

Análisis de relaciones intra-institucionales e interdisciplinarias de una universidad a partir de la producción registrada en *Microsoft Academic*: el caso de la Universidad Nacional de La Plata

José Federico Medrano¹, Sandra Miguel ²,
jfmedrano@fi.unju.edu.ar, smiguel@fahce.unlp.edu.ar

¹VRAIn / Visualización y Recuperación Avanzada de Información / Facultad de Ingeniería
Universidad Nacional de Jujuy (UNJu) - Ítalo Palanca 10, +54 (388) 4221587

² Instituto de Investigaciones en Humanidades y Ciencias Sociales (IdIHCS) (UNLP-
CONICET) / Facultad de Humanidades y Ciencias de la Educación / Universidad Nacional de
La Plata, Calle 51 e/ 124 y 125, (1925) Ensenada, Argentina

RESUMEN

El objetivo del trabajo es conocer las relaciones de colaboración entre investigadores y/o grupos de investigación de una misma institución, separando por áreas del conocimiento, empleando metodologías e indicadores de análisis de redes sociales. El estudio se realizará aplicado al caso de la producción científico-académica de la Universidad Nacional de La Plata indexada en el buscador académico *Microsoft Academic*. Se tomará una muestra del corpus total que corresponde a los registros con año de creación entre 2010-2019. Se analizará la colaboración entre científicos a partir de un estudio descriptivo de las redes de coautoría que existen entre ellos y se realizará un análisis textual de las relaciones entre los temas y ámbitos de investigación. Los resultados permitirán reflejar los vínculos entre investigadores de las mismas o distintas áreas y conocer los niveles de interdisciplinariedad y

relaciones entre unidades de investigación en la producción científico-académica de la institución. Con el trabajo no solo se espera contribuir al conocimiento de las relaciones de colaboración intra-institucionales e interdisciplinarias, sino también mostrar el uso de este tipo de base de datos bibliográfica como fuente de datos para este tipo de estudios métricos y de análisis de redes sociales.

Palabras clave: *Recuperación de Información; Microsoft Academic; Redes de Coautoría; Análisis de Redes Sociales*

CONTEXTO

La línea de investigación aquí presentada se encuentra enmarcada dentro del Proyecto tetra anual (2017-2020) con Código 11/H827, que

lleva por nombre “*La investigación sobre temas locales. Análisis de la producción científica sobre Argentina desde la perspectiva bibliométrica y altmétrica*”, aprobado y financiado por la Secretaría de Ciencia y Técnica de la Universidad Nacional de La Plata (UNLP). Este proyecto es dirigido por la Dra. Sandra Miguel, directora del Departamento de Bibliotecología de la Facultad de Humanidades y Ciencias de la Educación de la UNLP.

1. INTRODUCCIÓN

La colaboración científica es un aspecto de gran relevancia pues se sitúa en la base del desarrollo científico de cualquier disciplina. La colaboración científica es entendida como el desarrollo de actividades conjuntas de investigación entre investigadores, instituciones y países, en las que se comparten recursos intelectuales, económicos, físicos, etc. para la generación y transferencia de conocimientos, productos o servicios (Miguel & Arias, 2012).

En la UNLP la colaboración constituye uno de los ejes estratégicos de sus políticas científicas y tecnológicas. En este sentido, y a través de diferentes programas, promueve la vinculación científica, académica y tecnológica y la participación en redes nacionales e internacionales en las diferentes áreas del conocimiento. En la UNLP más del 80% de la producción científica con visibilidad regional e internacional se publica en colaboración entre autores de la misma institución, o de otras instituciones del país y/o del extranjero. Esto significa una capacidad instalada de cooperación científica de la Universidad, que si bien puede variar entre un campo temático y otro, se va consolidando como un nuevo modo de

producción de conocimientos en la mayoría de las disciplinas (Miguel, 2013).

Según (Moreno, 1997), la colaboración entre grupos tiene un efecto positivo sobre la producción científica, tanto cuantitativa como cualitativa, en especial cuando ésta se lleva a cabo por comunidades científicas y grupos del más alto nivel. Se ha demostrado que las investigaciones que se realizan en equipo son más productivas y adquieren mayor número de citas que cuando se trabaja en forma aislada o con escasa colaboración. Esta dinámica de generación de conocimiento se puede evaluar a través de redes de colaboración, que llegan a ser un reflejo de las relaciones dadas entre grupos, quienes pueden unir instituciones a través de sus vínculos laborales y académicos (Gutiérrez & Velasco, 2017).

Los estudios previos sobre la colaboración institucional han intentado medir la interacción de investigadores e instituciones a partir de las redes de coautoría reflejadas en la producción científico-académica recolectada en distintas bases de datos, al respecto las más empleadas han sido *Scopus*, la *Web of Science*, los Sistemas de Información de Ciencia y Tecnología nacionales, la producción registrada en el CV, en los portales de acceso abierto SciELO y RedALyC, entre otros. Sin embargo, dichos análisis no han sido tenidos en cuenta empleando bases de datos bibliográficas de libre acceso como el motor académico *Microsoft Academic (MA)*¹.

Microsoft Academic Graph (MAG) es un grafo de conocimiento de publicaciones académicas que brinda soporte a MA y está estructurado en torno a los siguientes tipos de entidades: publicación, autor, afiliación del autor

¹ <https://academic.microsoft.com/>

(institución), lugar de publicación (revistas y conferencias), campo de estudio (tema). Contiene fechas de publicación, así como pares de citas y, por supuesto, datos de coautoría. Debido a que MAG es un grafo de conocimiento, facilita poderosos análisis cuantitativos del resultado de la publicación, el impacto, la colaboración y más. Como lo indica (Hug & Brändle, 2017), MAG fue lanzado en Junio de 2015 y modela “*las actividades de comunicación académica de la vida real como un grafo heterogéneo*” (Sinha, y otros, 2015). Los datos para MAG se recogen principalmente de los *feeds* de metadatos de editores y páginas web indexadas por el buscador Bing² (el buscador de propósito general de Microsoft). Este enorme “grafo” puede ser accedido por dos vías, la primera de ellas a partir del motor de búsqueda académico *Microsoft Academic* y la otra por medio de la Academic Knowledge API (AK API)³.

Existen estudios que se encargaron de remarcar la utilidad y capacidad de emplear a MA como herramienta de análisis bibliométrico (Hug & Brändle, 2017; Hug, Ochsner, & Brändle, 2017; Thelwall, 2018; Kousha & Thelwall, 2018). Es en este sentido que MA resulta atractivo a investigadores por la estructuración de los datos que ofrece, la AK API actuando de puente entre el usuario y MAG, facilita enormemente las tareas de extracción y recolección de publicaciones, tarea nada sencilla para su gran rival *Google Scholar*.

Por ello, este trabajo plantea un enfoque novedoso empleando a MA como origen de datos para recolectar toda la producción científico-académica de la UNLP indexada en

esta enorme base de datos bibliográfica de libre acceso. Este conjunto de datos permitirá analizar las redes de coautoría que existen entre los investigadores, o grupos de investigación, del mismo modo se plantea el estudio de las relaciones entre los temas y ámbitos de investigación a partir de un análisis textual de dichas publicaciones. Esto será posible gracias a las entidades que se encuentran representadas en MAG y que permiten ser recolectadas a partir de distintas consultas por medio de la AK API.

2. LÍNEAS DE INVESTIGACIÓN Y DESARROLLO

La presente investigación se enfocará en dos aspectos claves:

- Análisis de relaciones intra-institucionales e interdisciplinarias de una universidad.
- Empleo de bases de datos bibliográficas de libre acceso.

Quizás el primero de ellos no es del todo novedoso, puesto que ya ha sido estudiado desde distintas ópticas y con diferentes enfoques, y en diferentes ámbitos (Muñoz & Delgado, 2016; Miguel & González, 2010; Alcaide, y otros, 2006), sin embargo es el segundo aspecto el que le aporta novedad y relevancia. Este tipo de bases de datos están siendo objeto de estudio por investigadores del campo de la Bibliometría y Ciencias de la Información, por representar una clara alternativa a las bases de datos tradicionales en esta materia, *Scopus* y *Web of Science* (Harzing, 2019; Haunschild, Hug, Brändle, & Bornmann, 2018; Thelwall, 2017).

² <https://www.bing.com/>

³ <https://www.microsoft.com/en-us/research/project/academic-knowledge/>

Por último, el análisis textual que se propone llevar a cabo incluye el empleo de técnicas de Procesamiento del Lenguaje Natural y Aprendizaje Automático para hallar patrones, tendencias y relaciones entre el contenido almacenado.

3. RESULTADOS OBTENIDOS/ESPERADOS

El objetivo de este trabajo es identificar los indicadores de colaboración científica y las redes de coautoría de los investigadores de la UNLP en la última década a partir de un estudio descriptivo de las redes de coautoría que existen entre ellos, empleando el motor académico *Microsoft Academic* como fuente de datos.

Realizando un análisis textual de las publicaciones (campos textuales como título, resumen y campo de estudio) se intentará establecer las relaciones entre los temas y ámbitos de investigación más comunes. Esto permitirá caracterizar desde una visión cuantitativa la producción por campo de estudio o área del conocimiento.

4. FORMACIÓN DE RECURSOS HUMANOS

El Equipo de Trabajo está conformado por docentes investigadores de la UNJu y UNLP. Los mismos llevan adelante esta línea de investigación desde hace años. Al ser un proyecto que involucra a profesionales de distintas instituciones, es un claro ejemplo de colaboración interinstitucional. Además, cada año se incorporan al proyecto alumnos avanzados de distintas carreras, quienes trabajan en temas relacionados. Del mismo modo, los integrantes del equipo participan en el dictado de

asignaturas/cursos de grado y postgrado de la UNLP, UNJu y UCSEDASS.

5. BIBLIOGRAFÍA

Alcaide, G., Zurián, J., Benavent, R., Arroyo, A., Orive, J., & Serrano, S. (2006). Redes de coautoría y colaboración de las instituciones españolas en la producción científica sobre drogodependencias en biomedicina 1999-2004. *Trastornos Adictivos*, 8(2), 78-114. Retrieved from <https://academic.microsoft.com/paper/2011648008>

Gutiérrez, J., & Velasco, N. (2017). Redes de coautoría como herramienta de evaluación de la producción científica de los grupos de investigación. *Revista General de Información y Documentación*, 27(2), 279-297. Retrieved from <https://academic.microsoft.com/paper/2771760668>

Harzing, A.-W. (2019). Two new kids on the block: How do Crossref and Dimensions compare with Google Scholar, Microsoft Academic, Scopus and the Web of Science? *Scientometrics*, 120(1), 341-349. Retrieved from <https://academic.microsoft.com/paper/2944642116>

Haunschild, R., Hug, S., Brändle, M., & Bornmann, L. (2018). The number of linked references of publications in Microsoft Academic in comparison with the Web of Science. *Scientometrics*, 114(1), 367-370. Retrieved from <https://academic.microsoft.com/paper/2763175323>

- Hug, S., & Brändle, M. (2017). The coverage of Microsoft Academic: analyzing the publication output of a university. *Scientometrics*, 113(3), 1551-1571. Retrieved from <https://academic.microsoft.com/paper/2605124925>
- Hug, S., Ochsner, M., & Brändle, M. (2017). Citation analysis with microsoft academic. *Scientometrics*, 111(1), 371-378. Retrieved from <https://academic.microsoft.com/paper/2522876520>
- Kousha, K., & Thelwall, M. (2018). Can Microsoft Academic help to assess the citation impact of academic books. *Journal of Informetrics*, 12(3), 972-984. Retrieved from <https://academic.microsoft.com/paper/2886862840>
- Miguel, S. (2013). *Serie Indicadores Bibliométricos UNLP: Informe 2013*. La Plata: Universidad Nacional de La Plata.
- Miguel, S., & Arias, R. (2012). La colaboración científica en la UNLP. Retrieved from <https://academic.microsoft.com/paper/798563433>
- Miguel, S., & González, C. (2010). Composición y dinámica de los grupos de investigación del Departamento de Bibliotecología de la Universidad Nacional de La Plata en el periodo 2000-2009. *I Jornada de Intercambio y Reflexión acerca de la Investigación en Bibliotecología*. Retrieved from <https://academic.microsoft.com/paper/409268175>
- Moreno, C. (1997). Técnicas bibliométricas aplicadas a los estudios de usuarios. *Revista General de Información y Documentación*, 7(2), 41-68. Retrieved from <https://academic.microsoft.com/paper/1597984441>
- Muñoz, E., & Delgado, Y. (2016). Detección de comunidades a partir de redes de coautoría en grafos RDF. *ACIMED*, 27(1), 90-99. Retrieved from <https://academic.microsoft.com/paper/2256194420>
- Sinha, A., Shen, Z., Song, Y., Ma, H., Eide, D., Hsu, B.-J., & Wang, K. (2015). An Overview of Microsoft Academic Service (MAS) and Applications. *Proceedings of the 24th International Conference on World Wide Web*, (pp. 243-246). Retrieved from <https://academic.microsoft.com/paper/1932742904>
- Thelwall, M. (2017). Microsoft Academic: A multidisciplinary comparison of citation counts with Scopus and Mendeley for 29 journals. *Journal of Informetrics*, 11(4), 1201-1212. Retrieved from <https://academic.microsoft.com/paper/2766265189>
- Thelwall, M. (2018). Microsoft Academic automatic document searches: Accuracy for journal articles and suitability for citation analysis. *Journal of Informetrics*, 12(1), 1-9. Retrieved from <https://academic.microsoft.com/paper/2768506257>