
CIENCIOMETRÍA, INFORMACIÓN E INFORMÁTICA EN CIENCIAS BIOLÓGICAS: ENFOQUE INTERDISCIPLINARIO PARA ESTUDIAR INTERDISCIPLINAS

LAYLA MICHÁN

A partir de 1950, las tecnologías de la información y la comunicación (TICs) iniciaron un desarrollo notable. Esto ha repercutido en la transformación de la comunicación y en las relaciones sociales, económicas, políticas y culturales del mundo, como consecuencia del incremento en las capacidades para generar, sistematizar, compartir, transmitir, analizar y difundir la información. Esta (r)evolución informática está caracterizada por la adopción de las computadoras como instrumento de comunicación, el uso del formato digital, la masificación y el uso del Internet; la personalización, automatización, actualización e inmediatez de la información a través de la Web, influye y es influida por el progreso científico y tecnológico del siglo XX. Referirse a información en pleno siglo XXI implica la mención de términos, métodos, teorías novedosas e innovadoras como: sociedad del conocimiento, sociedad de la información, globalización, infodiversidad, acceso a la información, e-ciencia, e-investigación, cómputo en *grids*, laboratorios, conocimiento basado en la literatura, minería de textos (*text mining*), web semántica, índice de impacto, cocitación, web 2.0 y 3.0, redes sociales, plagio, acceso libre, derecho al olvido, computación en nube (*cloud computer*), ontologías y dispositivos móviles, por mencionar las más frecuentes (Macías y Michán 2009).

La explosión de la información científica y la producción acelerada de una gran variedad de programas, aplicaciones, herramientas, utilidades recursos y servicios electrónicos para la investigación científica, disponibles a través de la Web, en donde se publica constantemente información digital diversa y colosal en diversos formatos que se sistematiza en bases de datos, catálogos o listas y puede consultarse de forma libre o restringida es una constante. Estos cambios, sin duda, han repercutido en la transformación de la práctica científica a varios niveles como: el propio objeto de estudio, la dinámica de la producción científica, las relaciones entre los científicos, los métodos y técnicas, la evaluación y la difusión científica. Se

Departamento de Biología Evolutiva, Facultad de Ciencias, Universidad Nacional Autónoma de México. / laylamichan@ciencias.unam.mx

ha reducido la energía, el costo y el tiempo requeridos para el análisis de la información científica; se han desarrollado nuevas técnicas analíticas, tecnologías de acceso y modelos de organización para explotar las colecciones digitales de manera innovadora como la bibliometría, la minería de textos (*text mining*), el análisis de redes y la semántica. Se diseñan a diario nuevas herramientas para realizar búsquedas más eficientes y precisas, así como para hacer metanálisis mejores y más extensos (Layla Michán, Muñoz-Velasco, Álvarez, y Macías 2011).

Se han producido formas sistémicas de analizar la información en cantidades colosales (terabites). El acceso inmediato a estos repositorios de datos que resguardan una inmensa variedad de temas y fuentes especializadas de diversas procedencias ha fomentado la integración y la multidisciplinariedad del conocimiento. Se ha iniciado el uso de estándares, lenguajes y ontologías que permiten la interoperatividad y la facilidad de compartir los datos digitales, lo que ha propiciado el aumento de la colaboración entre diferentes disciplinas y permitido que investigadores, estudiantes, educadores e instituciones localizadas en diversas áreas geográficas interrelacionarse. Además, el medio de comunicación científica por excelencia, la publicación científica a través de las revistas electrónicas, actualmente se han vuelto más dinámicas y permiten una comunicación más completa, eficiente e inmediata entre los grupos científicos.

Este fenómeno ha aumentado la eficacia y la exactitud en la obtención de información y ha generado nuevas oportunidades de investigación al incrementar el espectro de opciones técnicas y análisis factibles. La automatización ha permitido a los investigadores desarrollar modelos y simulaciones de fenómenos, lo que ha promovido la aceleración del conocimiento y la facilidad de su distribución.

La repercusión de esta transformación ha sido tal que se han acuñado nuevos términos para referirse a esta nueva forma de hacer ciencia: e-ciencia (*e-science*), e-investigación (*e-research*) y ciberinfraestructura (*cyberinfrastructure*). La e-ciencia resulta del uso y aplicación de la ciberinfraestructura en la práctica científica, la cual se caracteriza por la inter y multidisciplinariedad, colaboración, la participación de un gran número de investigadores (en algunos casos cientos) localizados en diversas regiones y con diferentes especialidades que forman grupos trabajo (Craddock, Harwood, Hallinan y Wipat, 2008; Layla Michán, Muñoz-Velasco, Alvarez, y Macías 2011; Rambo 2009; Watson, Trefethen y Vander Meer 2010). La ciberinfraestructura se refiere al entorno tecnológico-social que permite crear, difundir y preservar los datos, información y conocimientos mediante la adquisición, almacenamiento, gestión, integración, informática, minería, visualización y otros servicios a través de Internet. Incluye un conjunto interoperable de diversos elementos: 1) Infraestructura, los sistemas computacionales (hardware, software y redes), servicios, instru-

mentos y herramientas; 2) Colecciones de datos sistematizadas en bases de datos, como índices catálogos, repositorios, y 3) Grupos virtuales de investigación (colaboratorios y observatorios) (Michán, Russell, Sánchez-Pereyra, Llorens-Cruset y López-Betrán 2008; L. Michán y (comps.) 1999; Layla Michán 2009; Layla Michán, Castañeda Sortibrán, Rodríguez-Arnaiz, y Ayala 2010; Layla Michán y Llorente-Bousquets 2010; Layla Michán y S. Michán 2010).

Sin duda, en la actualidad, vivimos la era de las ciencias biológicas, y la influencia de las tecnología de la información y la comunicación han transformado nuestra visión de la realidad biológica radicalmente. Ya no se explora únicamente a través de los métodos tradicionales, como la observación, la comparación, los experimentos y los modelos *in vivo* e *in vitro*, sino que actualmente muchos de los descubrimientos se hacen *in silico*, procesando información digital con base en modelos reales. Este formato permite analizar y compartir la información, facilitar el acceso inmediato a una variedad extensa de colecciones de datos especializadas de diversas procedencias y temas. Esta nueva tecnología faculta una reinención de la ciencia ya que tiene la capacidad de abordar el desafío que implica manejar conjuntos de datos grandes y complejos. Como resultado de este complejo proceso, las ciencias de la vida han revolucionado y evolucionado rápidamente; su producción tiene un crecimiento exponencial, se utilizan métodos innovadores, y han surgido nuevos campos del conocimiento, como la informática biológica (*biological informatics*), ciencia que se enfoca al manejo y análisis de los diferentes tipos de información biológica, con ramas que incluyen la bioinformática, la neuroinformática, la informática de la ecología, la informática de la biodiversidad, la e-taxonomía o cibertaxonomía, informática de la nanotecnología, biología de sistemas, la genómica y varios enfoques relacionados denominados *omics* (proteómica, transcriptómica y metabolómica) (Kitching, Mayo y Scoble 2009; Zauner 2009).

Uno de los propósitos del proyecto "Información, informática y ciencia en ciencias biológicas" es investigar el desarrollo, tendencias del impacto de la revolución digital en biología y a través de la evolución que ha sufrido la propia información biológica, en especial las publicaciones académicas, así como las teorías, métodos, y aplicaciones informáticas más novedosas e innovadoras para la recuperación, manejo y metanálisis acorde con las necesidades y retos de nuestro tiempo.

Información. Nuestro propósito es entender las transformaciones tecnológicas, metodológicas y sociales de la información en formato digital producida por los investigadores en biología, sobre todo la representada en la literatura primaria (artículos), en la que se somete a evaluación y escrutinio de los pares académicos los nuevos conocimientos científicos y niveles superiores de organización, como las revistas y colecciones de

literatura. Además, nos interesan los procesos involucrados en la producción de la literatura como la publicación, evaluación, distribución y los permisos de acceso, por ejemplo.

Informática. Nuestra finalidad es identificar, ordenar y clasificar todos los recursos electrónicos disponibles en la Web más relevantes para la recuperación, el manejo y el análisis de literatura académica de manera eficaz y eficiente. Hacerlo con base en distintos tipos y características de la información digital, sus usos y aplicaciones, su importancia e implicaciones, en especial resaltan las bases de datos que contienen colecciones bibliográficas y el software especializado para procesarla, con una amplia gama de objetivos que van desde mejorar los procesos de recuperación de literatura, obtener nuevo conocimiento biológico o incluso aplicarlo en la toma de decisiones producto de la evaluación, gestión y política científica (Macías y Layla Michán 2009; Layla Michán, Muñoz-Velasco, Alvarez, y Macías 2011).

Cienciometría (también denominada bibliometría, infometría y/o cibermetría en muchos casos). Se refiere al análisis sistémico y simultáneo de la producción científica, con la finalidad de obtener mediciones que permitan establecer estructuras, comparaciones, relaciones y tendencias institucionales, grupales, regionales, nacionales y/o mundiales, respecto a los elementos teóricos, metodológicos y sociales de la práctica científica (Michan, et al. 2008; L. Michán y (comps.) 1999; Layla Michán 2009; Layla Michán, et al. 2010; Layla Michán y Llorente-Bousquets 2010; Layla Michán y S. Michán 2010).

De tal manera aplicamos un enfoque interdisciplinario y multidisciplinario entre la información (ciencias de la información y documentación), informática y ciencias de la computación, la cienciaometría y las ciencias biológicas. Estudiamos un fenómeno que integra a la información, la informática y la biología, a través del surgimiento y de nuevas disciplinas biológicas como la bioinformática, la neuroinformática, la e-taxonomía, por mencionar algunas. Para este fin utilizamos e integramos métodos, lenguajes y discursos de disciplinas distintas, lo que ha resultado en la participación de distintos especialistas en los proyectos, la formación de estudiantes de diferentes disciplinas, la impartición de cursos en diversas áreas y la publicación de resultados en foros y revistas disímolas. Es claro aquí que la interdisciplina es una de las características que nos define.

De hecho, no nos queda lugar a dudas de que es la manera más completa de entender objetos y fenómenos producto de la sincronía en la era actual que son en extremo mixtos, singulares, heterogéneos y convergentes.

REFERENCIAS

- Craddock, T., Harwood, C. R., Hallinan, J., & Wipat, A. (2008), "e-Science: relieving bottlenecks in large-scale genome analyses," *Nature Reviews. Microbiology*, 6 (12): 948-954. Nature Publishing Group. doi: 10.1038/nrmicro2031.
- Kitching, I., Mayo, S., & Scoble, M., with Godfray. (2009), "Taxonomy as an e-Science," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 367 (1890). doi: 10.1098/rsta.2008.0190.
- Macías, L. & Michán, Layla (2009), "Los recursos de la Web 2.0 para el manejo de información académica," *Fuente 1* (1): 18-27. Retrieved from http://fuente.uan.edu.mx/publicaciones/01-01/los_recursos_de_la_web_2.0_para_el_manejo_de_informacion_academica.pdf.
- Michan, L., Russell, J., Sánchez-Pereyra, A., Llorens-Cruset, A. & López-Betrán, C. (2008), "Análisis de la sistemática actual en Latinoamérica," *Interciencia* 33 (10): 754-761. INTERCIENCIA. Retrieved from http://www.interciencia.org/v33_10/754.pdf.
- Michán, L. y Llorente, J. L. (1999), "La taxonomía en México, en la segunda mitad del siglo XX, autores y revistas nacionales," *Publ. Doc. Mus. Zool. UNAM* 3.
- Michán, Layla (2009), "Las revistas y la institucionalización de la sistemática en América Latina," *Revista Mexicana de Biodiversidad* 80 (1): 105-117.
- Michán, Layla; Castañeda Sortibrán, A.; Rodríguez-Arnaiz, R. & Ayala, F. J. (2010), "Global *Drosophila melanogaster* Research: a bibliometric analysis," *Drosophila Information Service* 93. Retrieved from <http://www.ou.edu/journals/dis/DIS93/Michan232.pdf>.
- Michán, Layla & Llorente-Bousquets, J. (2010), "Bibliometría de la sistemática biológica sobre América Latina durante el siglo XX en tres bases de datos mundiales," *Revista de Biología Tropical* 58 (2): 531-545. Retrieved from <http://www.biologia.ucr.ac.cr/rbt/attachments/volumenes/vol58-2/01-Michan-Bibliometria.pdf>.
- Michán, Layla & Michán, S. (2010), "El desarrollo de la biogerontología y geriatría de inicios del siglo XX a la actualidad," in L. M. Gutiérrez Robledo & J. H. Gutiérrez Ávila (eds.), *Envejecimiento humano*. México: Institutos Nacionales de Salud, pp. 137-145.
- Michán, Layla, Muñoz-Velasco, I., Alvarez, Eduardo, & Macías, L. (2011), *Biomedical Web, Collections and Meta-Analysis Literature Applications*.
- Rambo, N. (2009), "E-science and biomedical libraries," *Journal of the Medical Library Association: JMLA* 97 (3): 159-161. doi: 10.3163/1536-5050.97.3.001.
- Watson, P., Trefethen, A. & Vander Meer, E. (2010), "e-Science: past, present and future II," *Physical and Engineering Sciences* 368 (1926): 4003-4003. doi: 10.1098/rsta.2010.0178.
- Zauner, H. (2009), "Evolving e-taxonomy," *BMC Evolutionary Biology* 9 (1). doi: 10.1186/1471-2148-9-141.