

## Análisis cuantitativo de la producción en investigación científica y tecnológica

Roberto M. Muñoz, Analía Guzmán, Martín M. Casatti,  
Calixto Maldonado, Juan C. Cuevas, Luis E. Damiano

Departamento de Ingeniería en Sistemas de Información  
CIDS – Centro de Investigación Desarrollo y Transferencia  
de Sistemas de Información  
Facultad Regional Córdoba – Universidad Tecnológica Nacional  
Maestro Marcelo López esq. Cruz Roja Argentina – Córdoba  
0351 – 4686385

[rmunoz@frc.utn.edu.ar](mailto:rmunoz@frc.utn.edu.ar), [aguzman@frc.utn.edu.ar](mailto:aguzman@frc.utn.edu.ar), [mcasatti@frc.utn.edu.ar](mailto:mcasatti@frc.utn.edu.ar),  
[cmaldonado@frc.utn.edu.ar](mailto:cmaldonado@frc.utn.edu.ar), [juancarloscue@gmail.com](mailto:juancarloscue@gmail.com), [luis.damiano@gmail.com](mailto:luis.damiano@gmail.com)

### Resumen

El proyecto caracterizará la producción científica y tecnológica desarrollada por las universidades de la República Argentina, por medio de la elaboración de una metodología de análisis cuantitativo a partir de la documentación producida por los investigadores, becarios y centros de investigación. Se realizará un análisis de la información contenida en la documentación y se desarrollarán técnicas y herramientas para la extracción de la misma y el almacenamiento en un medio, que posteriormente permita el correcto análisis cuantitativo. Parte de este desarrollo incluye una plataforma web para la gestión administrativa y académica de congresos relacionados a tecnología informática y sistemas de información en una primera etapa, obteniendo así la información de análisis en tiempo real a partir de los artículos enviados por los autores a los congresos.

**Palabras clave:** *cuantimetría – redes de colaboración – producción científica – métricas - patrones.*

### Contexto

El presente trabajo forma parte del proyecto de investigación y desarrollo que ha sido homologado por la Secretaría de Investigación, Desarrollo y Posgrado de la Universidad Tecnológica Nacional (UTN), reconocido con el código: SIUTNCO0007848. El mismo se desarrolla en el Centro de Investigaciones, Desarrollo y Transferencia de Sistemas de Información – CIDS, de la Facultad Regional Córdoba de UTN.

### 1. Introducción

La cuantimetría (scientometrics) como disciplina se dedica al “estudio de los aspectos cuantitativos de la comunicación científica, las prácticas de investigación y desarrollo, así como las políticas de ciencia y tecnología” [1]. Con la aparición de big data se logra que grandes conjuntos de datos puedan ser tratados y utilizados con fines analíticos y de

predicción, lo que hasta hace poco era imposible de procesar mediante el uso de herramientas tradicionales debido a su tamaño y complejidad. La combinación de “big data”, “data mining”, bibliometría y redes sociales, ha dado lugar a lo que se ha denominado como science mapping [2], cuyo objetivo es visualizar la estructura y las relaciones de la ciencia en general, y de las áreas científicas, disciplinas, líneas de investigación, autorías, etc., en particular.

Una cita es una referencia a un trabajo, ya sea propio o de otros autores o grupos, que se ha tomado total o parcialmente para construir el propio trabajo de investigación. En cierta medida el tipo y cantidad de citas que un trabajo de investigación registra es un indicador, con mayor o menor grado de fidelidad, del impacto que dicho trabajo produce en la comunidad científica. Un trabajo muy citado es la base para diversos papers científicos y líneas de investigación, mientras que un trabajo raramente citado no produce un efecto importante en la comunidad.

Otro análisis importante que se puede realizar por medio del análisis de citas es el de los patrones de colaboración entre instituciones o entre investigadores. Las redes de coautoría son una importante clase de redes sociales. El análisis de estas redes revela las características de las comunidades académicas que ayudan en el entendimiento de los trabajos científicos colaborativos y en la identificación de los investigadores prominentes [3]. Es importante destacar que el análisis de redes de coautoría puede ayudar a caracterizar una cierta población de investigadores, determinando cuales son las temáticas de interés y la producción total y relativa por área temática. Si se considera que las relaciones entre los investigadores se puede representar como un grafo no dirigido existe una métrica que es de

especial importancia a la hora de caracterizar dicho grafo. Dicha métrica es el grado de centralidad. La centralidad puede indicar, en un grafo que represente coautoría de trabajos de investigación, aquellos autores que tienen un grado alto de colaboración con otros autores. Por otra parte, una métrica asociada, la centralidad de intermediación, puede indicar aquellos autores que ofician de nexo entre otros investigadores o grupos de investigación.

Por otra parte, se pueden utilizar las redes de coautoría para evaluar la interrelación entre las instituciones académicas, el sector gubernamental y el sector productivo [7]. De esta manera es posible identificar el impacto generado por esa interrelación y se pueden diseñar políticas tendientes a reforzar esos vínculos para generar resultados beneficiosos a mediano y largo plazo.

Un estudio de extrema importancia, sobre todo en economías en donde la investigación científica y tecnológica no es prioritaria en las políticas de gobierno, es el análisis que establece relaciones entre la producción científica y tecnológica y el ámbito de los investigadores sobre todo desde el punto de vista socioeconómico [1] [4]. Desde este punto de vista también se puede analizar el surgimiento o decadencia de ciertas líneas de investigación.

Todas estas prácticas entran dentro de las incumbencias del denominado mapeo científico o en su nombre original “science mapping” [5].

“El objetivo central de “science mapping” es el de revelar la estructura y dinámica del conocimiento científico ...”, tal como menciona el trabajo de Chen, Dubin y Schultz [9]. Los análisis cuantitativos sirven como punto de partida para contar con información

cuantitativa, heterogénea y relacionada para poder plantear análisis de science mapping [8].

Hay que mencionar que, a la hora de optimizar el uso de recursos escasos, como pueden ser los lugares físicos para laboratorios, las plazas en las plantas de investigadores permanentes, los recursos económicos, de materiales o de insumos, es fundamental contar con un diagnóstico adecuado de la situación actual del ámbito en donde esos recursos se van a utilizar. En este tema en particular el análisis cuantitativo puede servir para tener una imagen actualizada y confiable de la infraestructura de I+D+i a los fines de asistir en la toma de decisiones [6].

## 2. Líneas de Investigación, Desarrollo e Innovación

En el presente trabajo se explorarán las siguientes líneas de investigación, las que pueden ser adaptadas o adecuadas de acuerdo a los hallazgos producto del propio trabajo de investigación:

- Estudio de la representación de información vinculada a la producción científica y redes de colaboración entre investigadores y centros de investigación.
- Revisión y/o implementación de mecanismos de extracción de información de mapeo a partir de documentos de difusión de resultados de investigación (papers).
- Estudio de características comunes en diversos tipos de instrumentos y medios de difusión de resultados de investigación y obtención de atributos que definen la temática, metodología y resultados obtenidos en dichos trabajos.
- Caracterización de los participantes de redes de colaboración, ya sea desde el

punto de vista personal o institucional y estudio de la representación de dichas características a fines de su consulta.

- Estudio de las bases de datos de grafo como un mecanismo idóneo para la representación de la información necesaria para un “science mapping” exitoso.
- Estudio de los mecanismos de consulta y visualización de la información, teniendo en cuenta la heterogeneidad y el posible gran tamaño de los resultados de las consultas.

## 3. Resultados esperados

El objetivo general del presente estudio es caracterizar la producción e impacto de las líneas de investigación que llevan adelante las universidades argentinas, por medio de una herramienta desarrollada ad-hoc.

Para ello se han identificado los siguientes objetivos particulares, para obtener los resultados esperados:

1. Modelar la representación de la información asociada a las tareas de I+D, que permita caracterizar todos los trabajos producidos por las universidades argentinas.
2. Estructurar un sistema de almacenamiento cuantitativo que permita la registración y consulta de la información registrada.
3. Procesar los documentos de los congresos, para extraer los atributos para el análisis cuantitativo e inferir lo que no se obtiene directamente de los textos.
4. Diseñar, desarrollar e implementar un sistema de gestión de congresos y conferencias para registrar automáticamente la información en el sistema de almacenamiento, a partir de los papers y artículos presentados, sin necesidad de interacción adicional o carga manual de datos.
5. Desarrollar una herramienta de consulta y análisis que permita obtener métricas asociadas a la producción científica, que posibilite

analizar los diversos patrones que pudieran surgir a partir del análisis de citas, autores y líneas de investigación.

Como se puede apreciar, los objetivos del proyecto se agrupan en dos enfoques muy particulares:

- El análisis y la caracterización de la investigación producida por las universidades argentinas, en base a información preexistente.
- La construcción de una herramienta de software que permita realizar dichos análisis sobre futuros trabajos y líneas de investigación.

#### 4. Formación de Recursos Humanos

El equipo de investigación está formado por: 6 docentes, de diferentes asignaturas de la carrera de Ingeniería en Sistemas de Información, 1 graduado y 3 estudiantes.

Los docentes investigadores explorarán los usos de una metodología, que les permitirá ampliar sus posibilidades, en cuanto a líneas de investigación de interés, tanto local como regional o internacional, y accederán a nombres y trabajos de posibles colaboradores en sus respectivos campos de estudio. De los docentes integrantes del equipo, dos de ellos realizarán su tesis en el marco del presente proyecto, la cual les permitirá finalizar los estudios de posgrado en la Maestría en Ingeniería en Sistemas de Información.

Dos estudiantes de grado, que aún cursan asignaturas y no poseen experiencias en investigación, podrán realizar sus primeras experiencias en este proyecto, y uno, que ya posee experiencia en proyectos, podrá ampliar los conocimientos adquiridos durante el cursado, en el desarrollo de las herramientas informáticas que se describen en los objetivos. Los 3 estudiantes podrán proponer la temática en sus Prácticas Supervisadas, contando con el

asesoramiento y seguimiento de los docentes integrantes del PID como Docentes Tutores.

La inclusión del reciente graduado se produce por su experiencia en anteriores Proyectos de Investigación académico y en el Laboratorio de la Carrera, como también por su rendimiento académico.

#### 5. Referencias

[1] RENDÓN, ESMERALDA CERVANTES AND VICTORIANO GARZA-ALMANZA (2015). "La cienciometría como herramienta para analizar el impacto de la investigación científica en una región." *Cultura Científica y Tecnológica* 48.

[2] M.J. COBO, A.G. LÓPEZ-HERRERA, E. HERRERA-VIEDMA AND F. HERRERA (2011). "Science mapping software tools: Review, analysis, and cooperative study among tools". *Journal of the American Society for Information Science and Technology*, vol. 62, nº 7, pp. 1382-1402.

[3] UMADEVI, V. (2013). "Case study–centrality measure analysis on co-authorship network" *Journal of Global Research in Computer Science*, 4(1), 67-70.

[4] LEYDESDORFF, LOET, AND ISMAEL RAFOLS (2011). "Local emergence and global diffusion of research technologies: An exploration of patterns of network formation." *Journal of the American Society for Information Science and Technology* 62.5: 846-860.

[5] EBENER, STEEVE, et al. (2006). "Knowledge mapping as a technique to support knowledge translation." *Bulletin of the World Health Organization* 84: 636-642.

[6] MCGRATH, W. (1989). "What bibliometricians, scientometricians and informetricians study; a typology for definition and classification; topics for discussion" *International Conference on Bibliometrics*,

Scientometrics and Informetrics (Vol. 2).  
Ontario: The University of Western Ontario.

[7] SÁNCHEZ, M. A., SCHMIDT, M. A., &  
OBIOL, L. C. (2018). “Redes de coautorías y  
patrones de colaboración institucional: análisis  
de un caso en el campo de la ingeniería”.

[8] ROBLES, E., & VINCK, D. (2010).  
“Mapeo de la evolución de la producción de  
ciencias y tecnologías emergentes:  
comparación de estrategias lexicales sobre las  
nanociencias y nanotecnologías”. In ESOCITE  
2010.

[9] CHEN, C., DUBIN, R., & SCHULTZ, T.  
(2015). “Science mapping”. In Encyclopedia of  
Information Science and Technology, Third  
Edition (pp. 4171-4184). IGI Global.