

Metodología para la Generación de Grafos de Colaboración Científica de una Institución Académica

Methodology for the Generation of Scientific Collaboration Graphs of an Academic Institution

Jared David Tadeo Guerrero Sosa¹, Víctor Hugo Menéndez Domínguez*¹, María Enriqueta Castellanos Bolaños¹, César Arturo Guerra García²

¹Facultad de Matemáticas, Universidad Autónoma de Yucatán

Periférico Norte Tablaje 13615 C.P. 97119, Mérida, Yucatán, México

²Universidad Autónoma de San Luis Potosí

Carretera Salinas - Santo Domingo 200 Salinas de Hidalgo C.P. 78600 San Luis Potosí, México

jared.guerrero@correo.uady.mx, mdoming@correo.uady.mx,
enriqueta.c@correo.uady.mx, cesar.guerra@uaslp.mx

*Autor para correspondencia

Abstract

Scientific collaborations have widely contributed to the scientific community. Several evaluating institutions consider this factor one of the most important for the allocation of resources and acknowledgments to researchers and research groups. This work proposes a methodology for the extraction of data associated with the scientific publications by an academic institution from a high-impact bibliographic database, using elements from the Educational Data Mining to generate collaboration graphs considering the possible multiplicity of profiles associated with a specific author. To validate the proposal, the internal and international external collaborations by the Universidad Autónoma de Yucatán were analyzed. The 109 most productive and collaborative professors distributed in 15 groups and the 25 most collaborative countries with the studied institution were identified.

Resumen

Las colaboraciones científicas han presentado una gran aportación a la comunidad científica. Diversas instituciones evaluadoras consideran este factor uno de los principales para la asignación de recursos y reconocimientos a investigadores y grupos de investigación. El presente trabajo propone una metodología que permite la extracción de los datos de las publicaciones científicas registradas en una base de datos bibliográfica de alto impacto generadas por una institución académica, utilizando elementos de la minería de datos educativos, con el propósito de generar grafos de las colaboraciones, considerando la posible multiplicidad de perfiles asociados a un mismo

Fecha de recepción: septiembre 1, 2020 / Fecha de aceptación: octubre 5, 2020

autor. Con la finalidad de validar la propuesta, se analizó la colaboración interna y externa internacional de la Universidad Autónoma de Yucatán. Se identificaron a los 109 profesores con mayor productividad y colaboración interna distribuidos en 15 grupos, además de los 25 países con mayor colaboración con la institución estudiada.

Keywords and phrases: Colaboración científica, Grafos de colaboración, Scopus, VOSviewer.

2020 Mathematics Subject Classification: 68R10, 97K30.

1 Introducción

Una práctica común en la comunidad científica es la colaboración entre dos o más individuos con la finalidad de generar conocimiento, diseminado a través de publicaciones con enfoques científicos, educativos, tecnológicos y de innovación.

Los motivos de las colaboraciones científicas son diversos. A nivel personal, un investigador colabora con la finalidad de obtener mayor exposición y progreso en la comunidad científica, resolver problemas grandes cuya solución abarca diversas líneas de investigación, además de recibir estímulos por una o más instancias evaluadoras debido a su labor [1].

A nivel institucional, existe evidencia de que, a mayor colaboración, mayor productividad, además de que la definición de grupos de investigación genera producción de mayor impacto (registrada en bases de datos de publicaciones indizadas), y por lo tanto, el índice de citación incrementa considerablemente [2].

Sin embargo, también se ha demostrado que incluso las colaboraciones persistentes y homogéneas no siempre culminan en la publicación de artículos relevantes, mientras que los grupos de colaboración transdisciplinarios han generado publicaciones de mayor aportación para la comunidad científica [3].

Se considera una publicación de alto impacto (o indizada) a toda aquella cuyo registro es localizable en una base de datos bibliográfica (o repositorio) cuyas publicaciones cumplen con una serie de indicadores de calidad [4], tales como el impacto de la investigación, la relevancia del origen de las citas, la relevancia de la revista, entre otros. Scopus y Web of Science se consideran las dos bases de datos bibliográficas más importantes a nivel mundial.

Tanto Scopus como Web of Science ponen a disposición los datos referentes a sus publicaciones almacenadas a través de diversas APIs. Algunas pertenecientes a Scopus y mencionadas en este trabajo permiten recuperar datos de afiliaciones [5], autores [6] y productos [7].

Incluso diversas instancias evaluadoras se encargan de otorgar reconocimientos y recursos evaluando la producción científica de un investigador incluida en Scopus y Web of Science. En

México, el Consejo Nacional de Ciencia y Tecnología (CONACYT), a través del Sistema Nacional de Investigadores (SNI) reconoce el nivel de los investigadores de acuerdo con su producción perteneciente a una de sus siete áreas de conocimiento [8].

Otra institución evaluadora es la Secretaría de Educación Pública (SEP) por medio del Programa para el Desarrollo Profesional Docente para el Tipo Superior (PRODEP). El PRODEP evalúa la productividad de los profesores de tiempo completo, así como de los cuerpos académicos (grupos de profesores de tiempo completo con una línea de investigación y un grado de consolidación otorgado de acuerdo con, entre otros factores, la relevancia de su producción). Aunque los productos científicos no son los únicos elementos que se evalúan para los reconocimientos, el PRODEP considera que son la mejor evidencia de un trabajo relevante con aportación científica, tecnológica y de innovación [9].

Para la realización de este trabajo se ha propuesto un modelo ontológico para la relevancia científica [10]; un análisis de la producción entre cuerpos académicos [11]; una propuesta de un sistema de recomendación para la relevancia científica, que considera, entre varios aspectos, la colaboración interna y externa [12]; así como el uso de una base de datos documental para el registro de la producción científica de alto impacto y disponible en repositorios de acceso abierto [13].

Este trabajo tiene como finalidad presentar una metodología que permita obtener, a través de la minería de datos educativos, los datos de la producción científica de una institución a partir de Scopus, para un posterior análisis del estado de la investigación, identificando los grupos de colaboración interna, así como localizar los países colaboradores con la institución en cuestión.

El resto del artículo se organiza de la siguiente manera: en el estado del arte se presentan los conceptos de la teoría de grafos y su uso en algunos estudios sobre la colaboración científica y aspectos de la colaboración internacional. En la metodología se describe una serie de pasos para la extracción de los datos pertenecientes a la producción indexada por Scopus, asociada a los miembros de una institución, para después exponer su implementación con la producción de los profesores de tiempo completo de la Universidad Autónoma de Yucatán. Finalmente, se indican las conclusiones y trabajo a futuro.

2 Estado del arte

2.1 Conceptos de la Teoría de Grafos

La representación de las relaciones entre varios elementos es posible a través de un grafo. En ciencias de la computación, un grafo es una abstracción matemática, con las siguientes características [14]:

- Grafo $G(V, E)$
- V es un conjunto de vértices
- E es un conjunto de aristas

La teoría de grafos cuenta con diversos elementos que facilitan el análisis entre los elementos y sus respectivas relaciones. Algunas de las propiedades más utilizadas para el análisis de la colaboración científica son:

- Grado de un vértice. Es el número total de aristas de un vértice. En otras palabras, la cantidad de relaciones del vértice [15]
- Peso. Si una arista tiene asignada un valor real, se le conoce como peso. Representa el valor de la relación de los vértices involucrados [16]
- Grafo ponderado. Es un grafo donde cada arista tiene asignada un peso [16]
- Grafo etiquetado. Un grafo es etiquetado si tiene etiquetas, ya sea en los vértices, en las aristas, o en ambos [14]. El etiquetado en los vértices de acuerdo con su clasificación facilita visualizar las relaciones entre elementos de diferente naturaleza

2.2 Herramientas de software para la generación de grafos

Existen diversas herramientas que permiten generar grafos a partir de conjuntos de datos con una estructura específica. Una de ellas es Gephi [17], la cual utiliza análisis estadísticos a partir de los conceptos de la teoría de grafos para estudiar el comportamiento entre los diversos elementos, ya sea un grafo dirigido o no dirigido.

Otra herramienta es VOSviewer, enfocada en la representación de redes de colaboración científica [18], cuya funcionalidad se resume en dos partes [19]:

1. Creación de grafos basados en datos de red. Permite construir redes desde cero y a partir de datos de revistas, investigadores, instituciones, países, palabras clave, entre otros, utilizando Web of Science, Scopus, Dimensions y PubMed
2. Visualización y exploración de datos. Presenta al grafo en red, en visualización superpuesta y en visualización de densidad

2.3 Estudios relacionados

Diversos estudios asociados con las colaboraciones científicas representadas a través de la teoría de grafos se han llevado a cabo.

Un estudio [20] identificó a través de un grafo, a partir de un conjunto de datos de la Oficina de Información Científica y Técnica (OSTI, por sus siglas en inglés) del Departamento de Energía (DOE, por sus siglas en inglés) de la Universidad de Tennessee, una fuerte relación entre los estudios sobre medio ambiente con otras áreas de estudio, tales como ciencias nucleares, fósiles y energía, considerando la ventana de tiempo de 1980-2000. En el mismo estudio, en otro grafo, analizando la producción generada entre el año 2000 y el año 2012, se localizó una mayor colaboración entre las ciencias nucleares con otras áreas de estudio, principalmente con la ciencia de la materia y energía.

Por otra parte, se ha propuesto un método para la identificación y análisis de colaboraciones entre investigadores de más de una institución a partir de la producción de Scopus con la finalidad de ubicar las disciplinas más productivas y con mayor colaboración [21]. A partir de los vértices que representan publicaciones, autores y grupos de colaboración, se analiza el estado de la investigación considerando los grupos de colaboración entre todos los miembros, los grupos que no colaboraron con otros investigadores, los investigadores que colaboran con extranjeros, entre otros.

Otro estudio identificó comunidades científicas a través del agrupamiento en grafos bipartitos, uniendo investigadores con revistas científicas, logrando localizar comunidades basadas en la similitud de actividad científica y la agrupación de revistas en función de cuáles son las más representativas de cada comunidad [22].

En lo que respecta a la colaboración con instituciones pertenecientes a otros países, existen algunos conceptos que evalúan esta producción. Algunos de ellos son [23]:

- Intensidad de la colaboración internacional. Es la cantidad de publicaciones de un autor con al menos un investigador perteneciente a una institución extranjera durante una ventana de tiempo
- Tasa de colaboración internacional. Es la relación porcentual de la intensidad de la colaboración internacional respecto al total de publicaciones del autor en una ventana de tiempo
- Amplitud de la colaboración internacional. Es el total de países extranjeros representados en una publicación en colaboración internacional de un investigador

3 Metodología

A través de la metodología de la minería de datos, específicamente en aplicaciones e-learning [24], se proponen los siguientes pasos para la extracción del conocimiento, presentados en la Figura 1.

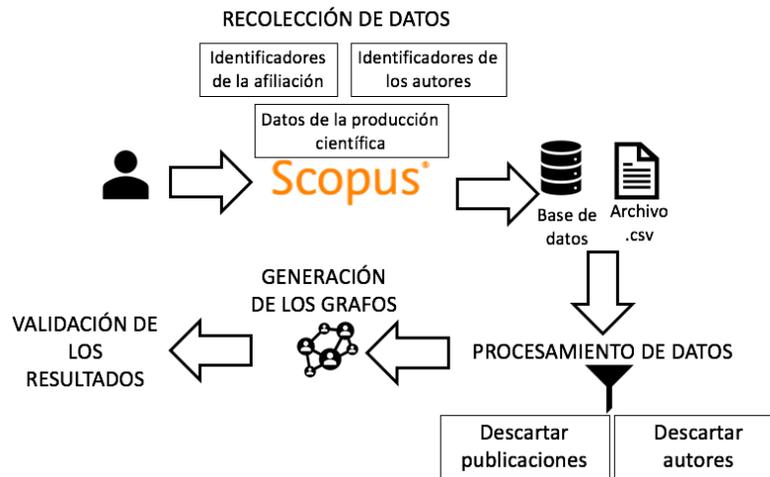


Figura 1. Fases de la metodología propuesta

3.1 Recolección de datos

Para mantener la organización de las publicaciones registradas, Scopus asocia cada producto con uno o más autores, y cada autor se asocia con una o más instituciones. Cada publicación cuenta con un identificador único, mientras que las afiliaciones y los autores pueden tener más de uno.

A continuación, se describe cada paso para extraer los datos de las publicaciones pertenecientes a los autores de una institución educativa o de investigación, considerando la unificación de los posibles múltiples identificadores de sus autores.

3.1.1 Registro de los identificadores de la afiliación

Se lleva a cabo una consulta en el buscador de afiliaciones de Scopus (Figura 2), ingresando el nombre de la institución, o bien, el nombre de la ciudad. Por cada resultado correcto, entrar al perfil y en la parte superior, se deberá localizar el identificador de la afiliación, acompañado por la leyenda *Affiliation-ID*.

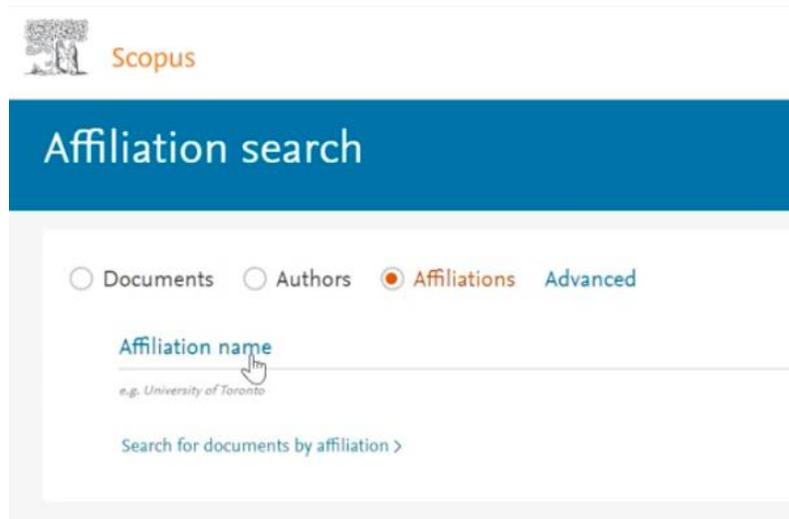


Figura 2. Buscador de afiliaciones de Scopus

Además, por cada perfil de afiliación, se deberá recuperar la producción asociada a la misma, ingresando en *Documents, whole institution* (Figura 3). Como la finalidad es analizar la producción actual, se trabaja con una ventana de tiempo, por lo que, en la herramienta de filtrado ubicada en la barra izquierda, se eligen los años pertenecientes al lapso establecido.

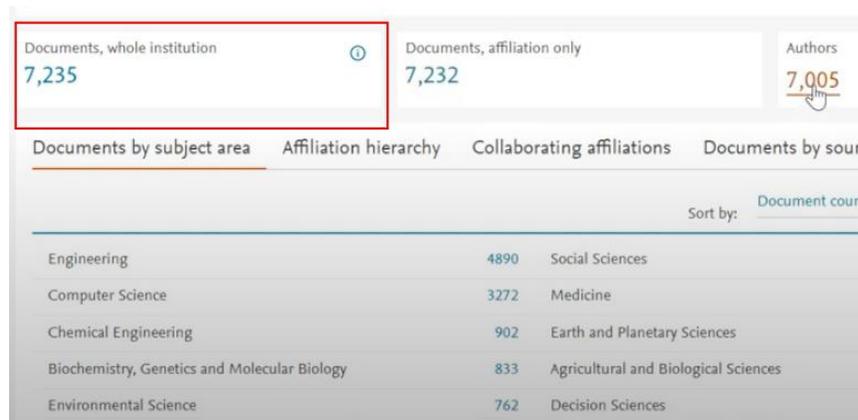


Figura 3. Documentos de una afiliación en Scopus

Posteriormente, en la parte superior de la lista de documentos, elegir la opción *All* para seleccionar todos los documentos, después clic en *Text export* y escoger todos los datos de los productos científicos. El documento deberá ser un archivo *.csv*.

3.1.2 Recuperación de los identificadores de los autores

Una vez registrados los identificadores de la afiliación, se utiliza una base de datos de los investigadores con las siguientes columnas:

- Nombre completo
- Nombre de la institución que le otorgó el último grado de estudios
- Área de estudio

A través de la API para la búsqueda de autores en Scopus, se lleva a cabo una consulta por cada autor con los siguientes parámetros:

- Nombre del investigador
- Identificadores de la afiliación

El resultado se devuelve en formato XML o JSON e incluye los registros de los perfiles que coinciden con la búsqueda. Cada registro de un perfil se compone por una serie de metadatos, entre ellos el identificador de Scopus del investigador. Si los valores de los metadatos del nombre del autor y del área de estudio coinciden con los datos del investigador, se considera como un perfil asociado y se almacena en la base de datos de los investigadores la serie de metadatos, actualizando el registro correspondiente.

Si la búsqueda no obtuvo éxito, entonces se deberá llevar a cabo una segunda consulta con los siguientes parámetros:

- Nombre del investigador
- Nombre de la institución que le otorgó el último grado de estudios

Si los valores de los metadatos del nombre del autor, del nombre de la institución que otorgó el último grado y del área de estudio coinciden con los datos del investigador, se considera como un perfil asociado y se almacena en la base de datos de los investigadores la serie de metadatos, actualizando el registro correspondiente.

3.1.3 Recuperación de los datos de la producción

Con los identificadores de cada investigador recuperados en la fase anterior, por medio de la API de Scopus para la búsqueda de publicaciones, se realizan las consultas correspondientes considerando como parámetros:

- Identificadores de Scopus del investigador
- Identificadores de la afiliación (permite excluir las publicaciones del investigador asociadas a otra institución a la que puede o pudo pertenecer)

Cada resultado contiene el registro de las publicaciones asociadas al autor consultado, las cuales deberán almacenarse en una base de datos de producción científica. Es posible que la autoría de algunas publicaciones pertenezca a más de un autor registrado en la base de datos de los investigadores de la institución, por lo que se deberá validar que el registro no exista en la base de datos de publicaciones antes de almacenarlo, con la finalidad de evitar la multiplicidad.

3.2 Procesamiento de datos

3.2.1 Descartar publicaciones

Después de contar con un documento en formato .csv y una base de datos, ambos con la producción de una institución, se deberán descartar aquellas que no pertenecen a los investigadores registrados en la base de datos. Ambos elementos cuentan con un campo llamado *EID*, que contiene un identificador único del producto.

Por cada registro en el archivo .csv, se realiza una consulta a la base de datos de la producción para identificar si los datos del producto asociado a la clave *EID* están registrados. Si la consulta no retorna algún resultado, el registro se elimina del archivo .csv.

3.2.2 Descartar autores externos y no registrados en la institución

Debido a la posible existencia de productos científicos cuya autoría también pertenece a personas no registradas en la base de datos de investigadores, se requiere omitirlos para analizar únicamente a los autores registrados.

El archivo .csv contiene los campos *Authors* (incluye los nombres de los autores) y *Author(s) ID* (incluye los identificadores de autores en Scopus).

Por cada identificador de autor de cada publicación, se consulta si existe en la base de datos de los investigadores de la institución. Si la consulta no retorna algún resultado, en el archivo .csv se elimina el identificador y el nombre del autor (de los campos *Author(s) ID* y *Authors*, respectivamente).

A cada autor restante en el archivo .csv se le asigna un identificador que incluye el nombre del área de conocimiento al que pertenece.

3.3 Generación de los grafos

Con los datos resultantes en el archivo .csv se generan grafos a través de la herramienta VOSviewer. El primer grafo representa a las colaboraciones entre los investigadores de la institución y tiene las siguientes características:

- Grafo $G1 (V1, E1)$
- $V1$ es un conjunto de vértices que representan a los investigadores de la institución
- $E1$ es un conjunto de aristas que representan a las relaciones de coautoría de los investigadores
- $G1$ es un grafo no dirigido
- $G1$ es un grafo ponderado
- $E1_{i,j}$ contiene el peso de la colaboración entre el investigador i y el investigador j . El valor del peso depende del número de colaboraciones entre ambos investigadores

Un segundo grafo representa a la colaboración internacional de la institución evaluada. Dicho grafo cuenta con las siguientes características:

- Grafo $G2 (V2, E2)$

- $V2$ es un conjunto de vértices que representan a los países que han realizado las colaboraciones internacionales
- $E2$ es un conjunto de aristas que representan a la existencia de colaboración entre dos países
- $G2$ es un grafo no dirigido
- $G2$ es un grafo ponderado
- $E2_{i,j}$ contiene el peso de la colaboración entre el país i y el país j . El valor del peso depende de la intensidad de la colaboración internacional entre ambos países

3.4 Validación de los resultados

La visualización de los grafos generados permite facilitar la toma de decisiones asociadas al impacto y grado de colaboración de una institución. A continuación, se enlistan algunas características de estos grafos para la realización de esta tarea:

- Las etiquetas de los vértices en el grafo de la colaboración entre los investigadores permiten localizar las colaboraciones multidisciplinarias. Además, a partir de los subgrafos se valida cómo han trabajado los grupos de investigación establecidos en la institución
- En el grafo de la colaboración entre los investigadores, la productividad de cada investigador es representada por el tamaño del vértice. A mayor productividad, mayor es el tamaño del vértice
- El peso de cada arista en ambos grafos representa la constancia en las colaboraciones
- En el grafo de la colaboración internacional, cada subgrafo representa una colaboración entre la institución evaluada y uno o más países

4 Implementación

Con la finalidad de evaluar la metodología propuesta en este trabajo, se realizó el caso de estudio de la producción generada entre el 2016 y el 2019 en la Universidad Autónoma de Yucatán (UADY), una importante institución educativa del sureste de México. Hoy en día, la UADY está compuesta por 15 facultades distribuidas en cinco campus, además de un centro de investigación enfocado en dos áreas de estudio [25]. Al día 1 de junio de 2018, en la UADY se registraron 824 profesores de tiempo completo y 78 cuerpos académicos [26].

Gracias a la generación del archivo .csv de la producción científica de la UADY registrada en Scopus, se encontró que la universidad generó 1190 publicaciones entre el 2016 y el 2019.

Se crearon las bases de datos de los profesores de tiempo completo y de la producción científica en MongoDB, una base de datos documental que facilita el almacenamiento en formato JSON [27]. A través de un script en Python se utilizó la API de búsqueda de autores en Scopus obteniendo los valores de los parámetros de la base de datos de MongoDB. Los identificadores de autor fueron insertados en los registros correspondientes.

Por medio de otro script de Python, haciendo uso de la API de Scopus para la búsqueda de publicaciones, se consultaron los productos asociados a los profesores registrados en la base de datos, así como su validación de inexistencia en la base de datos correspondiente para su inserción.

Sin embargo, al depurar los registros de las publicaciones que no pertenecen a los 824 profesores de tiempo completo, se encontró que generaron 900 publicaciones distribuidas de la siguiente manera: 213 en el 2016, 207 en el 2017, 284 en el 2018 y 196 en el 2019.

Por otra parte, se identificó que, de los 824 profesores, 315 cuentan con por lo menos una publicación en la ventana de tiempo establecida, de los cuales, 280 colaboraron con otros profesores de la UADY.

4.1 Colaboraciones entre los investigadores

Para identificar a los profesores más productivos y colaboradores, se consideraron para la creación del grafo aquellos cuya producción es mayor o igual a cuatro productos en cuatro años, compartiendo por lo menos cuatro coautorías.

VOSviewer presenta el grafo compuesto por subgrafos (o clústers). Se determinó que los subgrafos deben incluir por lo menos tres vértices. De los 280 profesores colaboradores, 109 cumplen con estas condiciones. Con los 109 profesores más productivos y colaboradores se generó un grafo, compuesto por 15 subgrafos, expuestos en la Figura 4. Cabe mencionar que tres profesores no se agruparon en alguno de los subgrafos. En la Tabla 1 se describen los subgrafos, el número de elementos asociados, las áreas de conocimiento involucradas, así como el grado medio y el grado medio con pesos, representando al promedio de colaboradores de cada profesor dentro de su subgrafo y al promedio de publicaciones de cada profesor en colaboración dentro de su subgrafo, respectivamente.

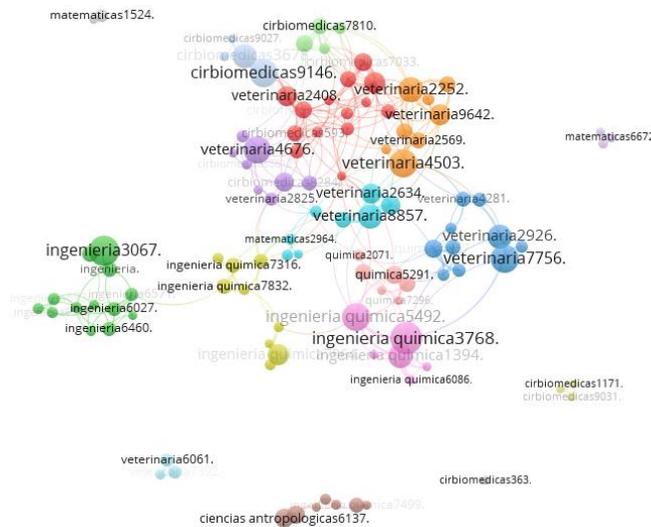


Figura 4. Grafo de colaboración interna de la UADY

Tabla 1: Descripción de los subgrafos de colaboración identificados en la UADY

Grupo	No. de profesores	Composición	Grado medio	Grado medio con pesos
-------	-------------------	-------------	-------------	-----------------------

1	14	5 de Biomedicina, 2 de Medicina y 7 de Veterinaria	6.14	14.57
2	13	13 de Ingeniería	5.07	12.61
3	10	5 de Química y 5 de Veterinaria	3.4	13.2
4	9	9 de Ingeniería Química	3.11	8.88
5	8	5 de Biomedicina y 3 de Veterinaria	3.75	8.25
6	8	1 de Matemáticas y 7 de Veterinaria	2.75	16.75
7	8	8 de Veterinaria	3.75	15.5
8	7	2 de Ciencias Antropológicas, 2 de Ingeniería Química y 3 de Odontología	2	5.42
9	6	5 de Ingeniería Química y 1 de Medicina	3.66	21.33
10	6	6 de Química	2.66	10.33
11	4	4 de Biomedicina	3	14.5
12	4	4 de Biomedicina	3	17.5
13	3	3 de Biomedicina	2	8
14	3	3 de Matemáticas	2	3.33
15	3	3 de Veterinaria	2	10

4.2 Colaboraciones con otros países

De las 900 publicaciones, se identificaron colaboraciones con 64 países. Se entiende como una colaboración con un país si una publicación ha sido realizada en conjunto con una o más instituciones del país en cuestión. Es decir, si en una publicación existen tres colaboradores de Estados Unidos, de los cuales, dos pertenecen a la Universidad de Standford y uno a la Universidad de Yale, se considera que se ha colaborado con un solo país. Se determinó identificar los países que cuenten por lo menos con cuatro colaboraciones con la UADY, cuyo resultado fue de 25. En la Tabla 2 se presentan los cinco países con mayor colaboración internacional con la UADY, las publicaciones en colaboración y el porcentaje respecto con la colaboración internacional total.

Tabla 2: Los cinco países con mayor colaboración con la UADY

Posición	País	Publicaciones en colaboración	% de colaboración
1	Estados Unidos	150	33.11%
2	España	57	12.58%
3	Francia	30	6.62%
4	Argentina	23	5.07%
5	Brasil	18	3.97%

En la Figura 5 se presenta el grafo de las colaboraciones internacionales, donde cada vértice representa un país y el grosor de cada arista depende de la intensidad de la colaboración internacional entre ambos países.

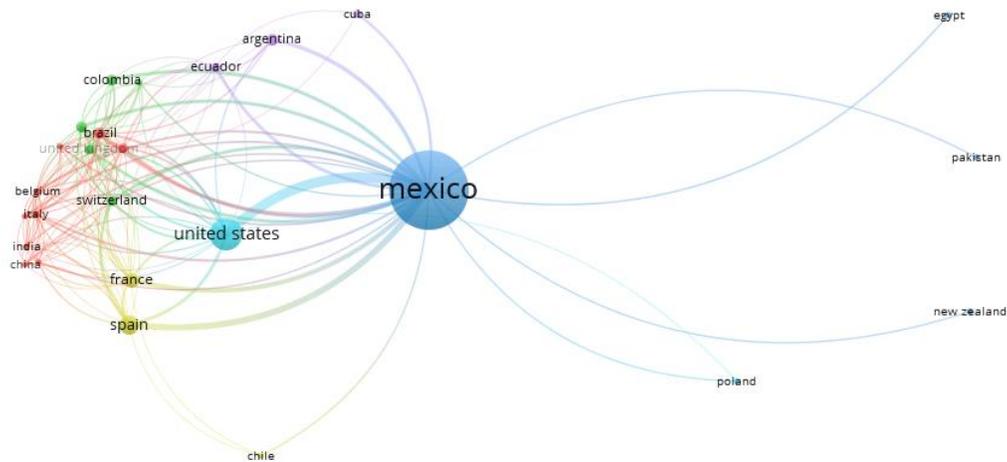


Figura 5. Grafo de colaboración internacional de la UADY

Una de las ventajas del grafo generado con VOSviewer, es la localización de redes de colaboración con otros países. Un ejemplo se presenta en la Figura 6, donde se observa que la UADY, al trabajar con instituciones argentinas, a la vez ha colaborado con instituciones de Estados Unidos, Francia, Brasil, Uruguay y Ecuador.

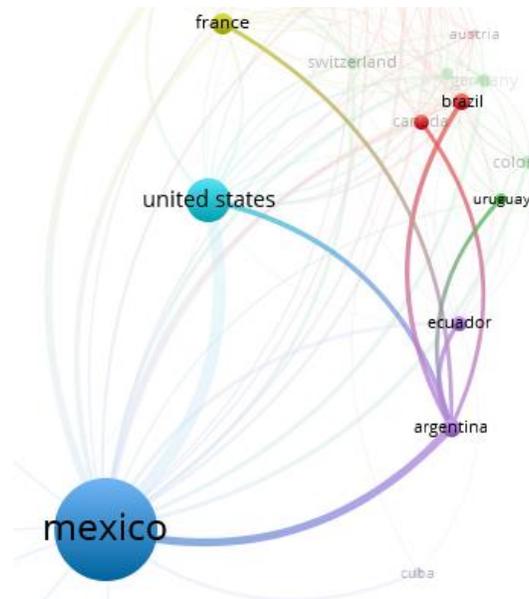


Figura 6. Ejemplo de red de colaboración internacional de la UADY

5 Conclusiones y trabajos futuros

Se espera que los grafos presentados permitan evaluar el estado de la colaboración entre los miembros de cuerpos académicos, así como entre los demás investigadores, facilitando la ejecución

de la toma de decisiones e incluso, en conjunto con técnicas de aprendizaje automático, recomendar a los investigadores colaborar con grupos que representen un gran impacto en la comunidad científica.

Se pretende que a partir de la extracción de los datos relacionados con las publicaciones científicas de una institución se identifiquen nuevos grupos que puedan llegar a consolidarse, además de localizar elementos que podrían aportar mayor relevancia a los grupos de investigación ya establecidos.

La propuesta presentada en este trabajo utiliza como fuente de datos a Scopus, la cual contiene un amplio catálogo de revistas y trabajos indexados. Sin embargo, se espera en el futuro anexar datos de otras bases de datos bibliográficas, como Web of Science, y en conjunto, llevar a cabo un análisis de las colaboraciones con mayor completitud.

Así como se ha presentado la unificación de los diversos perfiles de Scopus de los autores, es posible ejecutar una tarea similar con los diversos perfiles de las instituciones nacionales y extranjeras que han colaborado con la UADY, con la finalidad de detectar los productos resultantes de los convenios vigentes.

A partir de los términos utilizados tanto en las palabras clave como en los resúmenes, se propone elaborar grafos que permitan identificar la coocurrencia entre los diversos términos, con la finalidad de validar las líneas de investigación con las temáticas publicadas, además de conocer cuáles grupos de investigación son multidisciplinarios.

Referencias

- [1] D. D. B. Beaver, "Reflections on scientific collaboration (and its study): Past, present, and future", *Scientometrics*, vol. 52, núm. 3, pp. 365–377, 2001, doi: 10.1023/A:1014254214337.
- [2] E. Leahey, "From Sole Investigator to Team Scientist: Trends in the Practice and Study of Research Collaboration", *Annu. Rev. Sociol.*, vol. 42, núm. 1, pp. 81–100, 2016, doi: 10.1146/annurev-soc-081715-074219.
- [3] Y. Bu, Y. Ding, X. Liang, y D. S. Murray, "Understanding persistent scientific collaboration", *J. Assoc. Inf. Sci. Technol.*, vol. 69, núm. 3, pp. 438–448, 2018, doi: 10.1002/asi.23966.
- [4] J. Bar-Ilan, "Which h-index? — A comparison of WoS, Scopus and Google Scholar", *Scientometrics*, vol. 74, núm. 2, pp. 257–271, 2007, doi: 10.1007/s11192-008-0216-y.
- [5] Elsevier, "Affiliation Search API", 2020. <https://dev.elsevier.com/documentation/AffiliationSearchAPI.wadl>. Recuperado el 20 de marzo de 2020.
- [6] Elsevier, "Author Search API", 2020. <https://dev.elsevier.com/documentation/AuthorSearchAPI.wadl>. Recuperado el 20 de marzo de 2020.
- [7] Elsevier, "Scopus Search API", 2020. <https://dev.elsevier.com/documentation/ScopusSearchAPI.wadl>. Recuperado el 20 de marzo de 2020.
- [8] CONACYT, "Sistema Nacional de Investigadores", 2020. <https://www.conacyt.gob.mx/index.php/el-conacyt/sistema-nacional-de-investigadores>. Recuperado el 20 de marzo de 2020.
- [9] DGESU, "Programa para el Desarrollo Profesional Docente, para el Tipo Superior (PRODEP)", 2020. <http://www.dgesu.ses.sep.gob.mx/PRODEP.htm>. Recuperado el 20 de marzo de 2020.

- [10] J. Guerrero-Sosa, V. Menéndez-Domínguez, M. Castellanos-Bolaños, y J. Gómez-Montalvo, "Use of an Ontological Model to Assess the Relevance of Scientific Production", *IEEE Lat. Am. Trans.*, vol. 17, núm. 09, pp. 1424–1431, 2019, doi: 10.1109/TLA.2019.8931135.
- [11] V. H. Menéndez Domínguez, J. D. T. Guerrero Sosa, M. E. Castellanos Bolaños, y E. Zurita Gallegos, "Análisis de la producción de cuerpos académicos basado en teoría de grafos", *RIDE Rev. Iberoam. Para La Investig. Y El Desarro. Educ.*, vol. 10, núm. 20, 2020, doi: 10.23913/ride.v10i20.603.
- [12] J. D. T. Guerrero-Sosa, F. P. Romero Chicharro, J. Serrano-Guerrero, V. Menendez-Dominguez, y M. E. Castellanos-Bolaños, "A proposal for a recommender system of scientific relevance", *Procedia Comput. Sci.*, vol. 162, pp. 199–206, 2019, doi: <https://doi.org/10.1016/j.procs.2019.11.276>.
- [13] J. D. T. Guerrero-Sosa, V. H. Menéndez-Domínguez, M.-E. Castellanos-Bolaños, y F. Moo-Mena, "Document Database for Scientific Production", en *8th International Workshop on ADVANCES in ICT Infrastructures and Services*, 2020, pp. 129–132.
- [14] J. M. Sallán, J. B. Fonollosa, V. Fernández, y A. Suñé, "Teoría de grafos", en *Métodos cuantitativos en organización industrial I*, J. M. Sallán Leyes, Ed. Edicions UPC, 2002, pp. 137–172.
- [15] M. Álvarez y J. Parra, "Teoría de grafos", Universidad del Bío-Bío, 2013.
- [16] R. J. Trudeau, *Introduction to Graph Theory*. Dover Pub., 1993.
- [17] Gephi Consortium, "Gephi - The Open Graph Viz Platform", 2017. <https://gephi.org/> Recuperado el 20 de marzo de 2020.
- [18] N. J. van Eck y L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping", *Scientometrics*, vol. 84, núm. 2, pp. 523–538, 2010, doi: 10.1007/s11192-009-0146-3.
- [19] N. J. van Eck y L. Waltman, "VOSviewer manual", *Leiden: Univeriteit Leiden*. 2020, [En línea]. Disponible en: http://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.1.pdf.
- [20] H. Bozdogan y O. Akbilgic, "Social network analysis of scientific collaborations across different subject fields", *Inf. Serv. Use*, vol. 33, núm. 3–4, pp. 219–233, 2013, doi: 10.3233/ISU-130715.
- [21] F. G. Montoya, A. Alcayde, R. Baños, y F. Manzano-Agugliaro, "A fast method for identifying worldwide scientific collaborations using the Scopus database", *Telemat. Informatics*, vol. 35, núm. 1, pp. 168–185, 2018, doi: 10.1016/j.tele.2017.10.010.
- [22] C. Carusi y G. Bianchi, "Scientific community detection via bipartite scholar/journal graph co-clustering", *J. Informetr.*, vol. 13, núm. 1, pp. 354–386, 2019, doi: <https://doi.org/10.1016/j.joi.2019.01.004>.
- [23] R. Todeschini y A. Baccini, *Handbook of Bibliometric Indicators: Quantitative Tools for Studying and Evaluating Research*. Wiley, 2016.
- [24] M. E. Prieto, A. Zapata, y V. H. Menendez, "Data Mining Learning Objects", en *Handbook of Educational Data Mining*, C. Romero, S. Ventura, y M. Pechenizkly, Eds. CRC Press, 2011, pp. 481–492.
- [25] UADY, "Universidad Autónoma de Yucatán", 2019. <https://www.uady.mx/nuestra-universidad>. Recuperado el 20 de marzo de 2020.
- [26] UADY, "Los Cuerpos Académicos, pilar de las funciones sustantivas de la UADY", 2018. <https://www.uady.mx/noticia/cuerpos-academicos-pilar-de-las-funciones-sustantivas-de-la-uady>. Recuperado el 25 de marzo de 2020.
- [27] "¿Qué es MongoDB?". <https://www.mongodb.com/es/what-is-mongodb>. Marzo 25, 2020.