

Los retos de la transición hacia el ámbito digital

De acuerdo a algunos estudios realizados con anterioridad la bibliometría es la parte de la cienciometría que aplica métodos matemáticos y estadísticos a la producción científica con la finalidad de estudiar y analizar dicha actividad, para lo cual se auxilia de leyes bibliométricas basadas en el comportamiento estadístico mediante indicadores. En este sentido se observa que la literatura científica presenta un crecimiento exponencial y en el caso de los investigadores también se ha incrementado su número pero no todos son igual de productivos, razón por la cual hay que seguir de cerca el número de citas que reciben sus publicaciones.

De este modo, al medir la productividad científica, es decir la cantidad de publicaciones producidas por un autor, país o institución durante un periodo determinado, se han desarrollado diversas tendencias en las aplicaciones bibliométricas. De ahí la intención de la Dirección General de Bibliotecas de abordar dicha temática en su XIV Reunión sobre Revistas Académicas y de Investigación, realizada en el marco de la XXXV Feria Internacional del Libro de Minería, en febrero de 2014.

La reunión fue inaugurada por el director general de Bibliotecas de la UNAM, doctor Adolfo Rodríguez Gallardo, quien dijo que el tema general giró en torno a *Los retos de la transición hacia el ámbito digital* y su objetivo se centró en presentar experiencias de los diferentes actores que participan en el campo de las publicaciones electrónicas científicas.

De esta forma el ponente David Pendlebury, de la compañía Thomson Reuters, abrió la sesión de la primera mesa titulada "Tendencias en aplicaciones bibliométricas" y señaló que su intervención giraría en torno a una interpretación práctica debido a que en bibliometría hay una proliferación de nuevas tendencias o metodologías, también hay un interés por las medidas que tienen valor predictivo. Por su parte, la tecnología se ha encargado de proporcionar bases de datos de texto completo permitiendo un análisis detallado del contenido, contexto y entorno donde se llevan a cabo las citas. El *open access* también es un tema relacionado pues de alguna forma afecta a las publicaciones y se debe conocer hasta qué grado va a influir en las citas y su impacto.

Al hacer un poco de historia, Pendlebury dijo que hace 50 años fue creado el *Science Citation Index* y las citas que se registraron en aquel entonces tenían cierta autoridad porque provenían directamente de los científicos. Más tarde, hacia los años ochenta, se integraron las citas para poder rastrear el desempeño nacional de la ciencia; posteriormente, en 1990 llegaron los buscadores encabezando la lista Google y a partir de 2005 apareció el indicador denominado índice H, considerado como una métrica reciente que ha tenido gran influencia en esta materia.



Ahora bien, si se desea hacer un análisis de las citas es necesario observar en la cita el sesgo de la métrica, ya sea que se analice por instituciones, campos o periodos anuales. Entonces, la pregunta obligada sería el por qué no se utilizan ciertos factores para evaluar a personas y artículos; una posible respuesta sería porque la media en el factor de impacto lo determina el sesgo de la distribución y eso equivale a que la mayoría de las ocurrencias que se miden en lo que respecta a los artículos por individuo relacionados con el factor de impacto de una revista van a ser menores con la media. Probablemente no se está hablando de una métrica sólida, específicamente al valorar el desempeño, es por esta razón que recientemente se ha tenido gran interés en el uso de los percentiles como una manera de caracterizar exactamente en dónde se ve la distribución del artículo de una revista; el año de publicación también juega un papel importante en el análisis de las citas.

En la parte de la distribución, por ejemplo, si el porcentaje de los artículos de la institución aparecen clasifi-

cados en el primer 10% de las citas, quiere decir que es la base y el punto de referencia; si se está por debajo de éste significa que no se está cumpliendo con los estándares más altos de excelencia y se podría recurrir a las interpretaciones de los percentiles, en particular al analizar la producción por individuo.

Por último, David Pendlebury habló del acoplamiento bibliográfico; es decir, si dos artículos citan la misma referencia probablemente hay una relación, pero si dos artículos se citan como pares en repetidas ocasiones en otros artículos se trata de una cocitación. Este método se utiliza en una escala multidimensional para definir mapas de la ciencia, y cuando se realiza dicha tarea se pueden observar los niveles de actividad y áreas que van surgiendo o que se alejan, es como un mapeo anualizado.

Este método se inició hacia los setentas y se fue extendiendo y cobró mayor fuerza a principios del año 2000, debido a que los datos ya estaban dispuestos en grandes cantidades y de esa forma se podían utilizar conjuntos de datos más grandes y manejables desde la comodidad de las computadoras personales. Ahora se crean vínculos en las áreas de especialización por año comparándolo con otro, pues se necesitan representaciones visuales para mostrar la vasta información. Es por esto que la colaboración es un tema de interés, aunado a que existen más herramientas, mejores técnicas, y regresando un poco a la colaboración, se da entre instituciones y autores.

En seguida Salvador Gorbea, del Instituto de Investigaciones Bibliotecológicas y de la Información, se avocó a hablar de la minería de datos dentro del contexto de las métricas de la información; al respecto, puntualizó que no es fácil conjugar las tendencias transdisciplinarias que se están dando entre los estudios métricos relacionados con otros campos como la computación y la teoría de la organización, de ahí la necesidad de desarrollar un esbozo de un modelo teórico sobre el desarrollo de los estudios métricos de la información con la intención de mitigar la discrepancia que existe en lo referente a la terminología de cienciométrica, o habría que llamarle infometría, o elementos de carácter métrico y cuantitativo. Al respecto, se pudo concluir que estas tres dimensiones conforman un

gran campo al que se le podría llamar estudios bibliométricos de la información, sin existir discrepancia entre las tres dimensiones porque comparten objetos de estudio pero con temas diferentes.



De igual forma se encontraron trabajos que indicaban que la bibliometría, la ciencimetría y la webmetría pueden ser consideradas como manifestaciones de una sola área que es la métrica de la información. Al contextualizar, se determina que dentro de la minería de datos se encuentra la bibliometría, que es la aplicación de la minería de datos combinada con esta disciplina y con las herramientas estadísticas para extraer patrones de comportamiento basados en los recursos digitales de un sistema bibliotecario.

También se revisaron otras investigaciones que enfocaban su estudio a la aplicación de las “minerías” en el campo de las ciencias de la información y que al mismo tiempo proponían utilizar el término de bibliominería para identificar patrones de comportamiento informativo en las bibliotecas.

En seguida surgieron controversias en su identificación semántica, pero ambas especialidades se ocupan del análisis de datos mediante técnicas estadísticas para establecer patrones y tendencias en los datos. Y se diferencia una de la otra porque la bibliominería trata datos sobre actuación previa a la autorización de la información que está almacenada, y la bibliometría trabaja con datos relativos a la información que ha sido finalmente usada, producida y comunicada.

Pero, a final de cuentas el propósito común es generar conocimiento y regularidad en los flujos de la información documental e integración entre los estudios métricos de información, sin perder de vista que la bibliometría identifica irregularidades de producción y comunicación científica, mientras que la minería de datos identifica patrones del comportamiento de la información que no ha podido ser clasificada. Del mismo modo, hay que considerar que el cúmulo de datos que arroja la industria, la empresa y las organizaciones han propiciado un cambio de enfoque referente a los datos, es decir han dejado de ser solo objetos de almacenamiento, búsqueda y recuperación de la información para convertirse en una fuente de conocimiento.

Finalmente, el panelista indicó que este cambio de paradigma sobre el tratamiento métrico de la información, así como de las nuevas exigencias que se derivan de los sistemas de gestión del conocimiento, amplían y ponen a prueba el rol bibliotecario y de los profesionales de la información porque tienen que incorporar estos novedosos procesos como una herramienta útil para el perfeccionamiento de los sistemas bibliotecarios.

Posteriormente, Mike Taylor, de la Editorial Elsevier, dijo que aunque suene muy parecido la almetría con la bibliometría son términos diferentes. La primera es una colección de ideas y perspectivas, por lo tanto tiene datos muy distintos a la segunda, porque ésta considera la citación de páginas, es decir, las reseñas que se hacen de los artículos publicados en la World Wide Web. También, se cuenta con otras modalidades que aplican *bookmark* en el trabajo de los investigadores para organizar sus documentos; además, se puede saber quién ha guardado, leído o compartido un artículo y aunque es una actividad distinta al consumo del artículo ha ido ganando terreno. Lo mismo sucede con la citación de artículos en páginas globales como los periódicos, que son medios abiertos y no son consultados necesariamente por los académicos.

Otro indicador que tiene que ver con la almetría es el reuso de la información, porque cuando un investigador la hace disponible se puede saber cuántas personas la han consultado y publicado nuevos artículos a partir de otros

artículos, que más tarde se suben a las redes sociales. Debido a su magnitud se deben hacer análisis equiparables para llegar a una conclusión de lo que hace la gente con los artículos científicos, tarea nada fácil porque a diario se generan miles de tweets; no se puede trabajar en el análisis de citas como hace cuarenta años.

Esta necesidad de las personas de establecer redes sociales cobra cada vez mayor fuerza, porque basta con un *click* para que se suba el artículo y otro para conocer las reacciones y el impacto logrado. Es por esta razón que en la actualidad existen plataformas que a través de diversos colores indican la creciente actividad en las redes sociales y a las cifras que arrojan se les denominó como valor alométrico. Aunque sea una nueva modalidad se sabe que los *tweets* no son un buen elemento predecible en el campo de las citaciones porque no tienen que ver con el mismo parámetro de la citación.

Los *tweets* tampoco reparan en la longitud del artículo, los antecedentes del académico y la unicidad del tema entre otras facetas que tienen cierto impacto, por eso hay que tener cuidado sobre los sistemas que inciden en nuestra práctica cuando hablamos de métodos de “predicaciones”; no se debe tomar a la ligera. En cambio, los métodos académicos permiten aumentar la visibilidad de las investigaciones.

Otro aspecto que se analizó en las redes sociales es que identifican el perfil de los usuarios, saben qué les interesa, qué consultaron y con quiénes se relacionan; es una especie de monitoreo personal para luego realizar sugerencias que se quedan grabadas en la interfaz. Este aspecto de la almetría merece atención porque representa un uso potencial de las métricas alternas; quizá en unos años se puedan citar los datos de manera correcta así como sus vínculos, para que las herramientas prueben su utilidad o quizá en un futuro surjan nuevos indicadores que se inspiren en la almetría.

La segunda mesa de trabajo abordó el tema de las “Experiencias de estudios bibliométricos en América Latina.” Al respecto, SciELO Brasil reportó que entre las vicisitudes a las que se han enfrentado constantemente tienen que ver con las métricas de seguimiento



de desempeño, para lo cual se apoyaron en los estudios bibliométricos aplicados al análisis de las revistas que están alojadas en su plataforma, para saber tanto de las citas que recibe como de las que concede. Esta actividad forma parte del programa más importante de este proyecto, entre otras razones porque el servicio en línea de información científica de cada país (o lo que se conoce como Web of Science) se traduce en la capacidad de un país para producir ciencia. Por lo tanto, saber producir buenas revistas es parte del avance de las infraestructuras nacionales de investigación.

En cuanto a la participación del Laboratorio de Dinámica no Lineal de la Facultad de Ciencias de la UNAM, a través del Dr. Humberto Carrillo Calvet, estuvo enfocada a la explicación de métodos y aplicaciones científico-métricas a partir de un sistema de software, en el cual es posible llevar a cabo todo el proceso de descubrimiento de conocimiento en bases de datos y obtener representaciones visuales, denominado ViBliosOM (Visualización Bibliométrica con el SOM).

Esta herramienta –se explicó– desarrolla un ejercicio automático de minería de datos basado en algoritmos matemáticos de la estadística y la inteligencia artificial con el apoyo de redes neuronales, pues constituyen un recurso del campo emergente de la inteligencia computacional para el análisis exploratorio de datos y visualización de información. En este sentido, ViBliosOM se utiliza para el análisis de bases de artículos de investigación científica.

Por último, Antonio Sánchez Pereyra, de la DGB, habló sobre la pertinencia de una existencia real de una bibliometría de revistas latinoamericanas, porque se sabe que este tipo de revistas están clasificadas en la periferia política del mundo. Dicho de otro modo, se trata de una ciencia periférica que con grandes esfuerzos logró que la red regional SciELO y la empresa Thomson Reuters establecieran un convenio de colaboración para la creación de SciELO Citation Index (SCI) y de este modo esta base de datos fuera albergada en la plataforma de búsqueda y descubrimiento denominada Web of Science (WoS), con la finalidad de dar mayor visibilidad internacional a las revistas producidas en la región.

El panelista añadió que en este escenario SciELO México tuvo una importante actuación, pero no hay que perder de vista que en la DGB se cuenta con el portal BIBLAT, concebido desde sus inicios como una fuente de información que proporciona datos bibliométricos sobre la producción científica publicada en las revistas latinoamericanas y del Caribe e indizadas en las bases de datos CLASE y PERIODICA, distinguidas por su carácter multidisciplinario, consideradas como indicadores de producción que a su vez pueden dividirse en indicadores de coautoría, colaboración, por autor, por revista, por concentración y dispersión de la producción. Otra característica radica en su presentación estética, amigable y funcional.

Para la presentación gráfica de la información de las revistas se organiza por zona núcleo y luego van de menor a mayor producción; ambos grupos a pesar de que cumplen con todos los requisitos y su regularidad es constante no son suficientemente ponderados porque no se ubican en los índices más importantes del mundo.

En cuanto a la colección SciELO México, fue incorporada a la Scientific Electronic Library Online y en su totalidad cuen-

ta con mil 150 revistas, 32 mil fascículos, 500 mil artículos en texto completo de acceso libre y hasta la fecha se han sistematizado más de 10 millones de citas bibliográficas.

Añadió que el convenio antes mencionado implicó vincularse a un universo gigantesco de aproximadamente 24 mil 700 revistas y el proceso de antelación de SCI estuvo basado en una actividad laboriosa y compleja que tiene que ver con productos de gran relevancia, con la implementación de indicadores y el manejo de potentes herramientas bibliométricas disponibles en WoS. A su vez, este conjunto de bases de datos también enriqueció sus fuentes de información al recibir los registros de SciELO.

La tercera mesa versó sobre “la gestión y evaluación de recursos electrónicos” y en primer término se explicó el funcionamiento del Sistema Bibliotecario de la Universidad de los Andes, en Colombia, que hasta diciembre de 2013 suscribió un total de 85 mil títulos de libros electrónicos, pues durante los últimos seis años ésta fue la tendencia en el desarrollo de sus colecciones. En cuanto a los criterios de evaluación, se consideró el acceso simultáneo remoto por rango IP, la capacitación del personal en el uso del recurso, la modalidad de perpetuidad de los contenidos, las estadísticas de uso, el factor de impacto, la pertinencia de contenidos y el suministro de registro MARC, entre otros factores.



Los retos que se enfrentaron en este proceso tuvieron que ver con la negociación con los proveedores, la capacitación de los bibliotecarios, el factor tecnológico y la difusión del material. Al final se consiguió la optimización de los recursos económicos, la reutilización del espacio físico para ampliar los puestos de lectura, la facilidad de búsqueda y una óptima recuperación y administración de la información.

La segunda participación de esta mesa estuvo a cargo de Natalia Timiraos de BioMed Central quien mostró un nuevo producto llamado *Paper*, que funciona como herramienta de referencia que tiene la capacidad de clasificar, etiquetar y organizar el contenido. Fue diseñado para usarse en computadoras y diversos dispositivos móviles con la intención de localizar cualquier documento desde un simple reporte hasta proyectos e investigaciones, porque *Paper* es una solución de flujo de trabajo que permite poner en sincronía el iPad y las laptops.

El sistema también permite renombrar y reestructurar los documentos, o lo que es lo mismo se trata de personalizar el trabajo para poder encontrar el contenido una vez que se haya cargado la aplicación. Posteriormente se puede buscar y compartir el trabajo con los colegas.

Se puede decir que hay una similitud con *iTunes* porque los documentos también se visualizan por título o PDF. Otra función tiene que ver con el uso de código de colores o realizar una búsqueda dentro del mismo recurso y, próximamente, se está tratando de incluir repositorios institucionales.

En este orden de participaciones, Adacelia López Roblero, del Sistema de Información Bibliotecario de ECOSUR, reflexionó sobre el papel mediador que ejercen las bibliotecas en el campo de las adquisiciones y la formación de recursos humanos para que sea capaz de acceder a las bases de datos científicas. A la par de esta tarea, las bibliotecas se han tenido que ajustar a los tiempos modernos; esto significa que en los últimos diez años ha evolucionado de ser una biblioteca tradicional a una híbrida y luego a una digital.

El siguiente paso está encaminado de convertirse en una biblioteca más integral, mejor dicho totalmente digital y colaborativa; claro está, con los problemas que conlleva la obtención de tecnologías heterogéneas, que van desde los equipos de cómputo hasta los grandes servidores, o los que tienen que ver con el ancho de banda.

Esto por un lado, y por el otro se tienen las habilidades y competencias desiguales con los países desarrollados, y no sólo de los bibliotecarios sino también por parte de los usuarios. También, se tiene que batallar porque hay muchos recursos de información similares, una falta de representatividad de editores y una diversidad de modelos que presentan los proveedores.

Adacelia López añadió que por todas estas desavenencias diversas instituciones mexicanas de educación superior suscribieron un convenio para la creación de Ciberciencia, que es un proyecto de gestión de información que permite a los usuarios de los Centros Públicos de Investigación del Consejo Nacional de Ciencia y Tecnología el acceso a bases de documentos y recursos electrónicos internacionales, suscritos en consorcio por las instituciones mexicanas participantes, con el objetivo de garantizar el acceso a la información científica de calidad.

Por último, la cuarta mesa se abocó a la exposición de las experiencias concretas de algunas entidades educativas sobre el uso de sistemas de información basados en las llamadas herramientas de descubrimiento. ☞

Reseña: Ma. del Rosario Rodríguez León

Fotos: Julio Zetter Leal

Secretaría Técnica de Difusión y Relaciones
Dirección General de Bibliotecas - UNAM