

Indicadores cibernéticos: ¿Nuevas propuestas para medir la información en el entorno digital?

MsC. Ailín Martínez Rodríguez¹

RESUMEN

Se examina la evolución de las disciplinas métricas de la información hasta su inserción en el entorno digital. Se abordan algunas de las características de Internet, el World Wide Web y la comunicación científica en el Web que posibilitan dicha inserción. Se estudian algunas de las principales aplicaciones de los estudios cibernéticos y webmétricos, a partir del uso de indicadores particulares y se analiza su relación con los indicadores empleados en estudios métricos tradicionales de la información.

Palabras clave: Cibermetría, información, Webmetría, indicadores.

ABSTRACT

The evolution of metric information disciplines, up to their insertion in the digital environment is examined. Some of the characteristics of Internet, the World Wide Web and the scientific communication in the Web that enable such insertion are undertaken. Some of the main applications of the cybermetric and webmetric studies are analyzed, from the use of private indicators, and their relation with those employed in traditional metric information studies is analyzed.

Key words: Cybermetric, information, Webmetrics, marks.

Copyright: © ECIMED. Contribución de acceso abierto, distribuida bajo los términos de la Licencia Creative Commons Reconocimiento-No Comercial-Compartir Igual 2.0, que permite consultar, reproducir, distribuir, comunicar públicamente y utilizar los resultados del trabajo en la práctica, así como todos sus derivados, sin propósitos comerciales y con licencia idéntica, siempre que se cite adecuadamente el autor o los autores y su fuente original.

Cita (Vancouver): Martínez Rodríguez A. Indicadores cibernéticos: ¿nuevas propuestas para medir la información en el entorno digital? Acimed 2005;14(4). Disponible en: http://bvs.sld.cu/revistas/aci/vol14_4_06/aci03406.htm Consultado: día/mes/año.

El estudio conceptual de las ciencias métricas de la información se ha tratado por diversos autores, a partir del objeto de estudio de las disciplinas informativas.

La revolución tecnológica y científica ha marcado el ámbito informativo con una cualidad interdisciplinaria, de manera que las ciencias bibliológico-informativas, en múltiples ocasiones, se entremezclan y pierden los límites que las definen.

La matematización de las ciencias, como expresión particular de estas regularidades, tuvo sus primeras manifestaciones y el mayor auge en las ciencias naturales y técnicas. En cambio, este proceso es más lento en el terreno de las ciencias sociales, debido a la complejidad para la medición de los fenómenos que analiza.

La introducción de las matemáticas en las disciplinas sociales tiene sus antecedentes en la doctrina de *Augusto Comte*, el positivismo. Esta corriente no admitía como válidos científicamente otros conocimientos que no procedieran de la experiencia. Esta doctrina tuvo gran influencia en los siglos XIX y XX.

El desarrollo alcanzado por las ciencias tiene entre sus tendencias la aplicación cada

vez más frecuente de las matemáticas y las estadísticas a las llamadas ciencias sociales.

Gorbea representa el proceso de matematización del sistema de conocimientos bibliológico – informativo, como sigue:¹



Las matemáticas aplicadas a las ciencias sociales participan en la solución de problemas prácticos, a la vez, que contribuyen al desarrollo de modelos matemáticos que participan en el estudio de regularidades y tendencias de los sistemas científicos (Pérez Matos NE. Bibliometría o Bibliotecometría. 2003. Observaciones no publicadas).

En 1948, en la Conferencia Annual de la ASLIB en Leanington, Inglaterra, el bibliotecario-matemático hindú *Ranganathan* incorporó el término *Lybrametric*.

A pesar de definirse desde entonces, dicho término es actualmente poco utilizado y que se sustituye con frecuencia por denominaciones como bibliometría, informetría y cientiometría.

Según Pérez, “*todos estos términos se relacionan porque representan una ciencia general, las llamadas ciencias métricas, lo que contribuye a que, en múltiples ocasiones, los modelos, indicadores, índices y demás mediciones se utilicen indistintamente en una u otra ciencia, pero, según algunos autores se distinguen por el objeto de estudio y los objetivos que persiguen dichas ciencias con relación a obtener determinados resultados. En la literatura científica, la Bibliometría es vista como ciencia instrumental de la Bibliotecología, en tanto la Informetría pertenece al mundo de las Ciencias de la Información*”.²

La mención de Pritchard como el primero en definir la bibliometría se ha cuestionado por varios autores, entre ellos, *Ríos* que plantea que existe un conjunto de autores como *Hulme* (1923), *Otlet* (1934), *Zoltowski* (1955) y *Vickery* (1948) que, con anterioridad, habían hecho alguna referencia a este término.³

A pesar de ser *Lybrametrics* y *Bibliometrics* las primeras definiciones conocidas y de donde parte cualquier estudio métrico, en la literatura, se reconoce que la denominada Bibliografía Estadística fue la antesala de la Bibliometría.

Algunos autores sitúan también el término Bibliometría en el “*Tratado de la Documentación*” de *Otlet*, en 1934. Se plantea que este autor insistió en diferenciar la bibliometría de la bibliografía estadística, porque desde sus orígenes, la medida o cuantificación de la ciencia se realizó mediante técnicas estadísticas que se aplicaban a las fuentes de información.

Desde hace algún tiempo, se aplican métodos científicos, fundamentalmente matemáticos y estadísticos, al estudio de la Ciencia. A pesar de que la Ley de Lotka sobre la distribución de la productividad de los autores científicos, se publicó en 1926 y la Ley de Bradford, en 1948, se considera que fue, a partir de la aparición de la obra de *Price* “*Little Science, Big Science*”, en 1963, cuando los estudios métricos de la información científica comenzaron su expansión.

El núcleo de la obra de *Price* es la aplicación de los recursos de la ciencia a su propio análisis. El punto de partida de la llamada *Ciencia de la Ciencia* se considera que fue la obra del británico *Bernal*, “*The Social Function of Science*”, publicada en 1939.

Varios autores plantean que la aspiración de crear una “ciencia de la ciencia” o “cienciología” tropezó, desde sus inicios, con la dificultad de integrar los conocimientos lógicos, sociológicos, económicos, históricos y psicológicos obtenidos sobre el quehacer científico. La ausencia de un marco teórico apropiado para interrelacionar estos enfoques impidió crear una disciplina única para el estudio de la ciencia en su conjunto. Sin embargo, se reconoce que las cuestiones relacionadas con la información científica publicada, la estructuración del sistema de instituciones de la ciencia y la planificación de su labor estaban en el centro de la atención de la “ciencia de la ciencia” o la “cienciología”.

A partir de la obra de *Price*, se desarrolla un campo nuevo: el análisis estadístico y sociométrico de la literatura científica. A este campo se le denomina *Cienciometría*.

Los métodos bibliométricos y cientiométricos son muy similares, a veces perfectamente idénticos. Se basan en el uso de modelos estadísticos. De una parte, analizan el tamaño, crecimiento y distribución de la bibliografía científica (libros, revistas, etc.), con el objetivo de mejorar las actividades de información, documentación y comunicación científica. De otra parte, analizan los aspectos de generación y programación de la literatura científica para llegar al mejor entendimiento de los mecanismos de la investigación científica, considerada como una actividad social, así como la estructura y dinámica social de los que producen y utilizan esta literatura.

El surgimiento del término *Informetría* se atribuye al alemán *Nacke*, quien lo propuso, por primera vez, en el año 1979. Posteriormente, se creó un Comité de la Federación Internacional de Documentación (FID), que lo tomó en su definición, con objetivos bien definidos en relación con el suministro de datos científicos y técnicos, aunque ello no implicó en modo alguno una amplia aceptación del término, si bien se reconoció como campo general de estudio que incluye elementos de la bibliometría y la cientiometría, surgidas anteriormente.

Su aceptación actual data de la “*Conferencia Internacional sobre Bibliometría y Aspectos Teóricos de la Recuperación de Información*”, celebrada en Diepenbeek, Bélgica, en 1987, ocasión en la que se sugirió la inclusión de este vocablo en el nombre de la siguiente conferencia, programada para Londres, en 1989. La referida reunión se denominó entonces “*Conferencia Internacional sobre Bibliometría, Cienciometría e Informetría*”. El nombre del tercer encuentro, que tuvo lugar en 1991, en Bangalore, India, refleja para algunos la aceptación definitiva del término: “*Conferencia Internacional de Informetría*”.

Licea es muy explícita cuando plantea: “*en Occidente, el término Bibliometría fue adoptado de inmediato y sin voces disidentes. Más o menos en la misma época, en los países de la Europa Oriental, se adoptaba el término Cienciometría, para referirse a las técnicas aplicadas a la cuantificación y análisis de las actividades científicas incluida la publicación de libros y revistas. El pedante puede encontrar algunas diferencias de connotación entre estos dos términos, pero, aún si lo hace, las dos líneas de estudios parecen ser convergentes*”.⁴

Esta autora señala respecto al término *Informetría* que éste no ha generado ideas distintas, pero le augura futuro porque cubre formas documentales y electrónicas.

Podría afirmarse entonces que la aplicación y sistematización de la regularidad de matematización de las ciencias, en su expresión particular para la Bibliotecología y la Ciencia de la Información, se encuentran en pleno desarrollo, y por tanto, es objeto de amplia polémica científica, atendido a este proceso.

Más recientemente, *Cronin* propuso el término *Influmetría* para diferenciar los estudios cuantitativos de las comunicaciones informales entre científicos; *Turner*, *Infometría* en sustitución de *Informetría*, y *Kraupskopf*, *Epistemometría* en sustitución de *Cienciometría*. Además, comenzó a utilizarse el término *Patentometría* para el análisis cuantitativo de los documentos de patentes y, por último, el extraordinario impacto de Internet ocasiona que, en búsquedas recientes, se encuentre también, con bastante

frecuencia, el término Webmetría o Cibermetría'

INDICADORES BIBLIO-CIENCIO- INFOMÉTRICOS

En el terreno de los indicadores, las discusiones conceptuales se quedan un tanto al margen. Casi todos los autores se limitan a su aplicación sin delimitar si pertenecen a una u otra disciplina métrica y se limitan a apellidarlos en correspondencia con el enfoque seguido en el trabajo -indicadores bibliométricos, informétricos, cienciométricos, patentométricos, entre otros-, sin diferenciar mucho en el tipo de indicador sino más bien en la finalidad con la que se emplean.

Hasta mediados de la década de los setenta, los indicadores se enfocaban casi exclusivamente a los insumos y no fue hasta el desarrollo masivo de las bases de datos electrónicas a principios de los años 70, que fue posible contemplar la construcción de indicadores de producción científica a partir de sistemas automatizados, por medio de los servicios en línea, por CD-ROM y, en años recientes, en Internet.

Los indicadores, en términos generales, representan una medición agregada y compleja que permite describir o evaluar un fenómeno, su naturaleza, estado y evolución.⁵

Con los indicadores bibliométricos, se podrán determinar, entre otros, los siguientes 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.

Existen variadas clasificaciones para los indicadores biblio- ciencia- informétricos (tabla 1). Excepto en el caso de *Lydesdorff* y *Wouters*, quienes señalan expresamente que se trata de indicadores cienciométricos, el resto de las opciones se refieren a indicadores bibliométricos.

Tabla 1. Clasificación de los indicadores según Leydesdorff y Wouters,⁶ López López,⁷ King⁸ y Vinkler⁹.

Autores	Indicadores
L. Leydesdorff y P. Wouters (1999)	Indicadores como elementos formales de los textos
	<ul style="list-style-type: none"> • Palabras • Copalabras
	Indicadores como elementos formales de los subtítulos
	<ul style="list-style-type: none"> • Citaciones • Nombres de los autores
	Indicadores como elementos formales de los paratítulos
	<ul style="list-style-type: none"> • Encabezamientos temáticos • Direcciones institucionales • (...)
<ul style="list-style-type: none"> • Índice de productividad • Índice de productividad fraccionada • Índice de colaboración 	
Análisis de citas	

P. López López (1996)	<ul style="list-style-type: none"> ● Factor de impacto ● Índice de Visibilidad
	<ul style="list-style-type: none"> ● Índice de asilamiento ● Elance bibliográfico (bibliographic cuopling) ● Cocitación
J. King (King 1978)	<ul style="list-style-type: none"> ● Recuento de publicaciones ● Análisis de citas ● Impacto de las revistas ● Análisis de cocitación ● Análisis de copalabras ● Referencias bibliográficas (enlace bibliográfico)
	Indicadores de publicaciones
	<ul style="list-style-type: none"> ● Número de publicaciones ● Índice de calidad de las revistas ● Índice de actividad ● (...)
P. Vinkler (Vinkler 1998)	Índice de citación
	<ul style="list-style-type: none"> ● Factores de impacto ● Índice de inmediatez ● Factor de popularidad ● Frecuencia de citación ● Índice de atracción ● Nivel de impacto ● (...)

Las discusiones principales entorno a la utilización de los indicadores giran fundamentalmente en torno a la posibilidad de su aplicación y estudio en todas las áreas geográficas, sobre todo, ante las limitaciones presentes en las bases de datos internacionales.

NUEVOS ESCENARIOS PARA LOS ESTUDIOS MÉTRICOS DE LA INFORMACIÓN

El desarrollo de las tecnologías de la información y las comunicaciones ofrece nuevos escenarios para la realización de los estudios métricos de la información, sobre todo de aquella que circula por Internet, con las facilidades del WWW y los cambios que para el modelo tradicional de comunicación científica este desarrollo ha significado.

Internet ha supuesto una revolución sin precedentes en el mundo de la informática y las comunicaciones. Los inventos del telégrafo, teléfono, radio y computadora sentaron las bases para esta integración de capacidades nunca antes vivida. Internet es, a la vez, una oportunidad de difusión mundial, un mecanismo de propagación de la información y un medio de colaboración e interacción entre los individuos y sus pares independientemente de su localización geográfica.

Esta red constituye un universo de recursos de información y un espacio virtual de comunicación entre usuarios, cuyo mayor éxito es el desarrollo de herramientas avanzadas que facilitan el acceso y consulta de los más disímiles recursos de información en cualquier parte del mundo.

Ha pasado de ser un instrumento especializado de la comunidad científica, a ser una red de fácil uso que ha modificado las pautas de interacción social. Esta revolución cambia todos los parámetros de espacio/tiempo en el que se había ubicado hasta la fecha nuestra actividad: el trabajo, el consumo, los aprendizajes, el ocio, lo cotidiano, e incluso las relaciones humanas.

A finales de la década de los años 80, la interconexión de miles de redes de área local había convertido Internet en el mayor almacén de datos que jamás hubiese existido, pero también en el más caótico. Las posibilidades eran enormes, pero las dificultades resultaban frustrantes: formatos incompatibles, programas distintos, protocolos

heterogéneos, etcétera. Se imponía, entonces, la necesidad de simplificar el acceso a este caudal de información, hacerlo más sencillo y homogéneo.

La experiencia de la proliferación del conocimiento y el hecho de no poder abarcarlo todo no es nueva, no ha surgido con el desarrollo de las tecnologías y la conectividad. Ya en 1945, *Vannevar Bush* expresaba:

“La suma de la experiencia humana se está expandiendo a un ritmo prodigioso y los medios que utilizamos para seguir el hilo a través del consiguiente laberinto de ítems momentáneamente importantes son los mismos que usábamos en los días de los barcos de vela”.¹⁰

En su opinión, el problema no era tanto una cantidad excesiva de publicaciones como el nulo avance de las tecnologías con que se gestionaba su manejo. Con los rudimentos tecnológicos de su época en mente, Bush fue capaz de idear un sistema llamado *Memex* que permitiría archivar el conocimiento de un modo más eficaz: una especie de escritorio futurista en el que se guardarían, microfilmados, los libros, actas, etcétera.

Por muchos, *Bush* se considera un visionario. En 1945, sus ideas no eran técnicamente realizables. Ni lo eran aún en 1965, cuando Nelson, las ordenó conceptualmente. Fue Nelson quien acuñó el término hipertexto para referirse a *“un cuerpo de material escrito o gráfico interconectado de un modo complejo que no se puede representar convenientemente sobre el papel; puede contener anotaciones, adiciones y notas de los estudiosos que los examinan”*.¹¹ La idea es que el lector examine los nodos de una red, y pase de unos a otros mediante conexiones. El hecho de que los nodos pueden contener texto, pero también pueden integrar otros medios como imagen, sonido u otras formas es lo que se desea destacar con el uso de otro término complementario: hipermedia.

Durante las dos décadas siguientes, se vivió el auge de las computadoras, el almacenamiento digital y las redes. El propio *Nelson* cobró conciencia de la importancia de estas nuevas tecnologías para la realización del sueño de una red de elementos de información libremente accesible alrededor del mundo. Sin embargo, sus ideas sólo han llegado a concretarse con el World Wide Web.

Según *Gray, Bray y Hobbes*, entre otros, la importancia del Web como sistema de difusión de la información está avalada por el enorme crecimiento, tanto en número de sitios como en el número de páginas Web.

Es este el contexto que ha hecho necesario un nuevo análisis del modelo tradicional de comunicación científica.

El mundo actual es un mundo de redes de información electrónicas, de tecnologías multimedia que cambian el paradigma de la comunicación científica. La aparición de Internet, donde surgen nuevas formas de literatura gris como los foros de discusión, las prepublicaciones electrónicas,... mucho más fluidos que los documentos tradicionales y que escapan a todo tipo de control e identificación, representan un modo revolucionario de comunicación que conduce a la desaparición de soportes y medios considerados como clásicos en los procesos de transferencia de información, a cambio de canales mucho más rápidos y eficaces.

En la red, el concepto de publicar alcanza una consideración eminentemente cuantitativa. Una virtud de las tecnologías electrónicas es eliminar los impedimentos institucionales y económicos para la producción y circulación de documentos. Como se recuerda, a menudo, cualquiera puede producir un documento y hacerlo accesible a millones de lectores. Igualmente, el aumento del número de documentos no tiene precedentes y aumenta drásticamente la proporción de lectores y escritores. Además, las dificultades cuantitativas crecen debido al incremento de las cualitativas. Existen pocos sistemas para el control formal y cualificado de las contribuciones en la red.

Como plantea *Kircz*, el “ *objetivo de la comunicación científica es el registro,*

evaluación, diseminación y acumulación del conocimiento, hechos y percepciones humanas".¹²

La tecnología de la información aporta cambios de gran alcance a los sistemas de comunicación científica. El hecho de que las investigaciones sobre la comunicación académica cobraran nueva importancia, a partir de mediados de los años 90, puede atribuirse a una reestructuración progresiva del sistema de comunicación científica junto con un rápido crecimiento de la tecnología de la información, el trabajo en redes y de las publicaciones electrónicas.¹³

Tanto las comunicaciones formales como informales experimentan alteraciones radicales, de modo que la distinción entre ambas se vuelve cada vez más borrosa. Este desdibujarse de las divisiones establecidas es un elemento clave en el cambio de los medios impresos a los electrónicos. Esto influye no sólo en la manera en que se intercambia información sino también en las instituciones responsables del procesamiento y distribución de la información.

Según Russell, *"este tipo de fenómenos influye en la estructura general del sistema de comunicación científica, transforma las funciones y los papeles de los diferentes actores. Por esto, los modelos tradicionales de edición y comunicación académica están sujetos permanentemente a nuevos análisis en el contexto del panorama actual definido por las tecnologías de la información. La dirección que tomarán estos cambios durante los primeros decenios de este siglo constituyen materia de tentadoras especulaciones"*.¹⁴

Todos los estudios destacan la función fundamental de la comunicación en la actividad científica, primariamente definida como una interacción entre un frente de investigación complejo y volátil, que crea del nuevo conocimiento, y un sistema de comunicación estable y menos flexible que evalúa y difunde este conocimiento. La actividad científica es, en este sentido, una *"serie de creación de nuevos conocimientos"*.¹⁵ Así, se consagra una gran parte de las distintas actividades de los científicos a la comunicación y, sin este paso, la continuidad de la investigación científica no tendría sentido.

El modelo tradicional de publicación evolucionó hasta el modelo editorial actual que ofrece a la vez un soporte físico para la comunicación científica, y un lugar de poder social en construcción, cuyas normas compartidas permiten la construcción de estrategias de comunicación para la transmisión conveniente de los resultados y la estructuración conjunta de la comunidad científica. Sin embargo, las tendencias de evolución de la ciencia y las tecnologías de comunicación e información llevan a reconsiderar los fundamentos de este modelo. Así, la cuestión que desafía actualmente el lugar de todos los protagonistas de la cadena de la comunicación científica no es simplemente la transformación de un soporte, sino la del cambio más amplio del modelo de la publicación científica.

El modelo tradicional de comunicación científica mediante publicaciones periódicas en un ambiente mundial, donde hace tiempo la información es la materia prima más importante y en el que los desarrollos tecnológicos han acelerado las posibilidades del intercambio, aparece como un obstáculo antes que como una herramienta al servicio de la investigación. Se puede establecer un paralelo con la idea de Kuhn, sobre las revoluciones científicas: *"...las revoluciones científicas se inician con un sentimiento creciente de que un paradigma existente ha dejado de funcionar adecuadamente en la exploración de un aspecto de la naturaleza hacia el cual este paradigma había previamente mostrado el camino [...], el sentimiento de mal funcionamiento que puede conducir a una crisis es un requisito previo para la revolución"*.¹⁶

CIBERMETRÍA: ESTADO DEL ARTE

El crecimiento rápido y cada vez mayor de la información electrónica, junto con los amplios potenciales de las tecnologías y de los medios de información emergentes, han atraído la atención de los investigadores de la información con vistas a reflejar

sobre la medida y la métrica cuantitativas de las fuentes de información, de los servicios y de los medios en este entorno.

Los estudios sobre la naturaleza y las propiedades de Internet en general y del WWW en particular, a partir de la aplicación de métodos informétricos para analizar sus contenidos o su estructura de enlaces, no se hacen patentes hasta la segunda mitad de la década de 1990. A partir de esta fecha, se observa un crecimiento progresivo de los trabajos relacionados con la aplicación de las técnicas de medida a Internet; un aumento proporcional al crecimiento de los recursos de información disponibles en la red y, por consiguiente, a la necesidad de poder cuantificarlos, organizarlos e interpretarlos.

Se destacan las investigaciones de:

- *Arnzen* (1996) sobre las que denominó “cibercitas”.
- *Clausen* (1996) que dedicó su estudio a cuantificar los hábitos de los usuarios de Internet y sus actitudes hacia la red como recurso de información en Dinamarca.
- *McMurdo* (1996) que realizó un estudio cuantitativo orientado a los medios en Internet. Conteo de los host de Internet y sus dominios, la distribución de los host por dominios, el crecimiento del Web y las relaciones de transformación de las sedes Web están también entre los parámetros principales de su estudio. En este utiliza algunas fuentes de información estadística y demográfica sobre Internet.
- *Almind e Ingwersen* (1997) que realizaron una de las investigaciones principales sobre la métrica del ciberespacio, a partir de la aplicación de métodos informétricos al WWW, y que denominaron Webmetría. Su estudio comparaba la proporción danesa en el Web con relación a la de otros países nórdicos. Este estudio exploró, por primera vez, el número medio de hiperenlaces por página y la densidad de enlaces para los diferentes tipos de dominio. Para este estudio la metodología utilizada fue de análisis bibliométrico.
- *Abraham* (1997) que empleó el término Webmetría, y planteó la necesidad de aplicar técnicas de redes neuronales para lograr un mejor conocimiento del Web, y representar las conexiones de los nodos mediante números reales, que indicaran la fuerza de la conexión.
- En 1997 se realizó una investigación, en la Escuela Real de Bibliotecólogos de Dinamarca, dirigida a explorar por estudios cuantitativos ciertos fenómenos y acontecimientos actuales de la información. En este estudio se refiere la Internetmetría y está más orientado a la información que las investigaciones anteriores.
- *Ingwersen* (1998) que realizó una investigación en la que se informa sobre la viabilidad y fiabilidad en el cálculo del factor de impacto de las sedes Web. El estudio demuestra que el Factor de Impacto Web es calculable y fiable con la precaución necesaria para estimar el número de las páginas del Web que señalan a las páginas de una sede determinada.
- *Dahal* (1999) que aplicó las leyes bibliométricas al análisis del desarrollo de los sistemas de información en ciencia y tecnología en Nepal, donde finalmente emplea el término Cibermetría para explicar las técnicas empleadas.

Cronin, afirma que los modelos y métodos utilizados en los estudios biblio- ciencia- informétricos pueden aplicarse en el entorno electrónico de Internet:

*“Efectivamente, si la informetría estudia todos los aspectos cuantitativos de los procesos de información en general, e Internet se ha erigido como la mayor fuente de información electrónica, no es extraño que, desde mediados de la década de 1990, se apliquen los modelos y métodos informétricos al ámbito de Internet y, más concretamente, al WWW”.*¹⁷

Con el ritmo actual de incremento de los recursos electrónicos, el usuario final de dicha información se ha visto sobrepasado por la ingente cantidad y diversidad de contenidos, el aparente carácter caótico de su organización, su alta volatilidad, la enorme heterogeneidad formal de sus recursos y las dudas sobre la calidad última de ciertos documentos.

Es precisamente en este entorno, donde nace el concepto de Cibermetría, un nuevo campo de estudio que se identifica con diferentes denominaciones y cuya relación formal con la informetría queda reflejada en un artículo de la revista *Cybermetrics* de 1997, donde se expresa, *“la Informetría investiga los aspectos cuantitativos de los procesos informativos (comunicativos), especialmente de los textuales; constituye un arma cuantitativa de la Ciencia de la Información y de la Biblioteconomía; y engloba tanto la Bibliometría tradicional como las nuevas áreas relacionadas con la Cibermetría y la Webmetría.”*¹⁸

Uno de sus precursores fue *William Paisley* quien, en el año 1990, señaló la importancia de la aplicación de los métodos informétricos al campo de la comunicación electrónica:

“En el futuro, una gran proporción de todos los textos que actualmente aparecen en libros, revistas científicas, comerciales, y periódicos, estarán depositados en bases de datos electrónicas. De hecho, las bases de datos electrónicas incluirán, tanto los registros publicados como los no publicados. Esta amplia colección de información electrónica es el ámbito futuro de la investigación bibliométrica”.¹⁹

En el año 1995, *Bossy* señala que *“utilizando Internet como instrumento de comunicación se pueden descubrir y cuantificar los flujos de información establecidos entre los científicos, así como detectar y medir el grado de colaboración que existe entre ellos”*.²⁰

Para *Aguillo*, la Cibermetría *“ mide distintos aspectos de Internet utilizando técnicas cuantitativas bibliométricas que han mostrado ser especialmente potentes y pueden aplicarse en ciertas condiciones ventajosas a la descripción de recursos Web”*.²¹

En paralelo, se presenta el término Webmetría que se refiere a la aplicación de la informetría u otras técnicas cuantitativas al entorno del WWW.

Turnbull afirma que la *“Bibliometría es un método normalizado para el análisis de la información perfectamente aplicable al desarrollo y medición de la información disponible en el Web; la idea consiste en usar las aplicaciones bibliométricas tradicionales en el entorno Web con el fin de estudiar los modelos de comunicación, las distintas áreas de investigación, realizar estudios históricos sobre el desarrollo de una disciplina o dominio, y evaluar la investigación por países, instituciones o individuos”*.²²

Abraham y Foresta, incorporaron una nueva acepción al vocablo para referirse a *“una técnica que nace con el objetivo de elaborar mapas cognitivos y modelos matemáticos del 3W que permitan entender su funcionamiento como prototipo de organización social, cultural y política del ciberespacio”*.²³

Pueden observarse entonces, dos vertientes de análisis: por una parte la Cibermetría, con todas sus variantes terminológicas, que estudia la aplicación de las tradicionales técnicas informétricas a cualquier tipo de información disponible en Internet; y la Webmetría, más restringida y que forma parte de la anterior, basada en la aplicación de la Informetría, u otras técnicas de medida, específicamente a la información disponible en el entorno Web.

Las técnicas métricas se han enfrentado, casi desde su surgimiento, a problemas e incomprendimientos de orden teórico; la ignorancia de sus fundamentos, categorías y leyes, así como también, la poca atención a sus aspectos cualitativos, ha generado cierto distanciamiento de los niveles científicos.

Según *López Yepes* *“calidad y cantidad son la punta del iceberg del conflicto permanente en las operaciones de evaluación y ambas aparecen, en ocasiones, plenas de sombras y de ambigüedades. Es evidente que el número de publicaciones pertenece al ámbito de la cantidad –considerar que existen mensajes científicos no publicados-, pero no es tan cierto que el análisis de citas represente la calidad*

inexorablemente".²⁴

El número de artículos publicados por un autor no arroja información sobre la calidad de los artículos. La frecuencia con que un autor se cita no se traduce necesariamente en un estadio de calidad o experiencia técnica. Un autor puede citarse frecuentemente por varias razones, incluido el hecho de que muchos otros autores discrepan con sus escritos. La frecuencia con que un autor publica ofrece información sobre su productividad, pero no sobre la calidad del producto.

Sin embargo, estas no serían razones suficientes para eliminar de raíz los estudios cuantitativos. Un estudio estadístico no prueba definitivamente que una hipótesis es justa, pero permite verificar que los hechos observados en medio de un procedimiento determinado no la niegan. Los números, los gráficos, las tablas, aunque no hablan por sí solos, permiten descubrir interdependencias y diferencias entre fenómenos no apreciables a simple vista.

Parece evidente que la aplicación de la métrica y de las medidas cuantitativas a la información electrónica se está convirtiendo en un área significativa para la investigación.

Es cierto que las cifras rara vez, revelan algo que no se haya pensado previamente, pero la aplicación de la cibermetría, permite analizar los datos con vistas a revelar las líneas de fuerza, los ritmos de crecimiento, las tendencias, las mutaciones coyunturales, las expectativas de futuro y las demandas sociales de este fenómeno que es Internet.

APLICACIONES DE LA CIBERMETRÍA Y LA WEBMETRÍA

Internet, como enorme autopista de la información, ha proporcionado argumentos para el estudio.

- El número, el alcance y los temas de las redes de información.
- Distribución de las redes por países.
- Volumen y tipología de las colecciones de información disponibles, según tamaño y tipo.
- Distribución de los diversos tipos de redes.
- Evaluación de las redes, según su tiempo de respuesta y acceso.

En el caso del correo electrónico pueden realizarse las siguientes apreciaciones:

- Número de direcciones de correo electrónico.
- Distribución de direcciones por países e instituciones.
- Uso del correo en los sectores público y privado.
- Distribución de usuarios por profesiones o empresas.
- El volumen, el tipo y el tamaño de la información enviada mediante el correo.
- Las proporciones de diversos tipos de documentos enviados por correo electrónico.

Desde un punto de vista exclusivamente webmétrico, *Dhyani, Keong y Bhowmick* consideran que las técnicas de medición pueden aplicarse fundamentalmente a las siguientes categorías del World Wide Web.²⁵

- El número de sedes Web y de páginas de inicio en el mundo y también su distribución por países.
- Clasificación de las páginas Web por tipos de documentos.
- Número de páginas Web por dominios.
- Clasificación de páginas Web por el idioma de los documentos y por los modos de representación de la información.
- Estadísticas de uso y usuarios de las páginas Web en un período de tiempo determinado.
- El número de citas recibidas por cada página Web.

- Ordenar los Web más citados y páginas personales según el tipo de documento.
- Los tipos de colecciones electrónicas disponibles en cada sede Web.
- Factor de impacto del Web y productividad de los autores.
- Análisis del contenido de las páginas Web.
- Identificar la variedad de publicaciones electrónicas por el tipo, el idioma y la distribución geográfica.

El estudio métrico de los recursos electrónicos, que también podrían considerarse como elementos de la comunicación científica en este espacio, como libros, revistas, bibliotecas y fuentes de referencia, pueden elaborarse otras investigaciones que permitan reconocer la transición revolucionaria del mundo impreso al electrónico. Para tener una idea del análisis cuantitativo de estos recursos electrónicos, algunos de los principales aspectos que pueden tratarse son:

- Estadísticas de bibliotecas digitales.
- Número de revistas electrónicas según su temática e idioma.
- Número de revistas publicadas en ambos formatos (electrónico e impreso).
- Número de obras de referencia disponibles electrónicamente.
- Análisis de citas de revistas electrónicas.
- Utilización de las revistas electrónicas.
- Distribución de recursos electrónicos por tipo, país e institución.
- Productividad científica en el entorno electrónico.
- El crecimiento y obsolescencia de la literatura electrónica.

Estas son las áreas más reconocidas que pueden cuantificarse mediante medidas estadísticas y técnicas informáticas, que según *Shiri* pueden agruparse en cuatro áreas fundamentales: “*las redes de información, el correo electrónico, la WWW y los recursos electrónicos*”.²⁶

El estudio de la comunicación científica se asocia al de algunas regularidades presentes en los flujos de información como son el crecimiento y la obsolescencia de la información científica –el crecimiento de la literatura activa y de su vida media citada y citante–, así como otras más propias del proceso de comunicación como son el impacto y la visibilidad de las fuentes de información y de sus autores, mediante la identificación de relaciones entre investigadores, áreas y disciplinas, así como de aquellas que se generan entre los métodos formales e informales de comunicación.

Por medio del análisis de las múltiples aplicaciones o campos de acción, puede observarse que la Cibermetría es una materia relativamente reciente con un carácter multidisciplinario.

La mayor parte de los autores expresa la importancia de la aplicación de los estudios cibernéticos y webmétricos para la recuperación y presentación de la información.

Los estudios de visibilidad, de densidad, los análisis de citas, la investigación sobre el diámetro del Web; nos adentran en un mundo absolutamente novedoso y prácticamente inexplorado hasta el momento. Todo ello de basa en la utilización de los llamados Indicadores cibernéticos.

INDICADORES CIBERMÉTRICOS

Los estudios realizados por diferentes autores para desarrollar sus análisis informáticos del ciberespacio se han incrementado notablemente en los últimos años. La finalidad de estas investigaciones es obtener una serie de parámetros que permitan conocer la presencia de cualquier entorno en Internet, posibilitar el desarrollo de estudios comparativos que abarquen campos tan dispares como: dominios geográficos (.es, .fr., .uk, etc.) y funcionales (.net, .org, .edu, etc.), instituciones públicas y privadas, documentos electrónicos, espacios Web científicos y comerciales, contextos sociales, regiones geográficas, etcétera. Con los resultados obtenidos en estas investigaciones, es posible conocer la situación particular de un ámbito determinado y emprender actuaciones de alcance internacional, nacional, regional, local o sectorial.

No abundan en la literatura sobre el tema textos que aborden la aplicación de las métricas en el entorno digital desde un punto de vista teórico o conceptual. La mayoría de los trabajos que se encuentran constituyen análisis de algunos indicadores que permitan identificar el comportamiento de determinados fenómenos que sirven de base para tomar decisiones. Por esta razón, no pueden mencionarse propuestas metodológicas concretas, que incluyan las propuestas de indicadores cibernéticos utilizados. En este sentido, cada autor/investigador, ha aplicado aquellos que ha estimado relevantes para cumplir con los objetivos de su investigación.

La siguiente lista recoge los indicadores más frecuentes empleados para el análisis métrico de los recursos digitales en diversos trabajos sobre la temática.

- Indicadores de tipos institucionales -cantidad de páginas en determinados sectores.
- Indicadores regionales.
- Indicadores idiomáticos.
- Indicadores de tipología de sitios -sitios académicos, comerciales, de sectores públicos o privado.
- Indicadores de tamaño, en sus dos variantes, tamaño documental -número total de páginas comprendidas en un dominio) o tamaño informático -tamaño en bytes de una sede Web.
- Indicadores de densidad, también tiene dos variantes, densidad hipertextual -media de enlaces por página- y densidad multimedia -media de objetos multimedia por página.
- Indicadores de profundidad -número máximo de niveles de una sede.
- Indicadores de luminosidad -total de enlaces emitidos desde 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.
- Indicador de navegabilidad -número total de enlaces internos respecto al total de páginas.
- 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.
- Indicadores para el estudio del comportamiento de usuarios en la recuperación de información.

Aguillo realiza una agrupación de los indicadores según las variables o categorías más generales que permiten evaluar.²⁷

Categoría	Indicador
Infraestructura	Número de host, de servidores Web, de usuarios, de dominios, de sitios, de sitios institucionales, etcétera.
Tamaño	Número de páginas, de objetos, de objetos multimedia, de archivos ejecutables, tamaño de los archivos, distribución por lenguajes, evolución temporal, número de niveles, de enlaces por página, etcétera.
Calidad	Porcentaje de enlaces válidos, de errores de enlace, apariencia, etcétera.
Conectividad	Total de enlaces, de enlaces por página, número de enlaces internos, de enlaces externos, etcétera.
Visibilidad	Número de enlaces recibidos o externos, enlaces nacionales externos, enlaces internacionales externos, etcétera.

Impacto	Factor de impacto
Popularidad	Número de visitas

En forma general, estos indicadores pueden agruparse en tres grandes grupos o tipos de medida:

- Medidas descriptivas: Como su nombre indica miden fundamentalmente el tamaño o número de objetos encontrados en cada sede -riqueza de páginas, ficheros media o ricos, densidad de enlaces totales y únicos. Se utilizan para medir la penetración de Internet desde el punto de vista de los contenidos en países, regiones, organizaciones o grupos de individuos.
- Medidas de visibilidad e impacto: Se basan en el carácter hipertextual del Web y exploran los patrones de enlace entre páginas y sedes distintas. El número y diversidad de enlaces externos recibidos, su volumen respecto a los contenidos objeto de enlace (llamado Factor de Impacto Web) o índices que se construyen de acuerdo con el peso relativo de las sedes de origen de los enlaces, como el famoso PageRank de Google, que permite establecer listas ordenadas, según la jerarquía numérica de estos indicadores.
- Medidas de popularidad: El consumo de información medido en términos de número y características de las visitas de una desde el Web resulta un atractivo, aunque extremadamente complejo de implementar, método de evaluación. Es ciertamente interesante para estudios temporales, en los que la medida de la evolución resulta prioritaria para los correspondientes informes. Como se indica es complicado obtener valores absolutos, pero ciertos valores relativos con sesgos importantes pueden, no obstante, utilizarse en análisis comparativos.

Las dos primeras clasificaciones se corresponden con la realizada cuando se trata de los indicadores métricos tradicionales, entre los que se distinguen dos tipos fundamentales: indicadores de actividad -número y distribución de publicaciones, productividad de autores, etc.- e indicadores de impacto -número de citas recibidas, factor de impacto, etcétera.

Los indicadores que más se utilizan hasta la fecha para realizar estudios cibernéticos se adaptan especialmente a la clasificación de King (tabla 1), quien diferencia entre dos tipos básicos de indicadores o medidas.

MEDIDAS DESCRIPTIVAS O INDICADORES DE RECuentOS

En la cibermetría, estos indicadores descriptivos, además del conteo de los artículos producidos y publicados por la comunidad científica en el entorno electrónico de Internet, también incluye el recuento de diferentes aspectos de los recursos en la red.

Entre estas medidas, se encuentran:

- El tamaño medio de los documentos analizados.
- Los protocolos utilizados por los URLs de los documentos HTML analizados.
- Los tipos de ficheros.
- El recuento de dominios científicos.
- La tipología documental de las páginas Web.
- Los recursos: página Web con datos textuales o audiovisuales.
- El número medio de enlaces por página.
- La densidad media de enlaces.
- El tamaño documental.
- El tamaño informático.
- La densidad hipertextual.
- La densidad multimedia.
- La profundidad.

A estos, pueden añadirse otros elementos como: el número de revistas electrónicas según su temática e idioma, el número de revistas publicadas en formato electrónico e impreso, el número de obras de referencia disponibles electrónicamente, la distribución

de recursos electrónicos por tipo, país e institución, así como la productividad científica en el entorno electrónico.

Estos últimos elementos apuntarían, sobre todo, a la medición de la comunicación científica en el Web y como puede observarse constituyen sólo adaptaciones al entorno digital, porque se utilizan también en los estudios métricos tradicionales.

Medidas de calidad

a) Análisis de citas.

La evaluación del verdadero alcance del paralelismo entre los procesos de citación y "sitación" es objeto de una línea de investigación aún en sus primeras fases.

Aunque no existe una herramienta similar a las bases de datos del ISI y, en especial al SCI, resulta factible realizar un análisis de citas de los espacios Web mediante el estudio de sus enlaces, si se asume que los enlaces que salen de y entran en los espacios Web son equiparables, respectivamente, a las referencias y citas en las publicaciones científicas tradicionales.²⁷

Estas comparaciones han generado diferentes posiciones entre los autores:

a) A favor de la analogía: *McKiernan, Larson, Almind, Ingwersen, Vreeland, Björneborn, Cronin, Zhu, Hong y Hogle*. b) En contra: *Egghe, Harter, Ford, Kim y Raan*.

Algunos investigadores, como *Rousseau*, se sitúan en una posición intermedia al considerar que el estudio de la "sitación" (sitiation), término definido por *McKiernan* en 1996, para designar la relación entre los sitios de Internet, aunque es conceptualmente igual al de citación tradicional, tiene un sentido ligeramente diferente, porque los enlaces y las citas están motivados por cuestiones distintas. Aunque las motivaciones de la citación científica son diversas, según *Brooks*, se acepta que se trata de un proceso de persuasión científica; sin embargo, los enlaces electrónicos responden a un comportamiento "científico-social-tecnológico", según *Kim*.

Además de los distintos motivos que inducen a realizar, tanto una cita/referencia como un enlace, tradicionalmente el análisis de citas se ha aplicado a las publicaciones científicas, mientras que las "unidades de conexión" como las denomina *Berners - Lee* o enlaces hipertextuales no se establecen únicamente entre textos científicos. Por ello, algunos autores como *Borgman* y *Furner* consideran más apropiado utilizar el término "análisis de enlaces", debido a que su significado es mucho más amplio que el de cita/referencia y recoge todas las posibles modalidades de citación que pueden darse también en la Web.

Como se refirió, no existe un modelo de análisis de citas consolidado y alternativo al del ISI, pero existen algunas propuestas interesantes en el ámbito de las publicaciones electrónicas. Entre ellas, se encuentra **Citebase**, una base de datos "open access" que permite navegar por los archivos de *eprints* (prepublicaciones electrónicas) mediante citas hipertextuales y que posibilita su análisis. Esta experiencia piloto forma parte del proyecto de investigación " **The Open Citation Project** " cuyo objetivo es elaborar herramientas útiles para el análisis de citas en publicaciones electrónicas "open access".

b) Factor de impacto (FI)

El Factor de impacto Web (FIW) es uno de los primeros indicadores examinados en los trabajos webmétricos y, además, uno de los más estudiados. Los análisis sobre esta medida se inician con *Ciolek y Rodríguez y Gairín* quienes investigan el FIW de diversos dominios geográficos.

La fórmula del FIW se podría generalizar de la siguiente manera:

$$FIW = \frac{\text{Sitas recibidas por un espacio web}}{\text{Páginas del espacio web}}$$

Para determinar el numerador y el denominador de la fórmula, la mayoría de las investigaciones han aplicado las posibilidades de recuperación que ofrece el motor de búsqueda AltaVista.

El cálculo de ambos elementos (numerador/denominador) no es tarea fácil porque, a los problemas relacionados con el método de recopilación del número de sitas y de páginas Web, hay que añadir los vinculados con su propia naturaleza. Con respecto a las tipologías de FIW, *Smith* señala tres tipos:²⁹

- FIW total: número de sitas recibidas por un espacio Web dividido entre el número de páginas de dicho espacio.
- FIW externo: número de sitas externas recibidas por un espacio Web dividido entre el número de páginas de dicho espacio. El FIW externo probablemente sea el mejor indicador para medir la importancia de un espacio Web en la comunidad Web externa.
- FIW interno: número de sitas internas recibidas por un espacio Web dividido entre el número de páginas de dicho espacio. Un espacio Web con un elevado FIW interno funcionaría en gran medida como una Intranet.

Existen una serie de problemas relacionados con el FIW que se centran fundamentalmente en los métodos de recopilación de sitas y páginas Web, es decir, del numerador y denominador, respectivamente, y en la propia naturaleza de ambos elementos. Con respecto al método de recopilación, no todos los motores ofrecen iguales posibilidades para realizar un estudio webmétrico, y particularmente para calcular el FIW. Aparentemente, AltaVista es el motor que mejor se adapta al estudio del FIW, porque reúne una serie de características, señaladas por *Smith* como imprescindibles, para realizar un estudio de este tipo:²⁸

- Dispone de una de las mayores bases de datos.
- Permite buscar todas las páginas incluidas en un dominio de alto nivel (comando *domain:*), de bajo nivel (comando *host:*), y en conjuntos de directorios (comando *url:*).
- Incluye la posibilidad de localizar las páginas que sitan a un espacio Web determinado (comando *link:*).
- Ofrece la posibilidad de combinar resultados con lógica booleana.

ANÁLISIS DE CO-CITAS Y DOCUMENTOS RELACIONADOS

Rousseau introdujo el término *sitas*, producto de la fusión de los términos ingleses *site* [sitio web] y *cite* [cita]), y por derivación, se utiliza *co-sitación*.

Existen múltiples motivos por los que se pueden cositar dos W e b: la proximidad geográfica ó idiomática, ó la similitud de contenidos. Esta técnica es, en principio, una de las maneras más sencillas de obtener información relacional, entre dos sitios Web determinados.

Uno de los trabajos hallados en Webmetría sobre el análisis de co-sitas, se encontró en 1996 con *Ray Larson* quien realiza uno de los primeros estudios de co-sitación por medio de técnicas multivariantes para descubrir la estructura intelectual del Web más allá de los simples directorios jerárquicos tipo Yahoo, a partir del hecho de que la sita de páginas Web es un elemento fundamental que provee la conexión transversal de contenidos informativos de la red.

Otros indicadores

Aunque son pocos los trabajos encontrados hasta el momento, existen algunos dedicados a analizar la posibilidad de la adaptación de las leyes bibliométricas más

utilizadas, propuestas por *Lotka*, *Bradford* y *Zipf*, al entorno digital.

El Page Rank, incorporado en Google, también se incluyó por algunos autores entre los indicadores cibernéticos. Se trata de un sistema de clasificación de páginas Web desarrollado por los fundadores *Larry Page* y *Sergey Brin* en la Universidad de Stanford. Su finalidad es asignar un valor numérico a las páginas Web, según el número de veces que otras páginas las recomienden y según el PageRank que éstas ostenten. Es decir, establece la importancia de cada página Web. Su lógica es la siguiente: si una página Web enlaza otra página, es que la recomienda. Y si la recomienda, es que debe ser importante en el ámbito del tema que trata la primera página Web. Una recomendación que proviene de una página que, a su vez, es muy recomendada vale más que una recomendación que provenga de una página que casi nadie recomienda.

Google desea que, en las primeras posiciones de los resultados de búsqueda, se encuentren páginas de cierta relevancia y que éstas se recomienden por otras páginas que, a su vez, también sean relevantes. Para determinar el PageRank, Google analiza el número de enlaces que provienen de otras páginas Web y su PageRank.

Este indicador apunta hacia una de las principales aplicaciones que para los autores tienen los estudios webmétricos, la recuperación de la información.

HERRAMIENTAS TECNOLÓGICAS EMPLEADAS PARA DESARROLLAR ESTUDIOS CIBERMÉTRICOS

Entre las principales, se encuentran:

- Agentes mapeadores (gestores de sitios Web):
- Astra Site Manager 2.0 (<http://www.mercury.com/>)
- COAST Web Master 7.0 (<http://www.coast.com/>)
- Custo 2.0 (<http://www.netwu.com/>)
- Funnel Web Profiler 2.0 (<http://www.quest.com/>)
- Microsoft Site Analyst 2.
- Microsoft Content Analyzer 3.0 (<http://www.microsoft.com/siteserver/>)
- WebKing 4.1 (<http://www.parasoft.com/>)
- Web Trends 7.1 (<http://www.netiq.com/>)

- Verificadores de enlaces.

- Web Link Validator 3.5 (<http://www.relsoftware.com/wlv/>)
- Verificadores Online
- W3C Link Checker (<http://validator.w3.org/checklink/>)
- Xenu Link Sleuth 1.2f (<http://home.snafu.de/tilman/xenulink.html>)

- Volcadores de sitios Web.

- Website Extractor 9.07 (<http://www.asona.org/>)
- WebCopier Pro 4.0 (<http://www.maximumsoft.com/>)
- WebZip 6.0 (<http://www.spidersoft.com/>)
- Teleport Pro 1.29.2074 (<http://www.tenmax.com/>)

La combinación de diferentes herramientas propias de la cibermetría, como son los motores de búsqueda y el software comercial (mapeadores), para la extracción de datos cuantitativos de las sedes Web previamente identificadas permite obtener datos complementarios para la realización de estudios de este tipo. Mientras los primeros resultan más apropiados para la extracción de detalles sobre la repercusión que cada unidad tiene (visibilidad), los segundos lo son a la hora de construir indicadores sobre sus características más generales -tamaño, riqueza hipertextual, etcétera.

Varias disciplinas influyen notablemente en los sustentos de los estudios cibernéticos. Los autores se refieren entre otras, a las aportaciones de la informática, la física, las

matemáticas, las ciencias sociales; y, sobre todo, a la gran influencia de la teoría de grafos, si se considera el análisis del Web como grafo dirigido, que permite desarrollar métodos para el análisis de su estructura.

Otras teorías que ofrecen elementos para realizar estos estudios es el análisis de las redes sociales, las redes neuronales artificiales, la teoría de redes complejas y la teoría del caos. Algunas de ellas, se consideran para presentar los resultados de estas investigaciones.

Con todos estos elementos, se aprecia que existen determinados problemas para realizar investigaciones cibernéticas. Estas problemáticas están se basan, sobre todo, en: el carácter dinámico y cambiante del Web, con criterios como los de persistencia e intermitencia; el tamaño del Web que presenta grandes dificultades para poderse determinar, incluido el tamaño del Web invisible, que limita el valor de los resultados obtenidos y los métodos de extracción de datos, ante las dificultades observadas en los motores de búsqueda como principales herramientas utilizadas en este sentido.

No obstante, se aprecia un creciente trabajo en la aplicación de los estudios cibernéticos, en ocasiones, relacionados con indicadores para medir la sociedad de la información o del conocimiento. De alguna manera, se puede expresar que aún estas investigaciones se encuentran en estado de gestación y por ello, muchos se dedican actualmente a la búsqueda de los indicadores más apropiados para su implementación en cada caso.

CONSIDERACIONES FINALES

- Aunque existe determinado reconocimiento sobre la existencia de un cuerpo teórico de conocimientos relativos a la aplicación de métodos matemáticos aplicados a la Bibliotecología y la Ciencia de la Información, éstos constituyen campos de investigación en los que aún la polémica científica es frecuente.
- Los estudios métricos de la información constituyen un terreno aún en pleno desarrollo y por ello, cada día surgen nuevas posibilidades de aplicación para los resultados obtenidos. Cada vez, se utilizan más con fines de agregar valor a la información, no sólo en unidades dedicadas al trabajo intensivo con esta, sino también para la toma de decisiones en cualquier tipo de organización. Así, se convirtieron en herramienta importante para los sistemas de inteligencia empresarial, y útiles, sobre todo, en los aspectos relacionados con la vigilancia tecnológica.
- Es necesario reformar el modelo de la comunicación científica ante la presencia de los cambios que introdujo la evolución y desarrollo de las nuevas tecnologías de la información y las comunicaciones y más concretamente Internet y el World Wide Web.
- La aparición y desarrollo de Internet en general y el WWW en particular, llevaron a la implementación de indicadores métricos para medir la información en este entorno digital; ellos se reconocen en la literatura, como indicadores cibernéticos o webmétricos. Estos estudios son vitales por su contribución a la presentación y recuperación de la información.
- Aunque con importantes limitaciones y sesgos, las herramientas más adecuadas para extraer los datos necesarios para construir indicadores cibernéticos son los motores de búsqueda. Estos pueden cumplir las mismas funciones que las bases del ISI tienen en los estudios bibliométricos, aunque existen algunos aspectos que es necesario considerar: ningún motor ofrece una cobertura completa o exhaustiva del Webespacio, incluso presentan una distribución heterogénea de los contenidos indizados -ciertas áreas geográficas están peor revisadas. Los motores tienen un comportamiento irregular con resultados que varían en períodos muy cortos y sin razón aparente. Además, las cifras que proporcionan suelen redondearse y, en algunos casos, no son muy fiables. Sin embargo, un análisis de las prestaciones y características diferenciadas de cada herramienta puede permitir un uso combinado de ellas que minimice los efectos negativos citados.
- Los mecanismos de extracción de datos incluyen el uso de los llamados delimitadores, operadores capaces de filtrar según ciertas características los contenidos. Los más importantes son los de dirección (domain, site, url) y los de enlace (linkdomain, link), aunque los de tipo de fichero (filetype) e idioma

(languague) son también muy útiles. Algunos de estos delimitadores se pueden combinar con operadores booleanos. No todos los motores permiten estas opciones; en el 2005, sólo se podían usar con fines cibernéticos los motores: Google, Yahoo Search, MSN Search y Teoma.

- Los métodos cibernéticos suponen una importante contribución a la batería de herramientas disponibles para la evaluación de la actividad científica. Aunque originalmente propuestos para complementar y comparar los resultados obtenidos por medio de técnicas bibliométricas pueden ser también útiles para describir otro tipo de patrones relacionados con los procesos de comunicación científica no tan formales como la publicación académica. La propia estructura hipertextual del Web y la organización fuertemente institucionalizada de los contenidos ofrecen nuevas oportunidades para descubrir y describir patrones de comunicación inéditos entre los diferentes actores del sistema ciencia-tecnología-sociedad.
- De manera general, los indicadores cibernéticos abarcan cuatro áreas fundamentales: Internet, World Wide Web, correo electrónico y comunicación científica en el Web.
- Los indicadores referidos a Internet se utilizan como criterios para evaluar sitios Web. Entre ellos, se encuentran los que se refieren a: profundidad, conectividad, tamaño, navegabilidad. Estos son los indicadores cibernéticos que no presentan análogos en el terreno de los estudios métricos tradicionales. En cambio, los que se utilizan para analizar visibilidad e impacto, e incluso los de recuento, cuando se trata de elementos de la comunicación científica en el Web, son sólo adaptaciones a este entorno de los indicadores tradicionales - relacionados con análisis de citas, factor de impacto, productividad científica.
- Se puede plantear que estas adaptaciones responden no a un cambio en la esencia de estos indicadores, sino a cambios en la manera de representar, almacenar y recuperar la información en este sistema específico, factor que también puede ser determinante en las denominaciones de los estudios métricos de la información, en tanto cada tipo de estudio, en el cual se enmarcan los indicadores que se evalúan, responden a las características de los sistemas de información evaluados. En el caso de los indicadores cibernéticos, las características medibles de Internet, el World Wide Web y la comunicación científica en este entorno determinan las adecuaciones para el uso de los indicadores.

REFERENCIAS BIBLIOGRÁFICAS

1. Gorbea Portal S. Modelación matemática de la actividad bibliotecaria: una revisión. Investigación Bibliotecológica 1998;12(24). Disponible en: <http://www.ejournal.unam.mx/biblio/vol12-24/1B102402.pdf> [Consultado: 20 de diciembre del 2005].
2. Pérez Matos NE. La bibliografía, bibliometría y las ciencias afines. Acimed 2002;10(3). Disponible en: http://bvs.sld.cu/revistas/aci/vol10_3_02/Aci012002.htm [Consultado: 20 de diciembre del 2005].
3. Ríos D. La bibliometría: nivel de penetración en la enseñanza bibliotecológica y su aplicación en el campo bibliotecario en los países del MERCOSUR. Disponible en: <http://www.ifla.org/IV/ifla66/papers/162.127.htm> [Consultado: 15 de enero del 2006].
4. Licea de Arenas J. Indicadores de la actividad científica. Ciencias de la Información 1993; 24(1):2-6.
5. Martínez E, Albornoz M (ed). Indicadores de ciencia y tecnología: estado del arte y perspectivas. Caracas: Nueva Sociedad, 1998.
6. Leydesdorff L, Wouters P. Between texts and contexts: advances in theories of citation? Scientometrics 1999;44(2):173-92.
7. López López P. Introducción a la Bibliometría. Valencia: Promolibro, 1996.
8. King J. A review of bibliometric and other science indicators and their role in research evaluation. Journal of Information Science 1987;13:261-76.
9. Vinkler P. An attempt of surveying and classifying bibliometric indicators for scientometric purposes. Scientometrics 1988;13(5-6):239-59.
10. Bush V. As we may think. Atlantic Monthly 1945;176(1):101-8.
11. Nelson TH. A file structure for the complex, the changing and the indeterminate.

- En: ACM 20 th National Conference. Proceedings of the 1965 ACM 20 th National Conference, Cleveland , USA . New York C: ACM Press; 1965. p. 84-100.
12. Kircz J. Scientific communication as an object of science. Disponible en: <http://www.science.uva.nl/projects/commphys/papers/aceur.htm> [Consultado: 5 de abril del 2005].
 13. Borgman CL. Digital libraries and the continuum of scholarly communication. *Journal of Documnetation* 2000;56(4):412-30.
 14. Russell JM. La comunicación científica a comienzos del siglo XXI. Disponible en: <http://www.campus-oei.org/salactsi/rusell.pdf> [Consultado: 8 de abril del 2005].
 15. Vassallo P. The knowledge continuum - organizing for research and scholarly communication. *Internet Research: Electronic Networking Applications and Policy* 1999;9(3):232-42.
 16. Kuhn T. *La structure des révolutions scientifiques*. Paris: Flammarion, 1962.
 17. Cronin B. Bibliometrics and beyond: some thoughts on web-based citation analysis. *Journal of Information Science* 2001;27(1):1-7.
 18. CINDOC.CSIC. España. Cybermetrics: International Journal of Scientometrics, Informetrics and Bibliometrics. Disponible en: <http://www.cindoc.csic.es/cybermetrics/> [Consultado: 17 de febrero del 2005].
 19. Almind TC, Ingwersen P. Informetrics Analyses on the World Wide Web: Methodological Approaches to Webometrics. *Journal of Documentation* 1997;53(4):404-26.
 20. Bossy M. The Last of the Litter: Netometrics. *Solaris* 1995;(2). Disponible en: <http://www.info.unicaen.fr/bnum/jelec/Solaris/d02/2bossy.html> [Consultado: 20 de enero del 2006].
 21. Aguillo IF. Indicadores hacia una evaluación no objetiva (cuantitativa) de sedes Web. En: VII Jornadas Españolas de Documentación, octubre 19-21, Bilbao, España. Bilbao: Servicio Editorial del País Vasco; 2000. p. 233-48.
 22. Turnbull D. Bibliometrics and the World Wide Web. 2000. Disponible en: <http://donturn.fis.utoronto.ca/research/bibweb.html> [Consultado: 5 de abril del 2006].
 23. Abraham RH, Foresta D. Webometry: chronotopography of the World Wide Web. 1996. Disponible en: <http://thales.vismath.org/webometry/articles/prague.html> [Consultado: 16 de marzo del 2006].
 24. López Yepes J. La evaluación de la ciencia en el contexto de las Ciencias de la Documentación. Disponible en: <http://www.urgitxetas.net/articulos/html/0001.htm> | [Consultado: 26 de febrero del 2006].
 25. Dhyani D, Bhowmick SS, Keong Ng W. Web informetrics: extending classical informetrics to the Web. En: Proceedings of the 13th International Workshop on Database and Expert Systems Applications (DEXA'02), 2002. Washington DC : IEEE Computer Society; 2002. p. 351-5.
 26. Shiri AA. Cybermetrics; a new horizon in information research. Disponible en: <http://www.deck.com/people/pages/shiri/cybermet1.html> [Consultado: 3 de marzo del 2006].
 27. Aguillo IF. Indicadores cibernéticos: midiendo y evaluando los contenidos de la Sociedad de la Información. En: INFO 2004. La Habana: IDICT; 2004.
 28. Faba C, Guerrero V, de Moya Anegón F. Fundamentos y técnicas cibernéticas. Disponible en: <http://www.google.com/cu/search?q=cache:OskSAatp7-CJ:www.juntaex.es/consejerias/ect/dgsi/Doc> [Consultado: 24 de abril del 2006].
 29. Smith AG. A tale of two Web spaces: comparing sites using Web impact factors. *Journal of Documentation* 1999;55(5):577-92.

Recibido: 28 de agosto del 2006. Aprobado: 12 de septiembre del 2006.

MsC. *Ailín Martínez Rodríguez*. Bibliotecología y Ciencia de la Información. Facultad de Comunicación. Universidad de La Habana. Calle G No.506 entre 21 y 23. El Vedado. Plaza de la Revolución. Ciudad de La Habana. Cuba. CP 10 400. Correo electrónico: ailin@fcom.uh.cu

Ficha de procesamiento

Clasificación: Artículo de revisión.

Términos sugeridos para la indización

Según DeCS¹

BIBLIOMETRIA; ANALISIS CUANTITATIVO; CIENCIA/estadísticas y datos numéricos; CIENCIAS DE LA INFORMACION; INTERNET.
BIBLIOMETRICS; CUANTITATIVE ANALYSIS; SCIENCE/ statistical and numerical data; INFORMATION SCIENCE; INTERNET.

Según DeCI²

INFORMETRIA/desarrollo; BIBLIOMETRIA/desarrollo; CIENCIOMETRIA/desarrollo; ANALISIS CUANTITATIVO; INTERNET.
INFORMETRICS/development; BIBLIOMETRICS/development; SCIENTOMETRICS/development; CUANTITATIVE ANALYSIS; INTERNET.

¹BIREME. Descriptores en Ciencias de la Salud (DeCS). Sao Paulo: BIREME, 2004.

Disponible en: <http://decs.bvs.br/E/homepagee.htm>

²Díaz del Campo S. Propuesta de términos para la indización en Ciencias de la Información. Descriptores en Ciencias de la Información (DeCI). Disponible en: <http://cis.sld.cu/E/tesauro.pdf>

¹**Máster en Bibliotecología y Ciencia de la Información. Profesora Instructora. Bibliotecología y Ciencia de la Información. Facultad de Comunicación. Universidad de la Habana.**

Índice Anterior Siguiente