (Mancini, 1996), en el sentido de que aporta a la medición y el análisis de datos mediante métodos estadísticos, en este caso, para ayudar a definir tendencias o patrones de uso de los recursos electrónicos en bibliotecas y unidades de información. (AENTA, 2011)

Teniendo en cuenta las definiciones de Witten *et al.* (2011) y Fayyad *et al.* (2002), es necesario aclarar que minería de datos y KDD no hacen referencia al mismo concepto; el último circunscribe elementos más profundos, como la evaluación y la interpretación de los datos para, finalmente, encontrar el conocimiento.

El KDD, como proceso, involucra una secuencia de etapas claramente definidas, cada una fundamental para la transición de los datos en conocimiento. Para Vieira *et al.* (2009), todos los sistemas de KDD mantienen la misma esencia, *la minería de datos*; su factor diferenciador radica en la implementación y la presentación. Todos transitan las mismas etapas: recolección, depuración y análisis de datos, de donde se obtiene como resultado un “modelo descriptivo” que puede ser convertido en un “modelo predictivo”, de ser necesario. En la figura 1 se presenta el esquema de las etapas del proceso de descubrimiento del conocimiento en bases de datos KDD.

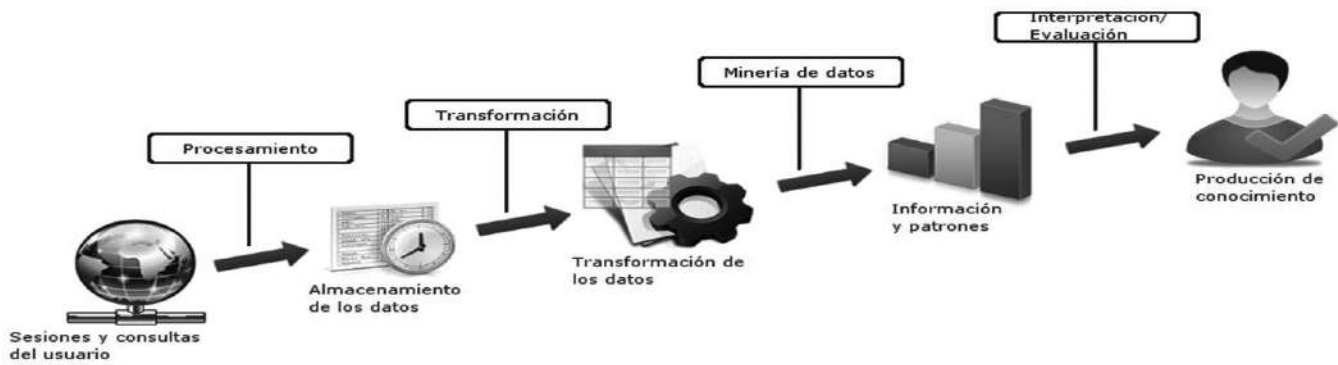


Figura 1. Esquema de las etapas del proceso de descubrimiento del conocimiento en bases de datos. (KDD)  
(Construcción propia)

Fundamentado en el modelo de *KDD*, la Biblioteca Fundadores, de la Universidad CES (Medellín), desarrolla ELOGIM, una aplicación informática concebida a través de recursos open source, que automatiza el proceso de almacenamiento y recuperación de registros estadísticos con respecto a las sesiones y descargas efectuadas en recursos electrónicos (bases de datos bibliográficas, libros electrónicos, repositorios digitales, catálogos en línea, etc.), permitiendo obtener estadísticas por criterios, y variables como facultad, programa, tipo de usuario o un usuario en particular, facilitando la administración de los recursos y servicios electrónicos en las unidades de información y dando lugar a informes estadísticos y reportes que derivan en indicadores de gestión (Contreras Contreras, 2005) para el apoyo en la toma de decisiones. Al ser concebido como herramienta web, basa su arquitectura cliente – servidor en recursos de acceso abierto y es desarrollado en lenguaje PHP – HTML y Java Scrip, en donde el servidor web es APACHE y el gestor de bases de datos es MySQL.

Para la elaboración de informes estadísticos, ELOGIM utiliza los datos almacenados en el archivo (.LOG) del EZProxy (software de validación y acceso a los recursos electrónicos de Online Computer Library Center – OCLC-) (OCLC, 2012), originados en el tráfico de sesiones de los usuarios al ingresar a cada recurso electrónico ofrecido por la biblioteca. Este archivo (.LOG) (Peter, 1996) puede ser configurado y parametrizado de acuerdo con la necesidad de información, y es a partir de él como se toman los datos, se crean los registros y posteriores reportes estadísticos, convirtiéndolos, mediante el cruce de variables, en información concreta como:

1. Ranking de recursos electrónicos: Listado de los recursos electrónicos de mayor uso.
2. Ranking de uso de los recursos electrónicos por usuarios, facultad y programa: listado de personas que más usan los recursos electrónicos y que puede ser discriminado por tipo de usuario, facultad o programa.
3. Efectividad en el recurso electrónico: listado comparativo entre las sesiones iniciadas y los documentos descargados.
4. Reconocimiento de direcciones IP de donde se accede a los recursos: el rastreo hace posible el seguimiento y control de las claves de acceso a los recursos electrónicos.
5. Periodos de acceso: definición de horas o periodos de tiempo con mayor tráfico en búsqueda y descarga de información por parte de los usuarios de los recursos electrónicos.

## Estructura

### ¿Cómo y de dónde se toman los datos?

ELOGIM está diseñado y estructurado para interoperar con la aplicación EZProxy, permitiendo la realización de rutinas de importación, procesos conjuntos y entrelazados, lo que da como resultado el funcionamiento integrado que hace de estas dos herramientas un mismo sistema, partiendo de los datos almacenados por EZproxy; funciona en forma paralela, facilitando

la convergencia entre ambas aplicaciones, sin que, en ningún momento, ELOGIM entre en conflicto con ellas o modifique los datos fuentes del código de EZproxy. La figura 2 presenta el modelo relacional de ELOGIM, donde se muestra el proceso de interacción desde que el usuario inicia la sesión en EZProxy hasta la generación de estadísticas y reportes.

ELOGIM utiliza varios filtros para la generación de los reportes, de la siguiente forma:

**Filtro 1. Seleccionar variables:** Son aquellos campos susceptibles de ser combinados por la aplicación. En el caso del diseño a la medida de ELOGIM para la Biblioteca Fundadores de la Universidad CES, se definieron como variables:

- **Facultad/institución:** Esta variable está compuesta por el listado de las facultades existentes en la Universidad CES y las instituciones adscritas que tienen acceso a sus recursos electrónicos.
- **Programa:** Esta variable está definida por los tipos de programas que ofrece la Universidad CES a través de sus facultades, como doctorados, maestrías, posgrados y pregrados establecidos.
- **Componente:** Determinado por el nombre específico de cada programa académico ofrecido por la Universidad. Ejemplo: Biología, Tecnología en Atención Prehospitalaria.
- **Tipo de usuario:** Clasificación que define la relación del usuario con la Universidad (estudiante, egresado, empleado, docente, externo y retirado).

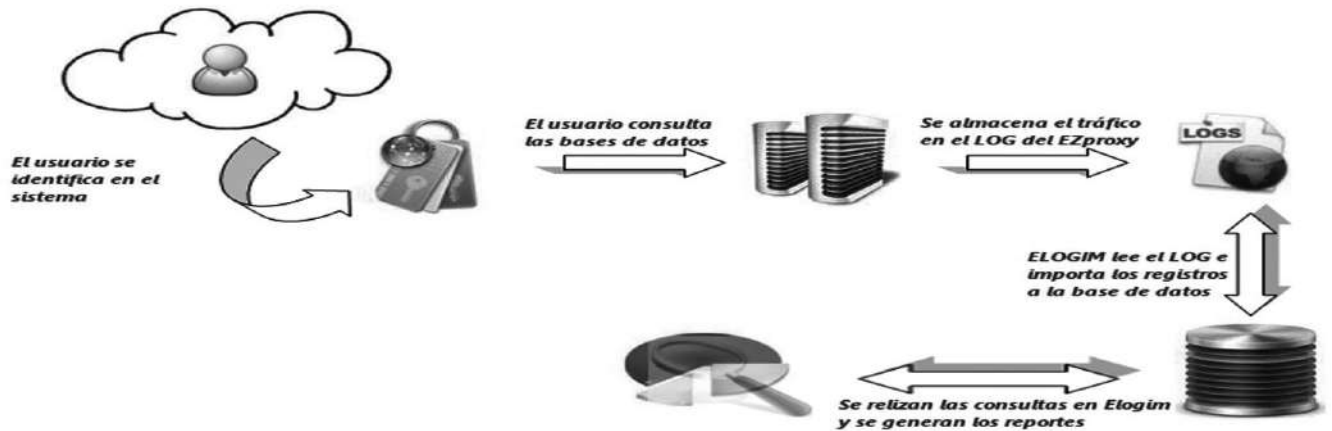


Figura 2. Modelo Relacional del proceso de ELOGIM (Construcción propia)

Figura 3. Representación gráfica de la selección de variables según el filtro 1.

En la figura 3 se observa el proceso de selección de variables a combinar, definido en una consulta según el filtro 1.

Los datos de estas cuatro variables están tomados directamente de FÉNIX (Sistema de Administración Bibliotecaria de la Biblioteca Fundadores), el cual importa los registros desde SARA (Sistema de Admisiones y Registros Académicos de la Universidad CES), y de sistemas que alojan los datos, como el registro de nómina (Midasoft), y el registro de los egresados. En este caso, ELOGIM solo consulta en las tablas de usuarios relacionadas con la base de datos FENIX; sin embargo, conserva el modelo establecido por el sistema SARA.

- **Recurso electrónico:** Esta es la única variable administrada directamente desde ELOGIM; en ella se definen los recursos electrónicos de que dispone la unidad de información, de los cuales se hace necesario generar estadísticas de uso e indicadores de gestión.

**Filtro 2: Tipo de resultados:** Este menú permite seleccionar el resultado y tipo de tabulación que se desea obtener en la consulta, definiendo seguidamente un rango de fechas. En la figura 4 se presenta la forma como se definen los resultados según el filtro 2.

Los tipos de resultados que se pueden obtener de la consulta son:

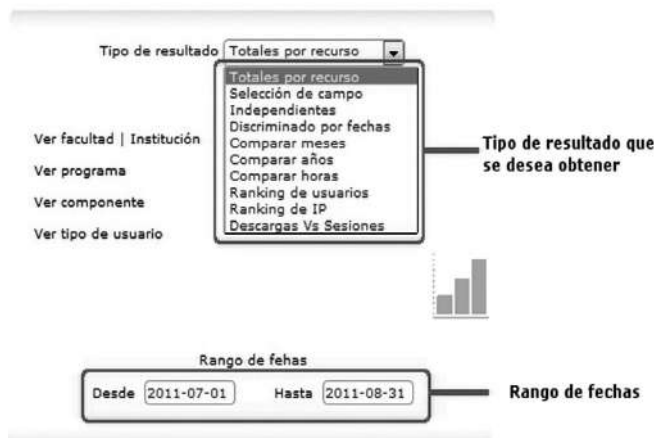


Figura 4. Representación gráfica de la selección de “Tipos de resultados” según el filtro 2.

- Totales por recurso:** Muestra solo el total por cada uno de los recursos digitales indicados en la consulta.
- Selección de campo:** Permite seleccionar una de las variables de las que dispone el ELOGIM, para visualizar los resultados por facultad, programa, componente o tipo de usuario.
- Independientes:** Este tipo de resultado muestra los registros relacionados en la consulta de manera independiente. En esta opción se visualiza cada campo del registro, lo que permite identificar datos como: fecha y hora del registro, usuario que efectuó la consulta y detalles particulares, como IP, facultad, recurso consultado.
- Discriminado por fechas:** Muestra el total de cada recurso, discriminado día por día, dentro del rango de fechas especificado.
- Comparar meses:** Muestra el total de cada recurso, discriminado mes por mes, dentro del rango de fechas especificado.
- Comparar años:** Muestra el total de cada recurso, discriminado año por año, dentro del rango de fechas especificado.
- Comparar horas:** Muestra el total de cada recurso, midiendo y agrupando los totales en cada hora del día, dentro del rango de fechas especificado.
- Ranking de usuarios:** Efectúa una consulta que entrega como resultado un listado, en orden de mayor a menor, de los usuarios que más usaron el recurso, de acuerdo con el criterio y variables de la consulta.
- Ranking de IP:** De acuerdo con el criterio y las variables de la consulta, arroja, en orden de mayor a menor, las direcciones IP desde las cuales se ingresó a los recursos.
- Descargas vs. sesiones:** Compara el total de sesiones iniciadas, con las descargas efectuadas en cada uno de los recursos digitales.

**Filtro 3. Menú para definir la presentación de resultados:** Este menú permite configurar el estilo o presentación



del tipo de resultado seleccionado. En la figura 5 se observa el filtro 3, en el cual se define la presentación de resultados. A continuación se define cada opción del menú.

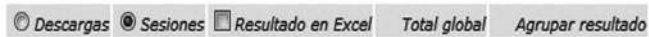


Figura 5. Filtro 3: Menú para definir la presentación de resultados

- a) **Descargas:** Según los criterios de búsqueda, se presentan como los resultados de las descargas efectuadas en el recurso.
- b) **Sesiones:** Según los criterios de búsqueda, se presentan como los resultados de las sesiones efectuadas en el recurso.
- c) **Ver en Excel:** Exporta el resultado de la consulta a un formato Excel.

d) **Total global:** Presenta el resultado, con los totales generales sin discriminar por recursos, de acuerdo con el criterio y las variables de la consulta (solo aplica para el tipo de resultado *Selección de campo*, *Discriminado por fechas*, *Comparar meses*, *Comparar años* y *Comparar horas*).

e) **Agrupar resultado:** Presenta los resultados, discriminados por recursos y campos o variables seleccionadas para visualizar, agrupando todo en una sola tabla (solo aplica para el tipo de resultado *Selección de campo*).

## Consultas

A la consulta: ¿Cuál es el recurso electrónico más usado en la Universidad CES por los estudiantes de pregrado de medicina en el último trimestre (octubre - diciembre) de 2010?, ELOGIM permite la selección de los campos referentes a la consulta, para generar luego el cruce de variables. La figura 6 muestra la forma como se realiza la consulta gráficamente.

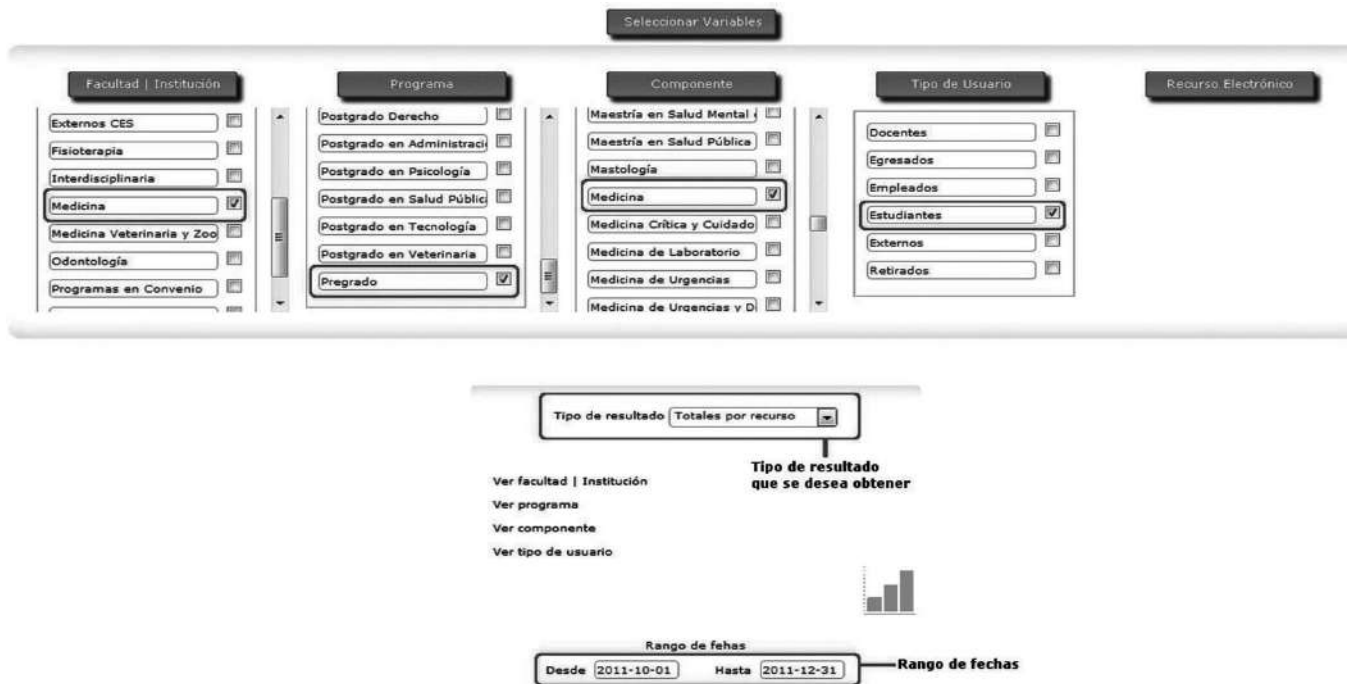


Figura 6. Composición gráfica de la consulta sobre el recurso más usado en la Universidad CES por los estudiantes del pregrado de Medicina en el último trimestre de 2010

## Resultados

ELOGIM presenta los resultados obtenidos en las consultas en listado en diferentes formatos, los cuales son predeterminados en el cruce de variables y exportables a Excel y a gráficos en forma de sectores, barras o curvas, lo que facilita la visualización y comprensión de

los datos, así como su uso en informes y reportes. La figura 7 enlista los diferentes tipos de gráficos que genera ELOGIM. Para efectos de la consulta anterior se selecciona el tipo de gráfico “Barras horizontales”.

La tabla 1 muestra la presentación de resultados para la misma consulta, en listado exportable a Excel.



Figura 7. Menú para la selección del tipo de gráfico

Tabla 1. Presentación de resultados en listado exportable a Excel

*Resultado de las Sesiones, en el rango 2010-10-01 y 2010-12-31*

<b>Facultad:</b>	Medicina
<b>Programa:</b>	Pregrado
<b>Componente:</b>	Medicina
<b>Tipo de usuario:</b>	Estudiantes
<b>Recurso electrónico:</b>	Todos

Recurso electrónico	Total de Sesiones
MDconsult	784
Pubmed	741
Ovid	391
Hinari	185
Academic Search Complete (Ebsco)	168
dynamed (Ebsco)	110
Cochrane	80
MEDLINE (Ebsco)	64
MedicLatina (Ebsco)	51
Business Searching Interface (Ebsco)	51
Scielo	50
Psychology and Behavioral Sciences Collection (Ebsco)	26
SPORTDiscus with Full Text (Ebsco)	20
Lilacs	11
MasterFILE Premier (Ebsco)	9
Fuente Academica (Ebsco)	8
Dentistry & Oral Sciences Source (Ebsco)	6
Pepsic	2
Library, Information Science & Technology Abstracts (Ebsco)	2
Hapi Online	2
Refworks	1
Regional Business News (Ebsco)	1
Noticiero oficial	1
GreenFILE (Ebsco)	1
<b>Total</b>	<b>2765</b>

En la figura 8 aparece la representación gráfica de resultados en diagrama de barras horizontales, según la consulta inicial.

A la consulta: ¿Cuál es el ranking de usuarios de los estudiantes de pregrado de medicina en el segundo trimestre (abril – junio) de 2011, en los recursos Academic

Search Complete (Ebsco), BMJ BestPractice, Business Source Complete (Ebsco) y Dentistry & Oral Sciences Source (Ebsco)?, ELOGIM permite la selección de los campos referentes a la consulta, para dar lugar luego al cruce de variables. En la figura 9 se puede observar en ELOGIM la consulta completa de todas las variables y filtros.

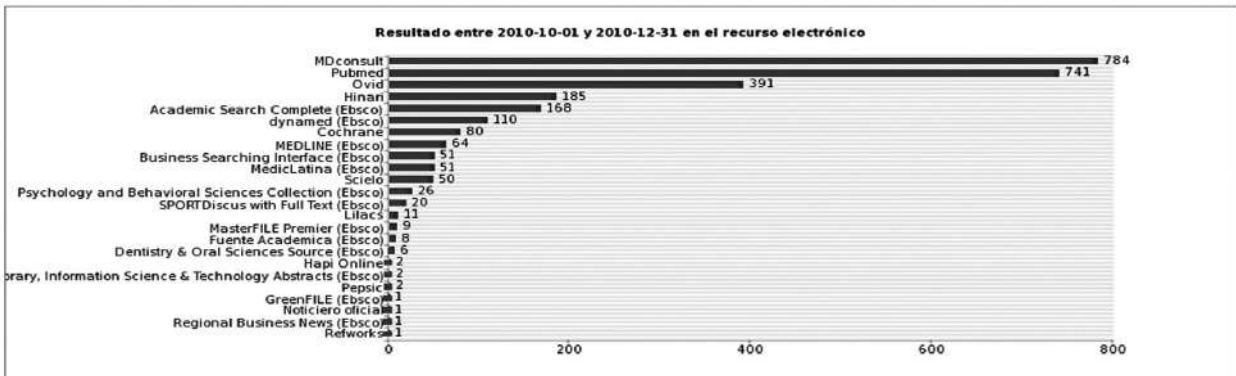


Figura 8. Representación gráfica de los resultados en diagrama de barras horizontales.



Figura 9. Representación gráfica de la consulta ¿Cuál es el ranking de usuarios de los estudiantes de pregrado de medicina en el segundo trimestre (abril – junio) de 2011 en los recursos Academic Search Complete (Ebsco), BMJ BestPractice, Business Source Complete (Ebsco) y Dentistry & Oral Sciences Source (Ebsco)



La tabla 2 muestra, en forma de listado, el resultado del ranking de usuarios, en donde se incluyen datos detallados del usuario, además de facultad, programa, componente y usuario, documento de identidad y número de sesiones del usuario en los recursos definidos en la consulta.

En la figura 10 se observa la consulta de ranking de usuarios en forma de sectores.

## Reflexión y análisis

ELOGIM, como herramienta derivada de la aplicación del método KDD, ofrece importantes beneficios, particularmente para la Biblioteca Fundadores, de la Universidad CES, beneficios que pueden ser replicados en las instituciones que acojan esta herramienta para extraer el conocimiento de sus bodegas de datos:

- **La gestión del conocimiento:** La gestión del conocimiento hace referencia al proceso sistémico de adquirir, organizar, comunicar y aplicar el conocimiento de los actores de una organización; en este contexto, ELOGIM permite determinar las necesidades de conocimiento de los usuarios de los recursos electrónicos de las bibliotecas, con miras a

suplirlas mediante el análisis de los datos sobre los recursos electrónicos que prefieren en sus búsquedas de información; Este proceso debe ser efectivo, pues el éxito de la gestión del conocimiento se basa en el adecuado proceso de la gestión de la información, es decir, que se entregue la información adecuada, en la forma correcta, al individuo u organización indicado, con un precio justo, en el momento y lugar oportuno. Todo lo anterior contribuye a la toma de decisiones acertadas.

- **La definición de patrones de comportamiento de los usuarios de los recursos electrónicos:** El análisis comportamental de los usuarios de cualquier producto o servicio es una tarea hacia la cual, en la actualidad, las organizaciones están dirigiendo su atención, pues el conocimiento de los clientes – usuarios, sus hábitos, patrones de compra o uso de servicios, permite diseñar y desarrollar programas y productos según sus necesidades. Las bibliotecas no son ajenas a este fenómeno de mercadeo y así, mediante los datos almacenados en ELOGIM, es posible identificar tendencias y patrones comportamentales que permiten predecir necesidades informacionales, generar perfiles o crear o modificar productos y servicios.

Tabla 2. Ranking de usuarios

*Resultado de las Sesiones, en el rango 2011-04-01 y 2011-06-30*

Facultad: Medicina   Programa: Pregrado   Componente: Medicina   Tipo de usuario: Estudiantes   Recurso electrónico: Academic Search Complete (Ebsco)   BMJ BestPractice   Business Searching Interface (Ebsco)   Dentistry & Oral Sciences Source (Ebsco)	<b>Criterios de la consulta</b>
--	---------------------------------

Facultad	Programa	Componente	Usuario	Identificación	Sesiones
Medicina	Pregrado	Medicina	Santiago Mejia Jaramillo	21214352	12
Medicina	Pregrado	Medicina	Ricardo Sambreño Perez	15052252	8
Medicina	Pregrado	Medicina	Gustavo Andres Palomino	15052252	7
Medicina	Pregrado	Medicina	Sebastien Giraldo Carrmona	1052986321	6
Medicina	Pregrado	Medicina	Fredy Perez Perez	1028325694	5

Resultado entre 2011-04-01 y 2011-06-30 en el recurso electrónico

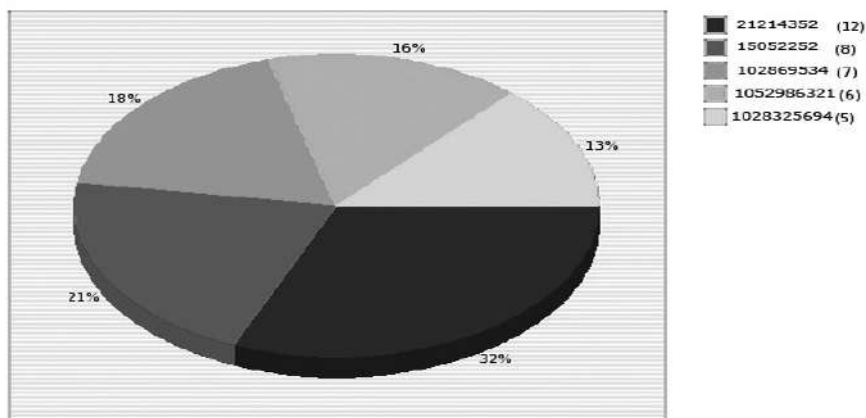


Figura 10. Ranking de usuarios en forma de sectores.

[Nora Ledis Quiroz Gil, Carlos Andrés Valencia]

- **La definición de indicadores de gestión:** La medición del desarrollo de procesos y los indicadores de gestión contribuyen a la gestión del conocimiento en las unidades de información, en la medida que sean un insumo útil de información que se fundamente en la toma de decisiones, mejore los procesos y los resultados y haga más fáciles las actividades bibliotecarias administrativas. Los datos, por sí solos, carecen de significado. Se hace necesario procesarlos y darles valor lógico, para convertirlos en indicadores que permitan una interpretación profunda de los procesos y evidencien la calidad de la gestión. ELOGIM posee un alto porcentaje de certeza en la calidad de los datos, gracias a su interconectividad con los demás sistemas universitarios, a la unificación del acceso de los usuarios a los recursos electrónicos mediante una sola contraseña, y la administración autónoma del perfil de cada usuario (ISO, 2008).
- **Los servicios al usuario:** ELOGIM facilita al usuario la administración autónoma del perfil de acceso remoto a los recursos electrónicos ofrecidos por la biblioteca. Para EZProxy, no había sido desarrollada antes una funcionalidad como esta. Antes, era necesario que un funcionario de la biblioteca interviniera directamente el código del sistema EZProxy y asignara o cambiara manualmente los datos de validación. En ELOGIM, este proceso es automático y lo realiza directamente el usuario desde la página web de la Biblioteca Fundadores.
- **La cohesión con otros servicios:** La creciente cantidad y variedad de recursos electrónicos al servicio de los usuarios de bibliotecas y unidades de información, exigía herramientas integradoras que facilitaran al usuario al acceso transparente a estos recursos. Algunas de esas herramientas integradoras, como VPN (Virtual Private Network) (Mejía Mesa, 2002), LDAP, SIP, Athens y Shibboleth (Afifi, 2004) y EZProxy, tienen como función habilitar para los usuarios el acceso a los recursos en red. Pero ninguna, con el proceso de *bibliomining* como lo hace ELOGIM, es capaz, además, de integrar los servicios a través de una sola validación del usuario de generar datos comunes susceptibles de ser analizados y transformados en conocimiento.
- **La innovación en los servicios:** ELOGIM posiciona la Biblioteca Fundadores de la Universidad CES como una unidad de información innovadora y origen de proyectos de vanguardia en pro del quehacer administrativo y de gestión en las unidades de información
- **El desarrollo de colecciones y bibliometría:** Conocer el comportamiento de los usuarios de los recursos electrónicos de información permite a las unidades de información tomar decisiones acertadas en cuanto al desarrollo de colecciones, minimizando el riesgo en las inversiones y la ejecución del presupuesto en recursos no adecuados o de bajo uso. ELOGIM posibilita el análisis exhaustivo de

estos aspectos y procura claridad sobre la necesidad, preferencia y uso de estos recursos por parte de quienes los consultan.

- **La actualización de procesos:** La agilidad en los procesos de gestión de la biblioteca y en la respuesta a solicitudes de información, específicamente para informes como los exigidos por el Ministerio de Educación Nacional en los trámites de obtención de registros calificados y de acreditación de programas académicos.

Además de los beneficios señalados, ELOGIM ha sido pensado como un desarrollo, no sólo útil sino perdurable; por eso, en futuras actualizaciones se proyecta:

- **La mejora de la generación de gráficas:** Para hacer las gráficas se usaron herramientas open sources como pChart (Pogolotti, 2011), para los gráficos de curvas, y Libchart (Libchart, 2011) para los demás. Estas herramientas ofrecen, hasta el momento, una paleta de 15 colores, lo que limita particularmente la graficación de las figuras por sectores.
- **La exportación a formato PDF:** En la programación actual del sistema no se contemplaron las librerías y componentes que permitan la exportación y presentación integrales de los resultados en archivos PDF.
- **El almacenamiento de gráficas en Excel:** Las librerías que generan los gráficos no permiten su almacenamiento al exportar las consultas a Excel, lo que obliga al usuario de ELOGIM a guardar el gráfico como imagen para ser incorporarla luego a cualquier tipo de archivo.
- **La dependencia de los datos:** En el caso de que alguna de las variables susceptibles de ser combinadas sea modificada por el área que las administra (por ejemplo, el nombre de algún programa o componente), el resultado de las consultas puede verse afectado. Esta limitante, en particular, es resuelta cuando el área respectiva envía previamente a la biblioteca los reportes de las modificaciones.

## Conclusiones

Teniendo en cuenta el incalculable volumen de información que puede ser almacenado en una base de datos, cualquiera sea su tipo o función, se hace necesario el empleo de métodos y herramientas que permitan su recuperación y análisis. El proceso del *KDD* en las unidades de información trae beneficios directos que residen en el valor de la información y en el conocimiento que se extraiga de los datos almacenados en las *data warehouse*, convirtiéndolos en indicadores de gestión que apoyen en la toma de decisiones y en la comprensión del alcance de los fenómenos que se dan en el entorno bibliotecario, entendiendo, de paso, los patrones de comportamiento informacional de los usuarios y de los recursos usados en la biblioteca, y dando opciones para gestionar la información desde una visión clara del entorno.

La técnica del *bibliomining* como proceso del *KDD* ayuda a las unidades de información en el ahorro de dinero, acerca al usuario, permite la asertividad en la adquisición de colecciones, colabora en la satisfacción de las necesidades de información de las comunidades y ofrece una visión veraz del uso de los recursos. El *KDD* debe ser la justificación, sustentada en datos, para la gestión de la información en cuanto a la defensa de presupuesto, solicitudes de financiamiento y de recursos. Todo lo anterior debe convertirse en el elemento clave para la gerencia del conocimiento; aun así, las técnicas de minería de datos, por sí solas, no son determinantes a la hora de tomar decisiones en términos bibliotecológicos, y deben ir acompañadas de técnicas clásicas que las refuercen, como los estudios de usuarios.

ELOGIM, por su parte, es un desarrollo actualmente en funcionamiento en la Biblioteca Fundadores de la Universidad CES. Se proyecta su posible aplicación en otras unidades de información e instituciones que requieran del análisis desde la bibliominería de los indicadores de uso de sus recursos bibliográficos electrónicos. Además, constituye una herramienta potencial, proveedora de datos para investigadores que quieran profundizar en el análisis de los resultados del proceso de *KDD*, abstrayéndolos del ámbito bibliotecológico y proyectándolos al entorno organizacional.

## Agradecimientos

Los autores agradecen a la Universidad CES por el apoyo en el desarrollo de ELOGIM.

## Referencias bibliográficas

- AENTA. (2011). *Sistema de gestión para evaluar y monitorear publicaciones científicas en la AENTA. Bibliometría*. Recuperado el 23 de Marzo de 2011, de <http://www.energia.inf.cu/bibliometria/Bibliometria/QueEsBibliometria.asp>
- AFIFI, Marianne. Shibboleth: Una solución de código libre para compartir recursos. En: World Library and Information Congress. [En línea] (70°, 22-27 de agosto, Buenos Aires, Argentina) Buenos Aires: IFLA General Conference and Council, Ponencia. 2004 [fecha de consulta: 03 marzo 2012] Disponible en: [http://archive.ifla.org/IV/ifla70/papers/055s\\_trans-Afifi.pdf](http://archive.ifla.org/IV/ifla70/papers/055s_trans-Afifi.pdf)
- BANERJEE, Kyle. Is data mining right for your library? *Computers in libraries*, 18 (10): 28-31, 1998.
- CONTRERAS CONTRERAS, Fortunato. Indicadores de gestión en unidades de información [en línea] Perú: Osrevi, 2005. Disponible en: [http://eprints.rclis.org/bitstream/10760/7008/1/1\\_10.pdf](http://eprints.rclis.org/bitstream/10760/7008/1/1_10.pdf)
- FAYYAD, Usama, GRINSTEIN, George & WIERSE, Andreas. Information visualization in data mining and knowledge discovery. San Francisco: Academic Press, 2002. 373p.
- ISO. ISO: 11620. Information and Documentation-library Performance Indicators. Ginebra: ISO, 2008.
- TRÉMEAUX, Jean Marc. Libchart: Simple PHP chart drawing library [en línea]. Francia: Libchart, 2011 [fecha

de consulta: 3 mayo 2012] Disponible en: <http://naku.dohcrew.com/libchart/pages/introduction/>

- MANCINI, Donna. Mining your automated system for systemwide decision making. *Library Administration & Management*, 10 (1): 11-15, 1996.
- MEJÍA MESA, Aurelio. Diccionario enciclopédico técnico actualizado. Medellín: Ditel, 2002. 748p.
- MICHAIL, A. Data mining library reuse patterns in user-selected applications. En: Automated Software Engineering (14°: 1999: Washington). Ponencia. Washington: International Conference Society. 1999. 24-33.
- OCLC. Software de autenticación y acceso del EZproxy. [en línea]. Dublín: OCLC, 2012 [fecha de consulta: 23 mayo 2012] Disponible en: <http://www.oclc.org/americalatina/es/ezproxy/default.htm>
- PETERS, Thomas. Using transaction log analysis for library management information. *Library Administration & Management*, 10 (1): 20-25, 1996.
- POGOLOTTI, Jean Damien. PChart - a PHP class to build charts [en línea]. Francia: Pchart, 2011 [fecha de consulta: 12 mayo 2012]. Disponible en: <http://pchart.sourceforge.net/>
- RIQUELME, José, RUIZ, Roberto & GILBERT, Karina. Minería de datos: Conceptos y tendencias. *Revista Iberoamericana de Inteligencia Artificial*, 10 (29): 11-18, 2006.
- VIEIRA BRAGA, Luis Paulo, ORTÍZ VALENCIA, Luis Iván & RAMÍREZ CARVAJAL, Santiago Segundo. Introducción a la minería de datos. Rio de Janeiro: E-Paper, 2009. 218p.
- WITTEN, Ian, FRANK, Eibe, & HALL, Mark. Data mining: practical machine learning tools and techniques. Burlington: Morgan Kaufmann Publisher, 2011. 650p.