

## CIENCIOMETRÍA EN EL INSTITUTO DE INGENIERÍA EUGENIO LÓPEZ ORTEGA Y ERIC ULISES GARCÍA CANO

En 2017 el Instituto de Ingeniería de la UNAM (IIUNAM) inició el proyecto titulado Cienciometría en el IIUNAM dirigido a analizar el trabajo de la institución a través de los documentos publicados por sus académicos.

De acuerdo a Leydesdorff (2001), la cienciometría estudia la producción científica con el fin de medir y analizar su desarrollo. Para este autor, el estudio actual del desarrollo científico se sustenta en tres dimensiones como se muestra en la figura 1.

En la figura 1, la cienciometría se ubica en la observación de los textos generados por los científicos (académicos) en las diferentes áreas de la ciencia y la tecnología.

Es posible distinguir diversas aplicaciones recientes de la cienciometría; por ejemplo:

- Evaluación de fuentes (revistas y congresos), de académicos y de las instituciones a las que pertenecen.

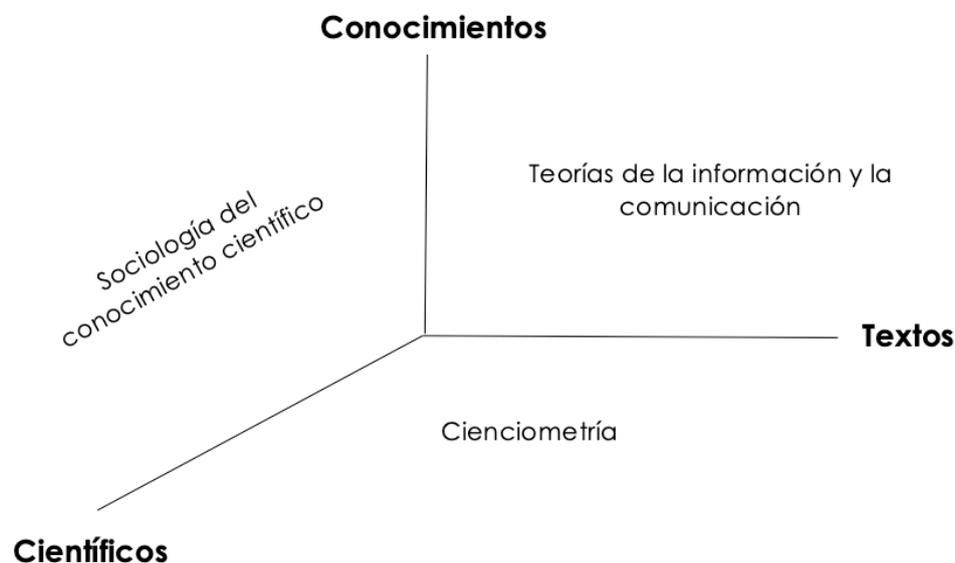
- Análisis del desarrollo de temas científicos y tecnológicos para identificar tendencias, subtemas y/o líneas de investigación emergentes.

- Diagnóstico de la situación actual de un tema de investigación de una institución dedicada al desarrollo científico y/o tecnológico, etc.

La cienciometría ha tenido un importante crecimiento en los últimos años motivado por el desarrollo de diversas bases de datos académicas que contienen información de la producción científica y tecnológica publicada en diversos medios (revistas, memorias de congresos, tesis de posgrado, etc.). Estas fuentes pueden ser de acceso restringido como Scopus y Web of Science (WoS) y otras de acceso abierto como Google Scholar y Crossref, entre otras.

### Proyecto Cienciometría en el IIUNAM

Este proyecto tiene como antecedente un trabajo con información contenida en la base de datos Scopus y está orientado a estudiar la producción científica y tecnológica de temas específicos de investigación a partir de la información contenida en revistas



Fuente: Tomado de Leydesdorff (2001)

Figura 1. Dimensiones para el análisis del desarrollo científico y tecnológico

que, de acuerdo con los académicos del IIUNAM, corresponden a las más relevantes en la difusión de nuevos conocimientos en los temas analizados (López Ortega, 2017).

Por su parte, el proyecto Ciencimetría en el IIUNAM está dirigido a analizar el impacto de la institución a través de los documentos publicados por los académicos e identificados en la base de datos Google Scholar (GS), también conocida como Google académico.

La base de datos GS es de libre acceso y posee mayor cobertura en comparación con ISI (WoS) y Scopus. Incluye todas las revistas académicas listadas en ISI y Scopus, incluso aquellas que no están listadas en estas bases de datos pero que están disponibles en internet. GS también incluye información de libros, memorias en congresos, documentos técnicos y reportes oficiales (Harzing, 2013).

Por esta razón, la información que ofrece GS representa mejor el impacto de un trabajo que aporta conocimientos relevantes, independientemente de la fuente de información en la que haya sido publicado.

El proyecto Ciencimetría en el IIUNAM consta de dos vertientes. Por una parte, la Unidad de Servicios de Información del IIUNAM (USI) emprendió un trabajo de promoción y facilitación de apertura y/o mantenimiento de perfiles de los académicos del Instituto en GS. Por otra parte, se analizó la manera de importar la información contenida en GS con el fin de desarrollar una base de datos que contuviese la información de los académicos del IIUNAM.

Con respecto a la primera vertiente del proyecto, a la fecha se cuenta con 137 académicos del IIUNAM con perfiles en GS. En la segunda vertiente se ha logrado importar la información de GS correspondiente a dichos académicos. A continuación se presenta una reseña del desarrollo de esta segunda vertiente del proyecto.

### **Desarrollo de una base de datos con información de GS y Crossref**

La base de datos GS tiene tres niveles de información de acuerdo con el detalle que presenta.

El primer nivel consiste en la identificación de autores (perfiles) que cuentan con una característica común. Por ejemplo, que estén asociados a una institución en particular; en nuestro caso, al Instituto de Ingeniería de la UNAM. Con base en esta selección, GS identifica a los perfiles de académicos que cumplen con dicha característica.

Al seleccionar cada perfil se despliega la lista de trabajos que GS ha identificado como de la autoría o coautoría del académico seleccionado. Este conjunto de información corresponde al segundo nivel.

El tercer nivel de información se despliega al seleccionar cada uno de los trabajos contenidos en la lista anterior. Al ingresar a este nivel, GS muestra información detallada del trabajo seleccionado, así como las veces que ha sido citado en el orden del año de publicación de las citas. En este tercer nivel, GS también señala el número de citas alcanzado por el trabajo seleccionado de acuerdo con la información contenida en otra base de datos académica (*Web of Science*). Esta información se muestra sólo en los casos en los que el trabajo seleccionado se encuentra indizado en esta última base de datos académica.

Con el fin de importar la información de los académicos del IIUNAM, se desarrolló un sistema de cómputo que permite ingresar a los tres niveles de información contenidos en GS. El sistema consta de tres partes:

a. Subsistema Importador (SI). Corresponde a una rutina de cómputo que accede a GS y solicita la información de los niveles dos y tres de los perfiles (académicos) que cumplen con alguna característica común establecida en el primer nivel de información; en este caso, con la pertenencia al IIUNAM.

Este subsistema tiene dos características a destacar. La primera consiste en controlar el tiempo de solicitudes de información a GS con el fin de evitar que el servidor de GS bloquee al SI al considerarlo un robot que importa los datos.

La segunda característica consiste en que, una vez importados los datos de un documento, el sistema accede a otra base de datos pública (Crossref) con el fin de comprobar y/o completar la información. En este proceso se agrega información faltante en la base de datos GS y se asigna el código identificador de objeto digital, conocido en inglés como (DOI) que corresponde al trabajo en cuestión.

b. Sistema Cargador (SC). Una vez que se ha descargado la información desde GS y verificada en Crossref, el SC la sube a una base de datos. En este proceso se depura parte de la información obtenida de GS. Por ejemplo, es frecuente que el nombre de las fuentes (revistas, congresos, etc.) en las que se publicaron los documentos aparezcan de diferente manera; por ejemplo, la revista *Water research and technology*, puede aparecer también como *Water res & tech*, entre otras formas de citarla. Con el fin de homogeneizar las fuentes se desarrolló un catálogo que identifica las diversas formas en que aparece el nombre de la fuente. De esta manera, al subir la

información a la base de datos, el SC previamente homogeniza el nombre de la fuente.

c. Base de datos *Crosscholar*. Corresponde a la base de datos que se ha construido con la información importada y depurada de GS y complementada con la información de Crossref. Provisionalmente se le ha llamado así debido a que integra datos contenidos en estas dos fuentes de información.

Los principales datos que integran esta base son:

- Título del trabajo.
- Autores y sus instituciones de adscripción.
- Tipo de publicación: artículo en revista, en memoria de congreso, libro, otro tipo de publicación.
- Datos de la fuente de publicación: nombre de la revista o congreso.
- Número de citas recibidas según el año.

La siguiente figura presenta un esquema del funcionamiento del sistema desarrollado.

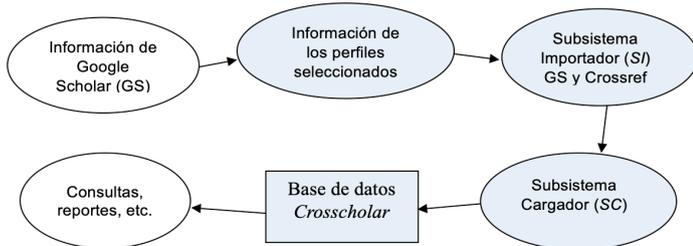


Figura 2. Esquema de funcionamiento del sistema desarrollado

## Resultados alcanzados

A la fecha se cuenta con 137 perfiles en GS de sendos académicos del IIUNAM, de los que se identificaron 6,949 trabajos de su autoría o coautoría. Estos trabajos han recibido un total de 74,683 citas, por lo que en promedio GS identifica 10.7 citas por trabajo.

Las siguientes figuras muestran algunos de los reportes que genera la base de datos *Crosscholar* desarrollada en el marco del proyecto *Cienciometría en el IIUNAM*.

La figura 3 muestra el número de trabajos publicados de la autoría o coautoría de los académicos del Instituto de Ingeniería de la UNAM durante el periodo 2010-2018. Asimismo, presenta el número de citas recibidas durante el mismo periodo. Se observa una tendencia creciente en dicho número de citas recibidas anualmente, con excepción de 2018, lo que seguramente obedece a información aún incompleta de dichas citas ya que los datos mostrados corresponden al 14 de diciembre de 2018.

La figura 4 presenta las principales fuentes (revistas, congresos, etc.) en las que publican los académicos del IIUNAM. Se muestra la cantidad de trabajos publicados en cada fuente, así como el número de citas recibidas.

La columna señalada como citas institucionales representa la parte que corresponde a los académicos del IIUNAM, de acuerdo con el número de autores y coautores en cada trabajo. Es decir, si en un trabajo publicado participaron cinco coautores y de ellos solamente dos son académicos del IIUNAM,



Figura 3. Número de trabajo y citas recibidas por los trabajos del IIUNAM (2010-2018)

## Fuentes

Fuente: scholar.google.com

14/12/18

Número de citas que corresponden al IIUNAM de acuerdo al número de académicos participantes en los trabajos

#	Título	Clasificación	JCR	Trabajos	Citas	Citas institucionales	País
1	<a href="#">WATER SCIENCE AND TECHNOLOGY</a>	Journal	<input checked="" type="checkbox"/>	138	3,106	1,399	
2	<a href="#">BULLETIN OF THE SEISMOLOGICAL SOCIETY OF AMERICA</a>	Journal	<input checked="" type="checkbox"/>	121	7,273	4,046	
3	<a href="#">AGU FALL MEETING</a>	Congress	<input type="checkbox"/>	117	18	5	
4	<a href="#">REVISTA DE INGENIERÍA SÍSMICA</a>	Journal	<input type="checkbox"/>	92	312	169	
5	<a href="#">CONGRESO INTERAMERICANO DE INGENIERÍA SANITARIA Y AMBIENTAL</a>	Congress	<input type="checkbox"/>	65	70	32	
#	Título	Clasificación	JCR	Trabajos	Citas	Citas institucionales	País

Showing 1 to 10 of 2,432 entries

Figura 4. Principales fuentes en las que se publican trabajos del IIUNAM

entonces del total de citas recibidas solamente le corresponde 40% a la institución.

La figura 5 presenta una comparación entre las citas reportadas por GS y las que señala otra base de datos académica como lo es Web of Science (WoS). La barra superior muestra el total de citas reportadas por GS de los trabajos asociados al IIUNAM. La segunda barra señala las citas que reporta GS de los trabajos en los que también WoS reporta citas. Es decir, de aquellos trabajos publicados en las fuentes contenidas en WoS. La tercera barra representa el número de citas reportadas por WoS de los mismos trabajos a los que se refiere la segunda barra. Por consecuencia, GS identifica 1.6 veces más citas de trabajos de los que reporta WoS de acuerdo a la información contenida en la base Crosscholar.

### Conclusiones

Con base en los resultados logrados, el proyecto Cienciometría en el IIUNAM sí arroja información relevante en cuanto a la

producción institucional. La base de datos académica Google Scholar, de la cual se importa la información, ofrece una amplia cobertura de los trabajos producidos por los académicos del IIUNAM, así como el impacto de ellos reflejado a través de las citas recibidas.

Se identifican las principales fuentes en las que los académicos de la institución publican sus trabajos, más allá de las reportadas por otras bases de datos académicas tales como WoS y Scopus.

En el caso de las citas, si se consideran los mismos trabajos identificados tanto en GS como en WoS, el número reportado resulta 1.6 veces superior en GS. Esta diferencia aumenta al considerar todos los trabajos producidos por los académicos del IIUNAM y que no necesariamente están identificados por WoS pero sí por GS. Esta proporción se eleva a casi 2.8 veces el número de citas a los trabajos asociados al IIUNAM.

El proyecto Cienciometría en el IIUNAM continúa su desarrollo con diversas acciones tales como la construcción de indicadores que expresen diversas características de la producción

## Google Scholar vs. WOS (2000 - 2018)

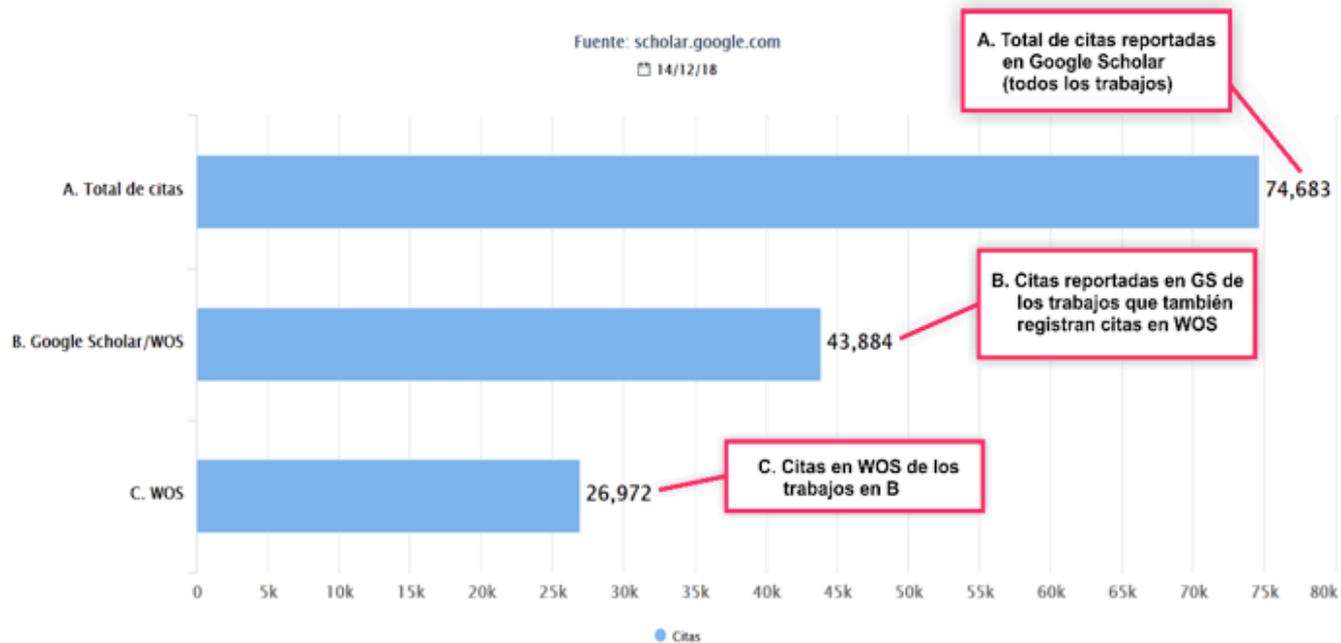


Figura 5. Número de citas de los trabajos del IIUNAM reportadas por GS y WoS

y publicación del trabajo de la institución, la creación de nuevos reportes con la información contenida en la base de datos, la consulta a través de internet por parte de los académicos y la consolidación del proceso de importación de información, entre otras acciones.

En el desarrollo del proyecto participaron los estudiantes Nadia Castillo Camarena y José Alfonso Domínguez Baleón.

### Referencias

Harzing, A. W. (2013). The publish or perish book. Tarma Software Research Pty Ltd, Melbourne, Australia. Recuperado el 7 de enero en: <https://harzing.com/popbook/ch1.htm>.

Leydesdorff, L. (2001). The challenge of scientometrics. The development, measurement and self-organization of scientific communication. 2<sup>nd</sup> edition. Universal publishers. USA.

López Ortega, E. (2017). Sistema de Cómputo para Inteligencia Tecnológica (SCIT). Gaceta del Instituto de Ingeniería, enero-febrero 2017, No. 123 (pp. 7-9).

