

# APLICACIONES Y PERSPECTIVAS DE LOS ESTUDIOS MÉTRICOS DE LA INFORMACIÓN (EMI) EN LA GESTIÓN DE INFORMACIÓN Y EL CONOCIMIENTO EN LAS ORGANIZACIONES

*Orlando Gregorio Chaviano* \*

## **Objetivos**

- Conceptualizar los estudios métricos de la información, su origen y etapas de formación, tipología de indicadores y aplicaciones específicas en unidades de información.
- Exponer las aplicaciones de los Estudios Métricos de la información (EMI) para la gestión del conocimiento y la información en unidades y centros de información especializados, a partir de los propios conceptos de información y conocimiento como bases de la gestión de información y del conocimiento.
- Analizar las herramientas de la gestión de información y del conocimiento y su importancia práctica en las unidades y centros de información especializados y para los profesionales de la información como vías para obtener liderazgos dentro de las organizaciones.

## **Introducción**

Los Estudios Métricos de Información, como campo de investigación emergente y de frontera de las Ciencias Bibliotecológica y de la Información, han tenido un desarrollo acelerado en los últimos tiempos, debido a su naturaleza multi, inter y transdisciplinaria. Esta característica vincula a este tipo de investigación con una muy variada gama de disciplinas en forma directa o indirecta.

Este carácter, se ha visto acentuado especialmente, por la interacción entre contenidos clásicos de la Matemática y la Estadística y las Ciencias Bibliotecológica y de la Información, potenciado además con el desarrollo de las Ciencias informáticas y otros métodos, procesos y tecnologías particulares de éstas, relacionados con el tratamiento, almacenamiento y análisis de los datos acumulados en las organizaciones. (Gorbea, 2006).

---

\* Grupo de Bibliotecas Públicas. Biblioteca Nacional de Colombia.  
Calle 24 No. 5-60. Bogotá. Teléfono: 243 0804  
E-mail: ochaviano@mincultura.gov.co; ogregorioch@yahoo.com

Al principio, el análisis cuantitativo de las bases de datos bibliográficas, por ejemplo, estuvo más orientada a la identificación de regularidades presentes de forma explícita en las variables bibliográficas, y más recientemente su interés se ha centrado en descubrir conocimiento tácito, es decir implícito pero no visible en los datos, con el propósito de identificar nuevos patrones, nunca antes hallados y que aportan nuevo conocimiento, que sin la implementación de una serie de procesos, de técnicas y tecnologías novedosas, no hubiera sido posible descubrir.

En este orden de ideas, resulta claramente visible la importancia de analizar las diferentes aplicaciones y perspectivas de los estudios métricos, tanto en la gestión de información como del conocimiento, profundizando en el impacto favorable que tienen para la organización del trabajo bibliotecario, la toma de decisiones y el prestigio de las unidades de información en general y los profesionales de la información en particular. Para realizar este tipo de análisis, es preciso partir elementos tan sencillos como datos, información y conocimiento expuestos a continuación.

### **Datos**

Constituyen el soporte físico donde descansa la información. Es decir, los datos expuestos fríamente sin elaboración de significado no aportan nada a la toma de decisiones, constituyendo solamente una representación simbólica.

### **Información**

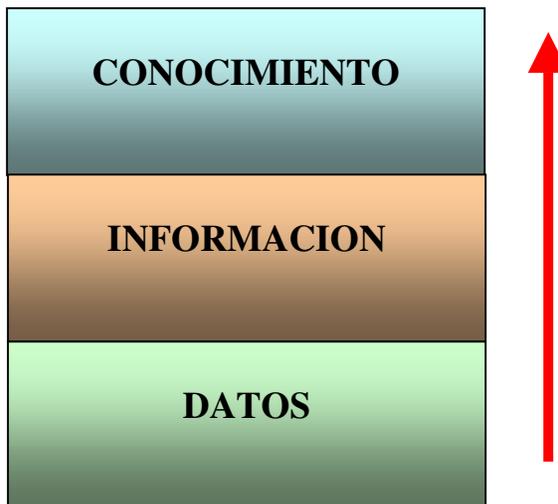
Construir un concepto de información, es tarea difícil. Un análisis importante lo realiza Davenport (2001), al decir que la información son datos dotados de pertinencia y propósito, con un bien marcado carácter contextual y de lectura del entorno.

Debemos partir de los datos para plantear que la información es un conjunto de datos que adquieren significado, es decir importancia y propósito, toda vez que la información en su ciclo de vida pasa por fases de generación, reproducción, distribución, adquisición, procesamiento, almacenamiento y asimilación.

### **Conocimiento**

Puede verse como información procesada que implica un proceso intelectual de enseñanza y aprendizaje. Para hablar de gestión del conocimiento, primeramente es necesario hablar sobre conocimiento. Davenport (2001) plantea que el conocimiento es una mezcla de experiencia estructurada, valores, información contextual e internacionalización experta que proporciona un marco para la evaluación e incorporación de nuevas experiencias e información. En las organizaciones queda presente en forma documentos, o bases de datos, rutinas, procesos, prácticas y normas institucionales.

En la sociedad de la información se manifiesta como el resultado de una información productiva que puede articularse como conocimiento artificial, como información electrónica útil, o como información en potencia desde una perspectiva pragmática, por lo que se estructura, concibe y articula en la trilogía: datos, información, conocimiento. Enfocada a modo de ejemplo:



### **Gestión de información**

Rojas (2003), afirma que hace más de veinte años se introdujo la gestión en las áreas relacionadas con la información. En la actualidad, empresas y organizaciones reconocen la información como un recurso indispensable para ampliar su competitividad, aumentar la calidad y la satisfacción de los clientes, así como para desenvolverse en el mercado global. Aproximadamente en la década de los años 80, comienza a hablarse del surgimiento de una nueva disciplina en el ámbito de la Información: *la gestión de la información*, que replantea el mercado del trabajo del profesional de la información y demanda un nuevo tipo de profesional con importantes responsabilidades en el diseño y el desarrollo de los sistemas de información en las organizaciones. Desde esa fecha han aparecido conceptos clave como: *information science, information management o information system*, que surgieron en gran medida por el acelerado proceso de producción de información en las organizaciones. (Rojas, 2006)

La gestión de información permite obtener y/o utilizar recursos básicos (económicos, físicos, humanos y materiales) para manejar información dentro y para la sociedad a la que sirve. Tiene como elemento básico, la gestión del ciclo de vida de este recurso y se desarrolla en cualquier organización. En particular, también se desarrolla en unidades especializadas que manejan este recurso en forma intensiva, llamadas unidades de información.

Ponjuán (2003), plantea que su objetivo central va a ser organizar y poner en uso los recursos de información de la organización (tanto de origen externo como interno) para permitirle operar, aprender y adaptarse a los cambios del ambiente. Los actores principales en la gestión de información son los mismos profesionales de información, en unión estrecha con sus usuarios. Los procesos principales de la gestión de información: la identificación de las necesidades de información, la adquisición de las fuentes informativas, su organización y almacenamiento, el desarrollo de productos y servicios, su distribución y uso, que son también la base de la creación del conocimiento durante la existencia productiva de la organización y por tanto, fundamento de la fase inicial de la gestión del conocimiento.

### **Gestión del conocimiento**

Diversos han sido los enfoques que en la literatura especializada han tratado las diferencias y semejanzas entre la Gestión de la Información y la del Conocimiento, los cuales han tenido como sustrato el propio debate que subyace entre los conceptos Información y Conocimiento. Solo puede residir en la mente de las personas y en el momento que abandona la mente humana se convierte en información, sin embargo, no todos los tipos de conocimientos pueden ser codificados y capturados (Saradell, 2005). El conocimiento en forma de habilidades y competencias puede ser sólo transferido de persona a persona. En esta dinámica de ideas, entran dos conceptos ampliamente estudiados:

- El *conocimiento tácito*, el que muchas personas creen que puede sólo existir en la mente humana, pero que en realidad incluye además las habilidades, competencias y experiencias de los individuos], que constituye un producto de la interacción entre las personas y su relación con su medio ambiente, además de resultar esencial para el desarrollo de una propuesta efectiva para la Gestión de Conocimiento.
- El *conocimiento explícito*, constituye el Know How y la información codificada y registrada en cualquier tipo de soporte, el cual constituye el objeto de los Sistemas de Gestión de Información.

De manera general, la gestión del conocimiento permite hoy día a las empresas, identificar la tipología de conocimiento a emplear para llevar a cabo sus actividades fundamentales, conocer donde están los nichos de conocimiento (externos e internos), garantizar la permanencia y disponibilidad de ellos y utilizarlos de manera eficiente.

### **Servicios de información para los centros de información**

Cuando analizamos de forma global la situación actual de los sistemas de información en unidades especializadas, es rápidamente visible la ausencia de servicios de valor añadido y personalizados de forma general. Tratar de establecer en las distintas unidades de información estos servicios sería transitar hacia una nueva fase de desarrollo. Esta fase se perfila como la situación deseada para las unidades

de información adscritas a centros de investigación pero que hoy día aún no se cumple.

Las unidades de información especializadas deben desarrollar productos y servicios de información para el uso interno o externo de la organización, no teniendo necesariamente, que contar con colecciones documentales propias, para la prestación de sus servicios especializados, sino que se apoya en la infraestructura informacional local, territorial, nacional, y/o internacional o en el Internet, lo que determinará en gran medida su categorización. Pueden además prestar servicios de información orientados a usuarios específicos a través de la prestación de servicios de información evaluada para la toma de decisiones, consultoría especializada, entre otros.

A la luz de esta dinámica, los estudios métricos se convierten en una herramienta indispensable para medir la información producida, tangible, pero, también, en una herramienta necesaria para producir información con valor a la luz de los evaluadores. Conocer como medir el valor de información permite conocer como dar valor a la misma. En este axioma se pone de manifiesto la importancia de las disciplinas instrumentales en la era de la información. (Cabrales, 2002).

### **Los Estudios Métricos de la Información**

Como ya explicamos, los Estudios Métricos de la Información constituyen un campo multiinter y transdisciplinario en el cual interactúan diversos métodos y modelos matemáticos y estadísticos con las disciplinas que integran el denominado Sistema de Conocimientos Bibliológico Informativo (Bibliotecología, Bibliografología, Archivología y Ciencia de la Información). A esta interacción se le atribuye el surgimiento de un conjunto de especialidades métricas (Bibliotecometría, Bibliometría, Archivometría e Informetría), las cuales han alcanzado cierta notoriedad, unas más que otras, en la literatura especializada en los últimos años.

Gorbea (2006), abordó la historia y desarrollo de los Estudios Métricos de la Información en un modelo teórico sobre esta especialidad. De acuerdo con los antecedentes históricos, el desarrollo de este tipo de estudio, así como el de las especialidades métricas vinculadas a las disciplinas de las esferas Bibliológico-Informativa y de la Ciencia, se puede dividir en tres etapas, a saber:

- **Etapla pre-disciplinar (1743-1897):** Desarrollo de estudios, de forma aislada, sobre las relaciones de citas, principalmente en la esfera jurídica; análisis cuantitativo sobre producción literaria, y, compilaciones de datos estadísticos sobre el comportamiento de las bibliotecas y la composición de sus colecciones.

- **Etapla disciplinar (1917-1979):** Denominación y definición de especialidades métricas clásicas que agrupan este tipo de estudio asociado a las disciplinas que le dieron origen.
- **Etapla de desarrollo disciplinar (1979- hasta la fecha):** Consolidación de las especialidades métricas en un cuerpo de conocimiento, surgimiento de nuevas especialidades y denominaciones, desarrollo matemático, terminológico, curricular e investigativo de este campo del saber.

Durante estas tres etapas de desarrollo este tipo de estudio ha puesto a prueba la factibilidad del uso de teorías, métodos, modelos e indicadores cuantitativos en la identificación de las regularidades de la producción y comunicación científica en muy diversas áreas temáticas y regiones, así como las reveladas en los sistemas bibliotecarios y de la información. Estos resultados han aportado información muy valiosa para la toma de decisiones en los Sistemas de Gestión de Información, constituyendo en estos últimos una herramienta novedosa de evaluación en su organización y planificación.

Paralelamente al desarrollo y consolidación de los Estudios Métricos de la Información, en los Sistemas de Gestión de Información también se han experimentado importantes cambios, estos sistemas han transitado de la gestión de conocimiento explícito, registrado en las fuentes de información, a la del conocimiento tácito; este último como el producto de la interacción entre las personas, su inteligencia, sus habilidades y sus experiencias en su relación con el medio ambiente y demás miembros de las organizaciones a las que pertenecen.

#### **Formación de las disciplinas métricas**

La utilización de las matemáticas en las ciencias sociales tiene su antecedente fundamental en el positivismo de *Augusto Comte*, filósofo y matemático francés. Esta doctrina filosófica tuvo una influencia especial en el siglo XIX, al excluir los conocimientos que no fueran los provenientes directamente de la experiencia del individuo. (Vallet, 2006).

A pesar de la existencia de nuevas disciplinas instrumentales, surgidas con el propio desarrollo científico-tecnológico de los últimos años, como la webmetría y la patentometría, con un sinnúmero de indicadores, la mayoría de los estudios toman como punto de referencia, la bibliometría, la cienciometría y la Informetría como disciplinas instrumentales de la bibliotecología, la cienciología y la ciencia de la información, respectivamente.

En la medida en que se desarrollaron las diferentes disciplinas -bibliotecología, documentación, ciencia de la información, etc.-, se desarrollaron también los estudios métricos. Desde principios del siglo XX, las matemáticas se emplearon con un enfoque métrico para la toma de decisiones y para la obtención de resultados cuantitativos, aunque no es hasta los años 60 que el término Informetría se utiliza

por primera vez. Puede afirmarse y es criterio del autor, que las técnicas métricas desde su surgimiento, se desarrollaron a la par de las disciplinas científicas a las que pertenecen. El propio desarrollo de la bibliotecología y la documentación y, más cerca en el tiempo, de la ciencia de la información, fueron el entorno en el que crecieron la bibliometría y la informetría.

A finales del siglo XX y en el presente, se han fortalecido dos nuevos términos métricos que se utilizan con relativa frecuencia: la webmetría, aparejada al desarrollo del WWW para medir la ciencia desde la perspectiva de indicadores propios de Internet y la patentometría, una herramienta útil para el análisis de oportunidades tecnológicas y del comportamiento de las patentes. Ambas son indispensables para el desarrollo científico y la toma de decisiones en las organizaciones modernas.

Antes de realizar un análisis de las diversas disciplinas métricas, se expondrán brevemente los aspectos teórico-conceptuales propios de cada una de las ciencias a las que pertenecen.

### **Bibliotecología**

Estudia los objetivos, principios, contenidos y uso social de los libros, aunque se ocupa además de la colección, almacenamiento y distribución de los registros impresos que forman parte de las bibliotecas, así como de investigar las leyes del desarrollo bibliotecario. Su objeto de estudio está marcado por la circulación bibliográfica y su utilización, en tanto medio de educación social; mientras que su tema de estudio analiza las regularidades del comportamiento y desarrollo de su objeto, es decir, la circulación y el uso de las fuentes presentes en la biblioteca.

### **Cienciología**

Existe como ciencia desde la década del 60 del pasado siglo y se encuentra muy ligada al desarrollo de la ciencia de la información. Se encarga de investigar el funcionamiento de los sistemas científicos para crear métodos que fortalezcan el potencial de la ciencia y el desarrollo científico. Como objeto de estudio analiza las características y regularidades del desarrollo de la actividad científica.

### **Ciencia de la información**

La Ciencia de la Información surgió, entre otras razones, debido a las limitaciones de la Documentación, a partir de la necesidad de una nueva definición para los nuevos enfoques que esta brindaba. El centro de problema no era más el documento sino la información en cualquier lugar y soporte.

Estudia la estructura de la información, la interacción ciencia de la información-sociedad, así como las propiedades de las fuentes y las regularidades de los procesos de su transmisión y procesamiento.

Su objeto de estudio comprende los datos, el lugar de las fuentes de información (documentales y no documentales) en la sociedad, la informatización de las instituciones, así como el procesamiento, organización, conservación, búsqueda y disseminación de la información.

### **Principales disciplinas métricas**

#### **Bibliometría**

La Bibliometría, como disciplina instrumental de la bibliotecología, consiste en la aplicación de las matemáticas y los métodos estadísticos para analizar el curso de una determinada disciplina científica, así como a su comportamiento. Aporta información cuantitativa para el análisis integral en el campo de la bibliotecología. Comprende la aplicación de análisis estadísticos que estudian las características del uso y creación de los documentos, el estudio cuantitativo de la producción de documentos, la aplicación de métodos matemáticos al análisis del uso de los libros y otros soportes, dentro y entre los sistemas de bibliotecas, así como el estudio cuantitativo de las unidades físicas publicadas y de las unidades bibliográficas. (Morales, 1995).

Ofrece, además, elementos cuantitativos para la organización y dirección de las bibliotecas, así como para el estudio de su efecto en la sociedad. Perfecciona la toma de decisiones y analiza las fuentes documentales. Y por ello, constituye una herramienta indispensable para la gestión bibliotecaria.

#### **Informetría**

Como disciplina instrumental de la Ciencia de la Información, estudia los aspectos cuantitativos de la información.

Permite, sobre la base de elementos cualitativos y cuantitativos, el análisis de los fenómenos y procesos relacionados con la información. Entre sus aplicaciones más importantes en el presente, están: analizar los flujos de información, la obsolescencia de la información y medir el nivel de informatización de la sociedad.

Comprende asuntos como el desarrollo de modelos teóricos y medidas de información para determinar las regularidades en los datos asociados con la producción y el uso de la información registrada; abarca la medición de aspectos de la información, el almacenamiento y su recuperación, incluye la teoría matemática y la modelación.

#### **Cienciometría**

En un principio, el término se refería sólo a la aplicación de métodos cuantitativos a la historia de la ciencia y el progreso tecnológico. Utiliza métodos matemáticos para el estudio de la ciencia y a la actividad científica en general, además de medir el nivel de desarrollo y el aporte de la ciencia a las diferentes esferas de la sociedad.

#### **Patentometría**

Como consecuencia del desarrollo de la ciencia y la técnica, se amplió el alcance de las disciplinas métricas a otros campos del quehacer científico. Un ejemplo es la aplicación de las técnicas métricas a la información de patentes conocida como patentometría. La importancia estratégica de las patentes como fuente de información, produjo la aparición de indicadores métricos para analizar este tipo de documento fundamentalmente para la búsqueda de oportunidades tecnológicas, así como para la evaluación de programas de I+D.

Constituye una herramienta de vital importancia para el análisis de oportunidades tecnológicas en las organizaciones que parte de la utilización de indicadores bibliométricos de patentes para analizar las actividades de gestión tecnológica en las empresas y organizaciones y que lleva a que las empresas logren una mayor orientación de su capacidad tecnológica y puedan agregarle valor a la información como producto.

### **Webmetría**

El crecimiento de Internet, de sus productos y servicios, del número de páginas Web, generó la necesidad de cuantificar y analizar, a partir de ciertos indicadores y criterios, todos estos componentes; así, apareció la *webmetría*, una disciplina métrica que se ocupa del análisis de los componentes del Web, pero con la misma filosofía de los indicadores bibliométricos.

A pesar de la existencia de las diferentes disciplinas métricas, surgidas como disciplinas instrumentales de diferentes ciencias, el término bibliometría se ha generalizado para la denominación de los estudios de esta índole.

### **Cronología mínima de los Estudios Métricos de la Información.**

#### **1917**

- *Cole y Eales* realizaron un análisis estadístico de las publicaciones sobre anatomía comparativa.

#### **1923**

- *E. Hulme* realizó un análisis estadístico de la historia de la ciencia.  
- Se acuñó el término bibliografía estadística.

#### **1926**

- Se publicó la Ley de *Lotka*, que describe una relación cuantitativa entre los autores y los artículos producidos en un campo durante un periodo de tiempo. Dicha ley revela la existencia de una distribución desigual, en tanto que la producción de la mayoría de los artículos se concentra en un número pequeño de autores de gran productividad.

#### **1927**

- *Gross* analizó las referencias hechas en un grupo de artículos publicados en *The*

*Journal of the American Chemistry Society*, donde, por primera vez, se utilizó el análisis de citas.

**1933**

- Zipf formuló, una nueva ley relacionada con la frecuencia con que se emplea una palabra dentro de un texto.

**1948**

- Se publicó la Ley de *Bradford*, que plantea que si un número determinado de publicaciones periódicas se ordenan en forma decreciente según su productividad de artículos y se dividen en tres zonas de análisis, cada una con la misma cantidad de contribuciones, entonces existirá una zona núcleo con un reducido número de fuentes donde se concentrará la mayor cantidad de artículos y dos zonas restantes con una mayor dispersión.

- *Ranganathan* acuñó el término *Librametry*, que puede entenderse como bibliotecometría.

**1954**

- Eugene Garfield propuso la creación de los índices de citas. Ellos permiten evaluar el desempeño de la investigación individual, institucional y por zonas geográficas.

**1963**

- *J. S Price* publicó "*Little Science, Big Science*", donde introdujo el estudio de los aspectos cuantitativos de la actividad científica.

**1969**

- *Alan Pritchard* empleó, por vez primera, el término bibliometría.

**1978**

- Se fundó la revista *Scientometrics*, una publicación insigne en el campo de los estudios métricos, que trata los aspectos cuantitativos de la ciencia de la ciencia, la comunicación y las políticas científicas a escala internacional.

**1979**

- *Otto Nacke* propuso, por vez primera, el término Informetría.

**1987**

-Se generaliza el término y se realiza la Primera Conferencia Internacional sobre Bibliometría y Aspectos Teóricos de la Recuperación de la Información, como se le denominó en aquel momento. Se le conoce actualmente como *Conferencia Internacional de la Sociedad Internacional de Informetría y Cienciometría*.

**1996**

- Comienzan a aparecer los estudios bibliométricos aplicados a Internet (cibermetría).
- Arnzen realizó un estudio en el que habla de las *cibercitas*.
- Clausen cuantificó los hábitos de los usuarios de Internet y sus actitudes hacia la red como recurso de información en Dinamarca.
- Gorbea Portal introduce el termino Estudios Métricos de la Información (EMI).

### **1997**

- Se creó el Centro de Estudios Informétricos de Dinamarca.

### **1998**

- Se fundó el Laboratorio de Estudios Métricos de la Información (LEMI) de la Universidad Carlos III de Madrid. Su objetivo va a ser el de participar, de forma activa en la definición y desarrollo de los campos relacionados con los estudios cuantitativos de la información.

### **Los Estudios Métricos de la Información y su relación con la Gestión de Conocimiento**

En las últimas décadas, el creciente cúmulo de datos almacenados en las empresas, la industria y otras organizaciones, unido a la necesidad de resolver problemas asociados a la toma de decisiones en materia de mercadotecnia, financiera y de inteligencia competitiva en las mismas, ha propiciado la aparición de una serie de procesos y técnicas orientadas a identificar patrones de comportamiento ocultos en los reservorios de datos.

Estas técnicas y procesos, no obstante a que han surgido en cuerpos de conocimientos tales como el Tratamiento Automatizado de Datos, la Gestión de Bases de Datos y otros campos de la Informática y la Computación, constituyen campos muy relacionados pero no incluyentes en las Ciencias Bibliotecológica y de la Información. Sin embargo, este tipo de técnicas y procesos convergen con estas disciplinas a través de los Estudios Métricos de la Información, lo cual le ha permitido a este tipo de estudio extender sus líneas de investigación e identificar otras regularidades y patrones de comportamiento antes no estudiadas con las variables de descripción de forma y representación de contenido de los documento, empleando para ello los métodos y herramientas de especialidades métricas ya conocidas como la Bibliometría y la Informetría.

Si se toma a la gestión del conocimiento como la forma de gestionar los activos intangibles con mayor valor para las organizaciones, entonces el flujo de información es la base. El control de flujo de información se convierte en actividad central de todo el proceso organizacional. Aquí entran nuevamente los Estudios Métricos de la Información y sus múltiples técnicas de análisis, de hecho está involucrada en todo proceso de agregación de valor a los servicios de información. Se establece como base, a la hora de determinar el tipo de información, la

dosificación, y la comparación de información para brindar a los usuarios de este tipo de servicio.

La identificación de regularidades presentes en el Flujo de Información Documental (FID), compilado en Bases de Datos Bibliográficas, tales como la obsolescencia, la concentración-dispersión, la productividad de autores científicos y otras presentes en los sistemas bibliotecarios y de la información, le han permitido a los Estudios Métricos de la Información aportar información valiosa para la toma de decisiones a los Sistemas de Gestión de Información.

La renovación de este tipo de estudio mediante su convergencia con los procesos de Descubrimiento de Conocimiento, antes descrito, abre ante sí una nueva etapa de incalculables resultados orientados en esta ocasión a la toma de decisiones en los Sistemas de Gestión de la Información y otros sistemas, pero fundamentalmente, a los de Gestión de Conocimiento, como una fase superior de gestión en estos tipos de sistemas.

Lo anterior se debe a que los nuevos sistemas de Gestión de Conocimiento, aunque también utilizan información, en la toma de decisiones y sus actividades gerenciales, centran su funcionamiento en la inteligencia, la experiencia, las habilidades o las capacidades cognitivas de los miembros de la organización y de su entorno, así como en el nuevo conocimiento que se encuentra oculto en los datos o que no aparece de forma explícita en las bases de datos que se disponen.

En estas nuevas relaciones, las fuentes que tradicionalmente se venían utilizando en el estudio Métrico de la Información, tales como bases de datos bibliográficas o factográficas se amplían y diversifican en otras de tipo textuales, Bibliotecas Digitales y en un estado superior de integración de los datos en los denominados Data Warehouse y otros recursos y contenidos en la WWW. Ello con el propósito de satisfacer las exigencias del proceso de Descubrimiento de Conocimiento y de las técnicas de la Minería de Datos, de texto, de la Web, de la Bibliominería y del OLAP. (Fernández, 2006).

En este tipo de estudio, de la selección de la fuente o del tipo de proceso utilizado, así como del fin último que se pretenda con los resultados obtenidos, dependerá su relación con las especialidades métricas en las que se alcance su explicación, además de la orientación que se le den a los resultados, es decir al tipo de Sistema de Gestión que los demande.

Las relaciones entre los procesos, especialidades métricas y Sistemas de Gestión hacia el cual se orientan los resultados pueden ser representadas en el modelo que se presenta a continuación.

### **Algunos procesos y técnicas.**

### **1. El KDD (Knowledge Discovery Database – Descubrimiento de Conocimiento en Bases de Datos) o Minería de datos.**

Este proceso o técnica, posee amplia cobertura bibliográfica. Herrera (2006) plantea que es el proceso, que se encarga de la preparación de los datos y la interpretación de los resultados obtenidos, los cuales permiten que dichos datos adquieran un sentido y aporten un nuevo conocimiento. Se diferencia de la Minería de datos pues esta es el proceso de descubrir patrones de información interesante y potencialmente útil, inmerso en una gran base de datos. Se encarga específicamente, de la preparación de los datos y la interpretación de los resultados obtenidos, los cuales permiten que dichos datos adquieran un sentido y aporten un nuevo conocimiento. Mientras que la minería de datos es el proceso de descubrir patrones de información interesante y potencialmente útil, inmerso en una gran base de datos. Está asociada a la Minería de datos.

Esta técnica permite:

- Limpieza de los datos (Para remover los datos inconsistentes que adicionan ruido)
- Integración de los datos (Combinación de múltiples datos fuentes, en datos de mayor nivel de integración)
- Transformación de los datos (Consolidar los datos en formas más apropiadas para ser utilizados en la Minería de Datos)
- Minería de Datos (Proceso esencial donde son aplicados métodos inteligentes para extraer patrones de los datos)
- Evaluación de patrones (Identificar los patrones que verdaderamente representan el conocimiento basado en algunas mediciones interesantes)
- Representación del Conocimiento (Presentar el conocimiento minado al usuario, mediante el uso de técnicas de visualización y representación de conocimiento).

En este conjunto de pasos interviene la creación de un reservorio de datos denominado en la literatura especializada como Data Warehouse, el cual es un repositorio de datos separado de las bases de datos operacionales de un sistema y que contiene una versión limpia, normalizada y anónima de los datos formateados específicamente para su análisis.

Constituye una actividad que guarda estrecha relación con la Informática, la Estadística y la Documentación y con el KDD. Es utilizada en numerosas disciplinas para el análisis de grandes cantidades de datos. En bibliotecas, su práctica está poco extendida a pesar las múltiples aplicaciones que tiene.

La Bibliominería es un proceso emergente y resultante de la combinación de la Minería de Datos, la Bibliometría, la Estadística y de las herramientas usadas para extraer patrones de comportamiento basados en los recursos digitales de un sistema de biblioteca. Aunque en la literatura especializada muchos trabajos que hacen

referencia en este contexto al término biblioteca no se refieren a una biblioteca tradicional, sino que están haciendo alusión a las bibliotecas implementadas en determinado software, como los sistemas de recursos que acumulan grandes volúmenes de información en las bibliotecas especializadas (Nicholson, 2003)

De la definición anterior aportada por los propios creadores del término se concibe a la Bibliominería como un proceso en el cual se deben considerar los pasos o etapas siguientes:

- Determinación de áreas focales o problemas específicos que requieren de la toma de decisiones en la biblioteca.
- Identificar las fuentes de datos apropiadas.
- Creación del Data Warehouse o repositorios de datos, confeccionado a partir de los datos seleccionados de las bases de datos del sistema y que, como ya se indicó anteriormente, contiene una versión de los datos limpios reformateados para su análisis.
- Selección de las herramientas apropiadas para el análisis, en este sentido el uso de la estadística multivariada y la inteligencia artificial han resultado ser una buena herramienta en la identificación del comportamiento de patrones no visibles por otros métodos en los datos.
- Descubrimiento de patrones y creación de reportes a través de la Minería de Datos, de acuerdo con lo antes expuesto sobre este concepto anteriormente.
- Análisis e implementación de los resultados.

Una de sus áreas es la bibliominería que se ocupa del análisis y cruce de datos mediante técnicas estadísticas para descubrir y establecer patrones y tendencias en los datos como ayuda a la toma de decisiones (obtener el dato por el dato, la mera presentación de resultados es una tarea vacua e inútil). Si bien en bibliominería se trata de datos sobre la actuación previa a la utilización de la información, mientras que la bibliometría trabaja con los datos relativos a la información que ha sido finalmente usada.

En el ámbito bibliotecario es ampliamente utilizada esta técnica. Dos de las aplicaciones relacionadas con los EMI son las siguientes:

- **Información al usuario:** Ayuda a entender el comportamiento de los usuarios para conseguir patrones que permitan a la biblioteca tener una idea de las necesidades de información y por lo tanto modificar y mejorar los servicios en base a los requerimientos particulares de los usuarios.
- **Toma de decisiones:** Permite conocer las tendencias de trabajo con las bases de datos y determinar tendencias de uso a corto y largo plazo. Ejemplo: a partir de análisis de estrategias de búsqueda.

**Aplicaciones más usuales en unidades de información:**

- Predicción de necesidades de los usuarios: ver la evolución de las temáticas consultadas puede ayudar a predecir cuáles serán consultadas posteriormente.
- Identificación de materiales no consultados: por un mal proceso de selección o por una catalogación o clasificación incorrecta.
- Justificación del mantenimiento o supresión de acuerdos de préstamo interbibliotecario: gasto del préstamo frente a gasto de adquisición.

## 2. Minería de Texto o Webmining

La Minería de Textos es una tecnología emergente vinculada con el Estudio Métrico de la Información, cuyo objetivo es la búsqueda de conocimiento en grandes colecciones de documentos no estructurados, es decir: consiste en descubrir, a partir de cantidades de textos grandes, el conocimiento que no está literalmente escrito en cualquiera de los documentos. Puede abarcar una amplia gama de dominios, desde aquellos de la recuperación y extracción de información, visualización de información, minería de datos aplicada a textos, etc. (Botta, 2007).

Esta técnica opera sobre bases de datos textuales no estructuradas con el objetivo de detectar patrones no triviales e incluso información sobre el conocimiento almacenado en las mismas. Permitiendo así el análisis léxico de los textos y especialmente la construcción automática de estructuras de clasificación y categorización que se codifican en forma de *tesauros*. De modo que estos sistemas pueden ayudar en la categorización de la información existente en una organización, en el filtrado y enrutado de información por ejemplo de e-mail, en la detección de información similar o relacionada con otra existente o para eliminar información duplicada. (Witten, 2005)

### Descripción de procesos:

- Adquisición de texto a analizar.
- Normalización.
- Usualmente en formato basado en XML
- Extracción de metadatos identificativos: autor, título, fecha, fuente,...
- Filtrado: Identificación de textos relevantes mediante un análisis de presencia de
- palabras predeterminadas.
- Análisis: Establecimiento de relaciones entre textos con base en los términos y
- categorías.
- Visualización: Uso de gráficos y diagramas

**Los Estudios Métricos de la Información y la gestión de información.  
Principales aplicaciones.**

Son muchos hoy día los campos de aplicación de los estudios métricos en la gestión de información. Las diferentes aplicaciones se han convertido en herramientas indispensables para la gestión bibliotecaria en general y para los profesionales de la información de manera particular.

Las herramientas métricas, juegan un importante papel dentro de las organizaciones, a la hora de gestionar la información. Toda vez que proporcionan elementos clave para la toma de decisiones y la formulación de políticas. La utilización de las métricas proporciona además, tener a mano la información puntual en el momento necesario. Su importancia y aplicabilidad práctica en la investigación científica actual, pero sobre todo en el propio desarrollo de la Ciencia de la Información y la Bibliotecología, y específicamente en la gestión de información, pueden resumirse en los siguientes aspectos generales:

- Estudio de la Producción y Comunicación Científica.
- Actividades bibliotecarias y de información. Industria editorial y de contenidos.
- Actividades docentes y de I+D.
- Actividades de vigilancia tecnológica.

### **Indicadores**

Los indicadores bibliométricos constituyen un elemento clave en los estudios métricos, toda vez que proporcionan información cuantitativa y objetiva sobre los resultados del proceso investigador, su volumen, evolución, visibilidad y estructura. Son muy útiles aplicados a comunidades con gran volumen de publicaciones, y pierden su significación cuando se trata de producciones muy pequeñas, por lo que no deben emplearse en las valoraciones de investigadores individuales.

La aplicación de los estudios métricos de la información, tanto en la gestión de la información, como en la gestión del conocimiento, parte en gran medida de indicadores específicos que permiten su utilización en unidades de información. Cada indicador está directamente relacionado con el tipo de estudio y sobre todo con la tipología de análisis (Bibliometría, informetría, ciencimetría, webmetría y patentometría). En cualquier rama de la ciencia, un indicador va a ser, una medida de evaluación, valoración difusión y consumo de la información científica. (Spinak, 1998).

Se expondrán algunas de las aplicaciones específicas de los estudios métricos de la información en la gestión de información en las organizaciones modernas a partir de la visualización de algunos indicadores de importancia estratégica:

- Indicadores socioeconómicos
- Indicadores de producción científica (indicadores de consumo y producción)
- Indicadores de colaboración-cooperación
- Indicadores de patentes
- Indicadores cibermetricos

### **Indicadores socioeconómicos**

- Indicadores de input (inversión): Miden los recursos dedicados a la ciencia y la tecnología y su evolución temporal.  
*Ej. Financiación, gastos de I+D y personal dedicado a I+D.*
- Indicadores de output (resultados e impacto): Miden cantidad de investigaciones, publicaciones, patentes....

### **Indicadores de producción científica** Factor de impacto, índice de inmediatez....

- Productividad por investigador, país, institución, año.
- Distribución por campo científico (materias)
- Número y distribución de publicaciones.
- Número y distribución de citas (consumos de información).
- Países que desarrollan o trabajan en determinado frente de investigación y el “*colegio invisible*” que lo integra. Idiomas que emplean en la publicación de sus resultados de investigación y el uso de información. Autores más relevantes y su productividad por ramas o disciplina. Instituciones en las que se genera la información, con cuáles otras se relacionan y dónde publican sus resultados. Cómo se organizan y colaboran los colectivos de autores para publicar sus resultados. (grupos de investigación, afiliaciones, niveles de colaboración).
- Tipología documental y estructura temática de sus contenidos. Crecimiento, concentración-dispersión, obsolescencia y uso de la información por disciplinas o temáticas.
- Impacto, visibilidad de autores, colectivos de científicos, instituciones y fuentes de información.
- Frecuencia de palabras y co-palabras en los contenidos de documentos (mapas semánticos). Tendencias de Investigación y Desarrollo, prospectiva tecnológica.

### **Indicadores de colaboración**

- Distribución de autores en un flujo de información.
- Distribución del número de publicaciones por autor.
- Índice de co - autoría.
- Colaboración institucional, por país (visibilidad).

### **Indicadores de patentes**

- Número y distribución de patentes solicitadas y concedidas. (*dinamismo de un campo, productividad*)
- Productividad de los innovadores, países, instituciones (*prestigio, asignación de recursos*).
- Productividad por el número de citas del documento de patente. (*medir la fortaleza y debilidad de la tecnología, medir impacto en la competencia, identificar liderazgo*) Firmas conjuntas en las patentes. (*índices y características de colaboración, redes, alianzas estratégicas*).

- Redes de citas (*identificar frentes de investigación, identificar estrategias tecnológicas, identificar líderes*). Co-ocurrencia de palabras. (*relaciones entre temas de investigación, identificación y transformación de temas*).

#### **Indicadores cibernéticos**

- Indicadores de tipos institucionales -cantidad de páginas en determinados sectores.
- Indicadores regionales e idiomáticos.
- Indicadores de tipología de sitios -sitios académicos, comerciales, de sectores públicos o privado.
- Indicadores de tamaño (tamaño documental -número total de páginas comprendidas en un dominio) o tamaño informático (tamaño en bytes de una sede Web).
- Indicadores de profundidad. Número máximo de niveles de una sede.
- Indicadores de visibilidad. Número total de enlaces externos diferentes recibidos por una sede, existen variantes para calcular la visibilidad nacional con límite a los enlaces recibidos en el mismo país, etcétera.
- Indicadores de validez hipertextual. Porcentaje de enlaces válidos respecto al total.
- Indicadores de cooperación (colegios invisibles).
- Indicador de diversidad. Distribución de las características de los enlaces recibidos por una página.
- Medidas de popularidad. Número y distribución de las visitas recibidas en un plazo determinado.
- Indicadores de impacto. Resultado de dividir el número total de enlaces externos diferentes recibidos por una sede por su tamaño expresado en número de páginas.

#### **Con los indicadores métricos, se pueden conocer entre otros aspectos:**

- El crecimiento de cualquier campo del saber.
- El envejecimiento de los campos científicos.
- La evolución cronológica de la producción científica.
- La productividad de los autores o instituciones.
- La colaboración entre científicos e instituciones.
- El impacto o visibilidad de las publicaciones en la comunidad científica.
- La dispersión de las publicaciones científicas entre las diversas fuentes.

#### **Estructuración y diseño de los estudios métricos de la información.**

- Identificar y conocer el objeto de estudio, familiarizarse con las diferentes estrategias de búsqueda, fuentes de información y características de la temática a estudiar.
- Búsquedas bibliográficas y proceso de salva de los ficheros resultantes en el formato seleccionado.

- Conversión del fichero resultante al modelo de software a emplear para el procesamiento.
- Creación de la base de datos para el procesamiento de información.
- Normalización de la información.
- Tratamiento estadístico.
- Interpretación, conclusiones y recomendaciones.
- Creación de bibliografías o índices alternativos. (producto final)

**Principales aportes de los Estudios métricos a las unidades de información especializadas.**

A partir de la utilización de técnicas y procedimientos, tanto de la gestión de información como de la gestión del conocimiento, las unidades de información pueden contribuir al desarrollo de las organizaciones. Algunos de los ejemplos se enuncian a continuación:

- Conocer el estado actual de las publicaciones de su institución, hacer comparaciones dentro de ella y analizar el estado respecto a otras instituciones.
- Dirigir la política editorial, así como establecer su relación con la competencia en frontera y fuera de ella.
- Conocer temáticas, autores, instituciones o países líderes que permitan diseñar, evaluar y dirigir proyectos de investigación.
- Analizar las influencias de tendencias y establecer pautas a través de estudios de diferentes centros.
- Lograr el rediseño de las tendencias a seguir y determinar el grado de desviación con respecto a los centros de excelencia.
- Conocer la dinámica tecnológica por campo de investigación, instituciones o países.

**Servicios deseados en los centros y unidades de información:**

- Conocer el estado actual de las publicaciones de su institución, hacer comparaciones dentro de ella y analizar el estado respecto a otras instituciones.
- Dirigir la política editorial, así como establecer su relación con la competencia en frontera y fuera de ella.
- Conocer temáticas, autores, instituciones o países líderes que permitan diseñar, evaluar y dirigir proyectos de investigación.
- Analizar las influencias de tendencias y establecer pautas a través de estudios de diferentes centros. (vigilancia).
- Lograr el rediseño de las tendencias a seguir y determinar el grado de desviación con respecto a los centros de excelencia.
- Conocer la dinámica tecnológica por campo de investigación, instituciones o países.

### **Consideraciones Finales**

Los Estudios Métricos de la Información (EMI) constituyen un campo de investigación que ofrece diversidad de variantes de aplicación hoy día, tanto para la gestión de la información como para la gestión del conocimiento y están involucrados en todo proceso de agregación de valor a los servicios de información.

Es necesario que las unidades de información especializadas en general y los profesionales de la información específicamente, comiencen a apoderarse de las herramientas necesarias para hacer uso de ellos y ponerlos en función de las necesidades de información de sus usuarios, creando servicios de información con valor agregado.

Los profesionales de la información deben y pueden posicionarse y obtener liderazgo dentro de sus organizaciones toda vez que cuentan con los conocimientos y herramientas necesarias para ello. Deben ser los encargados de conocer las formas de obtener, organizar y procesar la información necesaria para la toma de decisiones. En los estudios métricos de la información está la clave de ello.

### **Literatura consultada**

Botta, E; Cabrera, JE. 2007. La Minería de textos. Una herramienta útil para mejorar la gestión del bibliotecario en el entorno digital. VII Jornada Nacional Bibliotecaria Instituto de Información Científica y Tecnológica (IDICT).

Consultado Julio 13 de 2007. Disponible en [www.idict.cu/.../Botta%20Ferret,%20Eleazar%20\\_%20Mineria\\_%20Jornada.pdf](http://www.idict.cu/.../Botta%20Ferret,%20Eleazar%20_%20Mineria_%20Jornada.pdf)

Cabrales, G. 2002. Análisis de los actuales servicios de información para los centros de investigación. *Ciencia da Informacao*. 31 (2):52-59.

Davenport TH; Prusack L. 2001. *Conocimiento en acción. Como las organizaciones manejan lo que saben*. Buenos Aires. Prentice Hall.

Fernández Marcial, V. 2006. *Gestión del Conocimiento vs Gestión de la Información*”.--*Investigación Bibliotecológica*, 20(41): 44-62.

Gorbea-Portal. S. 2006. *Aplicaciones de los Estudios Métricos de la Información*. Notas de clases.

Herrera, R. 2006. *Bibliomining: minería de datos y descubrimiento de conocimiento en bases de datos aplicados al ámbito bibliotecario*. Artículo en prensa.

Hudomalj, E.; G. Vidmar. 2003. OLAP and bibliographic databases. *Scientometrics* 58(3):609-622.

Martínez Rodríguez, A. 2006. Indicadores cibernéticos: ¿Nuevas propuestas para medir la información en el entorno digital?. *Acimed*. 14(4). Consultado Julio 5 de 2007. Disponible en [http://bvs.sld.cu/revistas/aci/vol14\\_4\\_06/aci03406.htm](http://bvs.sld.cu/revistas/aci/vol14_4_06/aci03406.htm)

Morales Morejón M. 1995. La bibliotecología, la ciencia de la información y sus disciplinas instrumentales: su alcance conceptual. *Ciencia de la Información* 26(2):70-88.

Nicholson, S. 2006. The basis for bibliomining: frameworks for bringing together usage-based data mining and bibliometrics through data warehousing in digital library services. *Information processing & management*, 42 (3) 785-804. Consultado Agosto 13 de 2007. Disponible en <http://dx.doi.org/10.1016/j.ipm.2005.05.008>

Ponjuán Dante, G. 2004. Gestión de información. Dimensiones e implementación para el éxito organizacional. Nuevo Paradigma, Rosario, ARGENTINA. 208 p.

Rojas, Y. 2006. De la gestión de información a la gestión del conocimiento (en línea). *Acimed*. (14). Consultado: 10 de agosto de 2007. Disponible en [http://www.bvs.sld.cu/revistas/aci/vol14\\_1\\_06/aci02106.htm](http://www.bvs.sld.cu/revistas/aci/vol14_1_06/aci02106.htm)

Saradell López E, Pérez JA. 2005. La gestión del conocimiento en la nueva economía. Consultado Agosto 12 de 2007. Disponible en: <http://www.ouc.edu.cu/dt/20133.html>

Spinak. E. 1998. Indicadores cienciométricos. *Ciencia da Informacao*; 27(2):141-148.

Vallet, Juan. 2006. El positivismo científico de Augusto Comte y lo utilizado de él por la ciencia del derecho en el siglo XX. Consultado Junio 15 de 2007. Disponible en <http://www.racmyp.es/noticias/2006/2006-05-30%20-%20juan%20vallet%20de%20goytisolo.pdf>

Witten, IH; Frank, E. 2005. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco, 2ª edición.