

---

## LAS BIO-ONTOLOGÍAS: BIOLOGÍA, FILOSOFÍA E INFORMÁTICA

LAYLA MICHÁN  
EDUARDO ÁLVAREZ

En este texto nos referiremos a un caso interesante, exitoso e innovador de la interacción entre la ciencia, la filosofía y la tecnología: las bio-ontologías. En la biología del siglo XXI es necesario organizar, clasificar y especificar el conocimiento digital de un dominio biológico dado para asociarlo a las entidades biológicas correspondientes con un significado preciso (Schulze-Kremer y Smithin, 2005). Para lograr este objetivo ha sido útil e indispensable pensar a las ciencias de la vida junto a la filosofía y desde la visión de la tecnología. Junto con la filosofía, desde la ontología y la filosofía, para definir clases y relaciones claras, precisas y funcionales que representen a los seres biológicos y sus propiedades pertinentemente, pues ambas disciplinas, la biología y la filosofía, se enfrentan a la representación de entidades, ideas y eventos, así como a sus propiedades y relaciones, de acuerdo con un sistema de categorización para describir individuos (instancias), clases (conceptos), atributos y relaciones.

A su vez, la biología y la filosofía han tenido que pensar la tecnología para concretar y procesar esa información en lenguaje máquina, a partir del diseño de ontologías, un recurso electrónico que es una especificación explícita de una conceptualización, que consiste en un vocabulario controlado digital que contiene términos organizados jerárquicamente, relacionados lógicamente, acompañados de una definición, sinónimos y un identificador, que se utiliza para organizar, describir y dar significado a la información, comúnmente sistematizada en bases de datos. Una ontología en el lenguaje tecnológico es un recurso que está contextualizado dentro de la denominada "Web Semántica" que plantea el marcado de la información de forma "semántica", es decir, se usan etiquetas que expresan el significado de los elementos y no su formato (Mukherjea, 2005). El objetivo es mejorar el Internet ampliando la interoperabilidad entre los sistemas informáticos y usando programas de agentes inteligentes (algoritmos),

---

<https://orcid.org/0000-0002-5798-662X>, Laboratorio de Bioinformación, Departamento de Biología Comparada, Facultad de Ciencias, Universidad Nacional Autónoma de México. / [laylamichan@ciencias.unam.mx](mailto:laylamichan@ciencias.unam.mx)

<https://orcid.org/0000-0002-1572-164X>, Centro de Investigación en Economía Creativa, Centro de Diseño, Cine y Televisión, Ciudad de México. / [ealvarezl@centro.edu.mx](mailto:ealvarezl@centro.edu.mx)

cuya funcionalidad es buscar información sin operadores humanos. Esta estructuración de datos en la web se basa en la creación de ontologías y lenguajes semánticos con suficiente capacidad expresiva y de razonamiento para representar la semántica utilizando lenguajes como *Resource Description Framework* (RDF), *RDF Schema* y *Web Ontology Language* (OWL) (W3C, 2015).

Se puede describir la ontología de un programa mediante la definición de un conjunto de términos representativos. En una ontología las definiciones asocian los nombres de las entidades en el universo del discurso (por ejemplo clases, relaciones, funciones u otros objetos) con texto legible que describe lo que los nombres deben significar y los axiomas formales que restringen la interpretación y el buen uso de esos términos. Una ontología puede tomar una variedad de formas, pero necesariamente incluirá un vocabulario de términos y alguna especificación de su significado. Gruber (1995) define una ontología como “la especificación de conceptualizaciones, utilizada para ayudar a los programas y a los humanos a compartir conocimientos”.

Si una ontología se refiere a un dominio biológico se llama bio-ontología, y las ciencias de la vida son pioneras en incorporar este recurso electrónico de organización del conocimiento para la recolección y manipulación de datos biológicos y el uso de tales datos para hacer descubrimientos o predicciones biológicas (figura 1) (Stevens y Lord, 2018). Las ontologías proporcionan el marco conceptual en el cual las entidades y los procesos biológicos de trabajo pueden ser estructurados y compartidos en la Web para ser interoperables y normalizados (Hoehndorf, et al., 2015). Los roles desempeñados por las ontologías biológicas se clasifican en tres categorías, que no son mutuamente excluyentes, porque una ontología puede tener uno o varios roles (Bodenreider, 2008):

1. Gestión del conocimiento, incluyendo la indexación y recuperación de datos e información.
2. Integración de datos, intercambio e interoperabilidad semántica.
3. Soporte de decisiones y razonamiento.

La primera bio-ontología, y por mucho la más utilizada, es *Gene Ontology* (GO), un megaproyecto de cobertura mundial que se formalizó en el año 2000 (Ashburner, et al., 2000). Cuenta con los elementos esenciales de cualquier ontología de la Web Semántica como: 1) identificador único y nombre del término, 2) aspecto, 3) definición, 4) relaciones con otros términos, y 5) sinónimos que pueden ser exactos, amplios, estrechos, relacionados y personalizados. Fue diseñada con el objetivo de estandarizar la representación de los genes y los atributos del producto génico entre las distintas especies y las diferentes bases de datos (figura 2). Esta iniciativa proporciona un vocabulario controlado para permitir el uso de definicio-

nes textuales e identificadores, con la finalidad de unificar descripciones con respecto a tres elementos fundamentales de la biología celular y molecular: funciones, procesos y componentes celulares susceptibles a relacionarse con las entidades biológicas representadas en las bases de datos biológicas como literatura, genes, proteínas, especies, etcétera (GeneOntology, 2020). Cabe señalar que el alto rendimiento tecnológico no está limitado únicamente a los genomas, sino que ha influenciado a otras áreas de la biología, como literatura, fenotipos (para determinar características de organismos, frecuentemente resultado de mutaciones específicas), microarreglos (para determinar la expresión de genes), sustancias químicas y fármacos, para mencionar algunas.

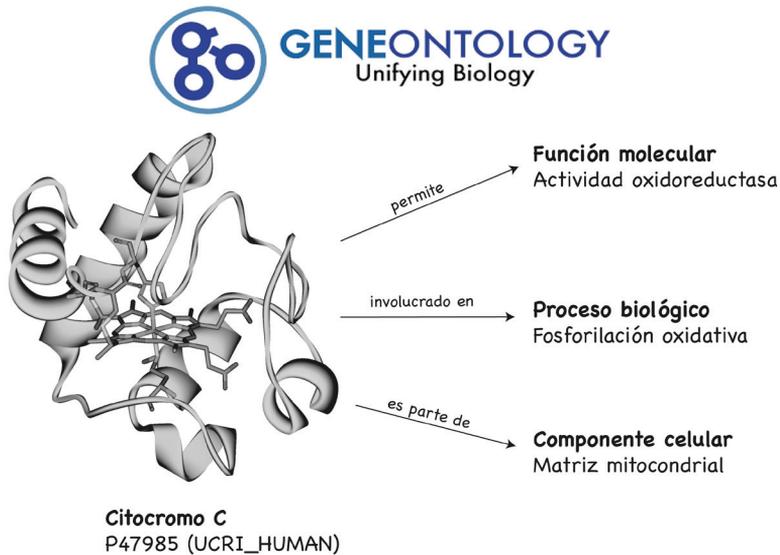


FIGURA 1. Representación ontológica del citocromo C humano. A través de los reportes de literatura y otras bases de datos se pueden predecir y/o anotar algunas funciones biológicas.

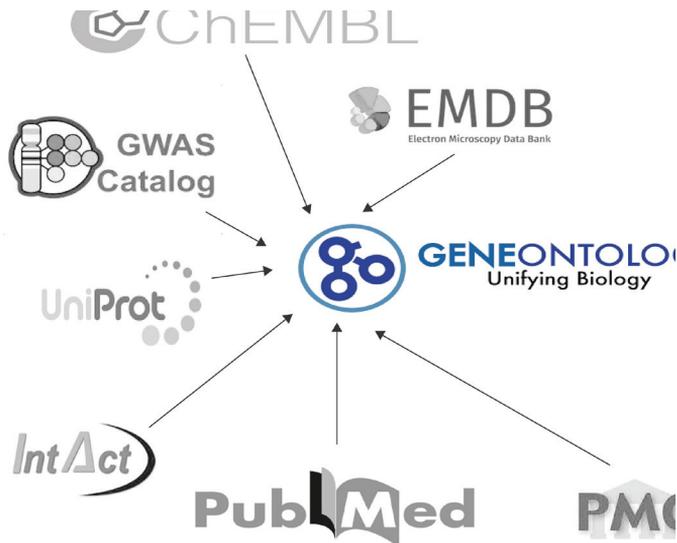


FIGURA 2. Diferentes bases de datos interrelacionadas con *Gene Ontology* para obtención de información, anotación semántica y representación del conocimiento.

De tal forma, en la práctica actual de la biología en la Web es común consultar y utilizar vocabularios acordados por las comunidades de biólogos con una estructura semántica para intercambiar información. La cantidad de ontologías biológicas ha aumentado exponencialmente desde la publicación de *Gene Ontology*, la cifra actual está en los miles y muchas de ellas se pueden consultar en servicios como <https://www.bioontology.org/>, <https://www.ebi.ac.uk/ols/index> y <http://www.obofoundry.org/>.

Este fenómeno de trabajo multi, trans e interdisciplinario que implica y fomenta la participación social necesaria para unificar criterios, acordar estándares, usar lenguajes comunes, intercambiar información y ponerla en acceso abierto es, sin duda, un tema de interés que vale la pena investigar, describir y fomentar en nuestra región.

Por otro lado, la aplicación de las ontologías en una diversidad de ámbitos y para un sinnúmero de aplicaciones tecnológicas de vanguardia en la Web, nos obliga a repensar y considerar la importancia de la comunicación, la

información, el lenguaje, las palabras, el significado, la pragmática y lo cognitivo, ahora desde la neurobiología, y generar así un ciclo permanente entre las distintas aproximaciones que nos permiten entender con más detalle y mejor la realidad de la cual formamos parte, a través del flujo de la información humana y biológica. Nunca la filosofía, la ciencia y la tecnología habían estado tan reunidas y eso, no cabe duda, es producto de la creatividad humana.

## REFERENCIAS

- Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., Harris, M. A., Hill, D. P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J. C., Richardson, J. E., Ringwald, M., Rubin, G. M., & Sherlock, G. (2000), "Gene Ontology: tool for the unification of biology," *Nature Genetics* 25(1): 25–29. <https://doi.org/10.1038/75556>
- Bodenreider, O. (2008), "Biomedical Ontologies in Action: Role in Knowledge Management, Data Integration and Decision Support", *Yearbook of Medical Informatics* 17(1): 67–79. <https://doi.org/10.1055/s-0038-1638585>
- GeneOntology (2020), *GO term elements*. <http://geneontology.org/docs/GO-term-elements>
- Gruber, T. R. (1995), "Toward principles for the design of ontologies used for knowledge sharing?" *International Journal of Human-Computer Studies* 43(5-6): 907–928. <https://doi.org/10.1006/ijhc.1995.1081>
- Hoehndorf, R., Schofield, P. N., & Gkoutos, G. V. (2015), "The role of ontologies in biological and biomedical research: a functional perspective", *Briefings in Bioinformatics* 16(6): 1069–1080. <https://doi.org/10.1093/bib/bbv011>
- Mukherjea, S. (2005), "Information retrieval and knowledge discovery utilising a biomedical Semantic Web", *Briefings in Bioinformatics* 6(3): 252–262. <https://doi.org/10.1093/bib/6.3.252>
- Schulze-Kremer, S., & Smithin, B. (2005), "Ontologies for the life sciences", In Lynn, J., Little, P., Dunn, M., and Subramaniam, S (eds.). *Encyclopedia of Genetics, Genomics, Proteomics and Bioinformatics*. London: John Wiley and Sons. <https://www.wiley.com/en-us/Encyclopedia+of+Genetics%2C+Genomics%2C+Proteomics+and+Bioinformatics%2C+8+Volume+Set-p-9780470849743>
- Stevens, R., & Lord, P. (2018), "Ontologies and Life Science Data Management", in Liu, L., & Özsu, M.T. (eds.), *Encyclopedia of Database Systems*. NY: Springer. [https://doi.org/10.1007/978-1-4614-8265-9\\_631](https://doi.org/10.1007/978-1-4614-8265-9_631)
- W3C (5 de febrero de 2015), *Semantic Web For Life Sciences*. <https://www.w3.org/wiki/SemanticWebForLifeSciences>