

<https://doi.org/10.23913/ride.v10i20.603>

*Artículos Científicos*

## **Análisis de la producción de cuerpos académicos basado en teoría de grafos**

*Analysis of the Production of Academic Groups Based on Graph Theory*

*Análise da produção de órgãos acadêmicos com base na teoria dos grafos*

**Víctor Hugo Menéndez Domínguez**

Universidad Autónoma de Yucatán, México

[mdoming@correo.uady.mx](mailto:mdoming@correo.uady.mx)

<https://orcid.org/0000-0003-3587-1263>

**Jared David Tadeo Guerrero Sosa**

Universidad Autónoma de Yucatán, México

[jaredgs93@gmail.com](mailto:jaredgs93@gmail.com)

<http://orcid.org/0000-0001-7999-9870>

**María Enriqueta Castellanos Bolaños**

Universidad Autónoma de Yucatán, México

[enriqueta.c@correo.uady.mx](mailto:enriqueta.c@correo.uady.mx)

<https://orcid.org/0000-0001-6294-5948>

**Esmeralda Zurita Gallegos**

Instituto Tecnológico Superior de Felipe Carrillo Puerto, México

[141k0057@itscarrillopuerto.edu.mx](mailto:141k0057@itscarrillopuerto.edu.mx)

<https://orcid.org/0000-0002-3450-9599>

## Resumen

En muchas universidades mexicanas, grupos de profesores conforman cuerpos académicos con la finalidad de generar nuevos conocimientos con base en la colaboración de sus miembros. Estos cuerpos académicos son clasificados siguiendo criterios establecidos por el Programa para el Desarrollo Profesional Docente (Prodep). Sin embargo, su valoración puede resultar subjetiva por la ausencia de herramientas que faciliten dicho proceso. Este trabajo se enfocó en analizar la producción y la colaboración de tres tipos de cuerpos académicos a partir de la información almacenada en la base de datos Scopus, haciendo uso de la visualización de datos, aplicando la teoría de grafos. La metodología utilizada se basa en la minería de datos educativos. Los resultados permitieron observar una correspondencia entre la estructura del grupo y el cumplimiento de la producción solicitada por el Prodep; los elementos más fuertes y más débiles de cada cuerpo académico fueron localizados con base en su participación y nivel de contribución. El aporte de esta investigación es el análisis de la producción y colaboración científica de un cuerpo académico haciendo uso de la teoría de grafos, lo que permite automatizar el proceso de evaluación y, de esta manera, reducir su interpretación subjetiva.

**Palabras clave:** colaboración científica, cuerpos académicos, producción científica, teoría de grafos, visualización de datos.

## Abstract

In many Mexican universities, groups of professors form academic groups with the purpose of generating new knowledge based on the collaboration of its members. These academic groups are classified according to the criteria established by the Program for Professional Development of Teachers (PRODEP), which is based on the relevance of production and the degree of collaboration; however, such evaluation may be subjective due to the absence of tools that facilitate that process. The present work is focused on analyzing the production and collaboration of three types of academic groups from the information stored in the Scopus database using data visualization, specifically, graph theory. The used methodology is based on educational data mining. The results allowed to observe a correspondence between the structure of the group and the fulfillment of the production requested by the PRODEP, locating the strongest and the weakest elements of each academic group based on their participation and level of contribution. Our research value is the production and scientific collaboration analysis of an academic group, using a data visualization

tool, in this case, graph theory, which allows to automate the evaluation process by reducing its subjective interpretation.

**Keywords:** scientific collaboration, academic groups, scientific production, graph theory, data visualization.

## Resumo

Em muitas universidades mexicanas, grupos de professores formam órgãos acadêmicos para gerar novos conhecimentos com base na colaboração de seus membros. Esses órgãos acadêmicos são classificados de acordo com os critérios estabelecidos pelo Programa de Desenvolvimento Profissional de Professores (Prodep). No entanto, sua avaliação pode ser subjetiva devido à ausência de ferramentas que facilitem esse processo. Este trabalho teve como objetivo analisar a produção e colaboração de três tipos de órgãos acadêmicos, com base nas informações armazenadas no banco de dados Scopus, utilizando visualização de dados, aplicando a teoria dos grafos. A metodologia utilizada é baseada na mineração de dados educacionais. Os resultados permitiram observar uma correspondência entre a estrutura do grupo e o cumprimento da produção solicitada pelo Prodep; Os elementos mais fortes e fracos de cada corpo acadêmico foram localizados com base em sua participação e nível de contribuição. A contribuição desta pesquisa é a análise da produção e colaboração científica de um corpo acadêmico utilizando a teoria dos grafos, que permite automatizar o processo de avaliação e, assim, reduzir sua interpretação subjetiva.

**Palavras-chave:** colaboração científica, órgãos acadêmicos, produção científica, teoria dos grafos, visualização de dados.

**Fecha Recepción:** Septiembre 2019

**Fecha Aceptación:** Enero 2020

## Introducción

En la comunidad científica es una actividad natural la colaboración entre dos o más miembros (González y Gómez, 2014) que en conjunto aplican sus conocimientos para generar nuevos, los cuales pueden reflejarse en productos de carácter científico, educativo, tecnológico o de innovación. Incluso en México ha existido la preocupación por esta actividad, lo que ha dado lugar a una política que estimula la colaboración científica en las instituciones públicas de educación superior por medio de la creación de los cuerpos académicos (López, 2010). La forma más sencilla y eficiente de conocer el desarrollo y evaluar a un cuerpo académico es a través de sus publicaciones en conjunto (Dirección General de Educación Superior Universitaria [Dgesu], 2018). Existen estudios que se han enfocado en representar la colaboración científica por medio de la teoría de grafos: desde la construcción del grafo y sus resultados fundamentales (Newman, 2001a) hasta la aplicación práctica de la teoría del camino más corto, redes ponderadas y la centralidad del grafo (Newman, 2001b), así como el uso de un sistema para la explotación visual de redes de colaboración (Giatsidis, Berberich, Thilikos y Vazirgiannis, 2012) y el análisis de las redes de colaboración científica considerando factores sociales, su evolución y agrupamiento tópico (Staudt, 2011). De igual manera, se ha estudiado este ámbito para áreas específicas, como la biomédica, a través del análisis de las características de la producción científica en Latinoamérica, donde las redes de colaboración se graficaron para obtener la centralidad de cada red o grafo (Huamaní, González, Curioso y Pacheco, 2012).

Este artículo presenta un análisis del nivel de producción científica de tres distintos grupos de docentes por medio de la teoría de grafos. Dichas tres agrupaciones cuentan con un estatus diferente: un cuerpo académico en formación (CAEF), un cuerpo académico en consolidación (CAEC) y un cuerpo académico consolidado (CAC), los cuales se encuentran afiliados a una universidad pública mexicana. Los grafos que se obtienen funcionan como una herramienta de visualización para la evaluación de la productividad y rendimiento de los cuerpos académicos. Los resultados obtenidos podrían tener múltiples aplicaciones; por ejemplo, la medición de la relevancia en la producción de un determinado grupo de docentes, identificar el grado de colaboración que existe entre cada cuerpo académico o ayudar a mejorar el desempeño de estos, entre otros.

El resto del artículo está organizado de la siguiente manera: en el estado del arte se describen los conceptos necesarios para comprender el caso de estudio. Posteriormente se presenta la metodología propuesta. En el caso de estudio se retoman los conceptos del estado del arte y, en conjunto con la metodología, se presenta el procedimiento y los resultados del análisis de la producción

científica de los cuerpos académicos en cuestión. Finalmente, se discuten los aspectos interesantes de los resultados y las conclusiones.

### **Estado del conocimiento**

Un cuerpo académico no es únicamente una fracción de la comunidad científica, ya que tiene ciertas características específicas. De forma general, es un conjunto de profesores-investigadores que tienen en común una o varias líneas de estudio y tienen como objetivo tanto la aplicación como la generación de nuevos conocimientos a raíz del trabajo en conjunto. En México, los cuerpos académicos se clasifican en tres grupos y cada uno tiene características específicas para las universidades estatales y afines, institutos y universidades tecnológicas (Dgesu, 2018). Las características que se presentan a continuación corresponden a las universidades estatales y afines, debido a que el caso de estudio se aplica con cuerpos académicos de una universidad perteneciente a ese grupo.

- **Cuerpo académico consolidado (CAC).** Es el nivel máximo que puede alcanzar un cuerpo académico. Sus características son las siguientes: 1) la mayoría de sus integrantes cuenta con la máxima habilitación académica para generar o aplicar de forma innovadora e independiente el conocimiento (doctorado); 2) cuentan con una amplia experiencia en la docencia y en la formación de recursos humanos; 3) la mayoría de sus integrantes cuentan con el perfil deseable definido por el Programa para el Desarrollo Profesional Docente, para el Tipo Superior (Prodep); 4) alto compromiso con la institución por medio de la colaboración y producción científica y académica; 5) demuestran una intensa actividad académica manifiesta en congresos, seminarios, mesas, talleres, etc., de forma regular y frecuente, y 6) sostienen una intensa participación en redes de intercambio académico (Dgesu, 2018).

- **Cuerpo académico en consolidación (CAEC).** Es el nivel intermedio en el que se puede clasificar un cuerpo académico. Se caracteriza por los siguientes aspectos: 1) más de la mitad de sus integrantes cuentan con doctorado; 2) cuentan con productos académicos con reconocimiento debido a su buena calidad, derivados de las líneas de investigación consolidadas; 3) al menos la tercera parte de sus integrantes cuentan con el perfil deseable definido por el Prodep; 4) participan de forma conjunta en líneas de investigación o aplicación innovadora del conocimiento; 5) cuentan con amplia experiencia en docencia y en formación de recursos humanos, y 6) colaboran con otros cuerpos académicos (Dgesu, 2018).

- Cuerpo Académico en Formación (CAEF). Como su nombre lo indica, se refiere a los cuerpos académicos que nacen a partir de una o más líneas de investigación y se encuentra en una etapa temprana. Sus características son: 1) los integrantes están identificados; 2) al menos la mitad de sus integrantes cuentan con el perfil deseable definido por el Prodep; 3) tienen definidas las líneas de generación o aplicación del conocimiento que cultivarán, y 4) tienen identificados los cuerpos académicos afines al que proponen y de alto nivel para establecer contacto (Dgesu, 2018).

Lo que se busca es que los cuerpos académicos tengan un verdadero impacto en la sociedad, por lo que es necesario que den a conocer sus trabajos en publicaciones de alto impacto. Cada cuerpo académico tiene una o más líneas de investigación y las publicaciones que han generado en conjunto pueden localizarse en repositorios que se enfoquen en dichas áreas de investigación. Algunos de estos repositorios son los siguientes: dblp para ciencias de la computación (dblp team [dblp], 2018), arXiv para física, matemáticas, ciencias de la computación, finanzas cuantitativas y estadística (Cornell University, 2018) y Merlot para recursos educativos abiertos (California State University System [CSU], 2018), entre otros. Pero las instituciones de nivel superior se enfocan no únicamente en las áreas de investigación mencionadas y no todos los repositorios cuentan con publicaciones de alto impacto. Entonces es necesario consultar alguna solución que cumpla con ambas características, y Scopus es la herramienta adecuada. Scopus es la mayor base de datos de citas y resúmenes de bibliografía revisada a través del método por pares. Contiene registros de revistas científicas, libros y actas de conferencias en una gran variedad de campos de la ciencia (Elsevier, 2018). Scopus presenta herramientas inteligentes para localizar, analizar y visualizar la investigación, las cuales facilitan la búsqueda de autores e información adicional como las instituciones y publicaciones. También cuenta con su propia interfaz de programación de aplicaciones (API, por sus siglas en inglés), que permite recuperar información a través de servicios web (Elsevier, 2017).

Por otra parte, un elemento fundamental en la investigación con datos es su visualización. En efecto, es válido pensar que es algo sencillo de realizar, pero la complejidad se encuentra en producir buenas representaciones visuales. En general, la visualización de datos tiene como finalidad facilitar la exploración y comunicación de los datos (Grus, 2015), así como para su presentación. A pesar de que la visualización de datos se ha aplicado mayormente en la estadística, es gracias a la computación que se obtienen más beneficios al usar herramientas que permiten acceder a información de forma intuitiva con base en datos (Chen, Härdle y Unwin, 2007); además, el usuario, por medio de la observación, es capaz de comprender la información útil o relevante del

aspecto a estudiar, debido a que la interacción con datos proporciona un nivel mayor de comprensión (Aparicio y Costa, 2015). Una técnica útil para la visualización y relación entre datos es el grafo. En ciencias de la computación, un grafo es una abstracción matemática, representada como  $G = (V, E)$ , donde  $V$  es un conjunto de vértices y  $E$  es un conjunto de aristas. De este modo, los grafos son útiles para modelar relaciones entre elementos y permite la resolución de problemas asociados al contexto en el que se encuentren, y requiere un proceso menos costoso que incluso la programación lineal. Además, para representar un grafo existen varias opciones: representación gráfica, representación por medio de una matriz asociada o de adyacentes, y otras representaciones, como el diccionario del grafo (Sallán, Suñé, Fernández y Fonllosa, 2002). Como el enfoque del presente artículo es representar grafos de manera gráfica y analizar sus resultados, no se profundizará en los otros tipos de representación. La representación gráfica de un grafo consiste en presentar cada vértice como un punto o círculo, generalmente. Aunque se puede representar por medio de otras figuras, dependiendo de lo que se desea graficar. En la teoría de grafos existen conceptos interesantes que son objeto de estudio. Entre ellas se encuentran:

- Grado de un vértice. Es el número total de aristas que inciden en dicho vértice. Se denota como  $g(v)$  (Álvarez y Parra, 2013).
- Peso. Cuando a cada arista de un grafo  $G$  se le asigna un valor real  $w(e)$ , se le conoce como peso (Trudeau, 2017).
- Grafo ponderado. El grafo  $G$  junto con los pesos en sus aristas (Trudeau, 2017).
- Grafo etiquetado. El grafo  $G$  cuenta con etiquetas ya sea en los vértices, aristas o en ambas. Cuando el grafo tiene etiquetas en ambas, se dice que es un *grafo completamente etiquetado* (Sallán *et al.*, 2002).
- Grafo dirigido. Un grafo  $G$  donde es necesario representar el vértice origen y el vértice destino, representando las aristas por medio de flechas (Sallán *et al.*, 2002).
- Grafo no dirigido. Las relaciones entre los vértices de un grafo  $G$  se representa por medio de líneas (Sallán *et al.*, 2002).

Los conceptos enlistados no son los únicos; solo se mencionan los que son de utilidad para el caso de estudio del presente artículo. Existen herramientas que permiten crear grafos a partir de tablas en archivos con extensión .xls y .csv. Una de ellas es NodeXL, que se extiende a Excel para generar gráficos de red y grafos (Social Media Research Foundation, 2018). Otras alternativas son Gephi (Gephi Consortium, 2018) y Google Fusion Tables (Google, 2018).

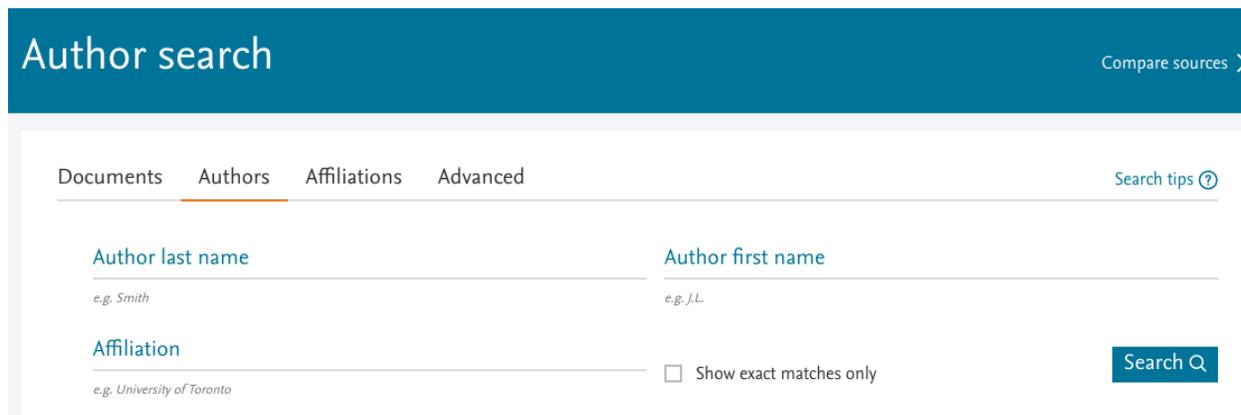
## Metodología

Haciendo uso de la metodología de la minería de datos, específicamente en aplicaciones de *e-learning* (Prieto, Menéndez y Zapata, 2010), se propone una de extracción del conocimiento, la cual consiste en las siguientes etapas.

### Recolección de datos

Una vez que se identifiquen los cuerpos académicos y sus miembros, se recolecta toda la información acerca de sus publicaciones utilizando Scopus, ya que, como se ha mencionado, esta herramienta posee datos de revistas científicas y otras publicaciones de alto impacto, y se espera que un cuerpo académico cuente con ellas para elevar su estatus, o en caso de ser consolidado, mantenerlo. Se realiza una consulta en el sitio web de Scopus por cada miembro del cuerpo académico, ingresando su nombre y la institución a la que pertenece, tal como se observa en el formulario de la figura 1.

**Figura 1.** Formulario de búsqueda de perfiles en Scopus



The image shows the 'Author search' interface on the Scopus website. It features a dark blue header with the text 'Author search' and a 'Compare sources >' link. Below the header, there are four tabs: 'Documents', 'Authors' (which is selected), 'Affiliations', and 'Advanced'. To the right of these tabs is a 'Search tips ?' link. The main search area contains three input fields: 'Author last name' with the example 'e.g. Smith', 'Author first name' with the example 'e.g. J.L.', and 'Affiliation' with the example 'e.g. University of Toronto'. There is also a checkbox labeled 'Show exact matches only' and a blue 'Search Q' button.

Fuente: Scopus (2019)

### Procesamiento de datos

Esta etapa se encuentra dividida en cuatro rubros.

1) Descartar los datos duplicados obtenidos en la búsqueda. Al realizar una revisión en Scopus de cada miembro del cuerpo académico es lógico que, al tratarse de colaboraciones, existan publicaciones que pertenecen a dos o más miembros del cuerpo académico.

2) Descartar las publicaciones que no tienen relación con algún otro miembro del cuerpo académico, así como las publicaciones individuales.

3) Asignar del grado de colaboración de cada miembro del cuerpo académico participe en cada publicación. Para esto, se utiliza como base el orden de los autores en la lista de autoría, considerando que este depende del grado de contribución de cada autor en la publicación, siendo el primer autor el que más aportó, y así sucesivamente. No se consideran a los investigadores externos al cuerpo académico. Por ejemplo, si en un artículo participan cinco autores, de los cuales dos pertenecen al cuerpo académico, y el primero, tercero y cuarto autor no, el segundo autor se considerará como primer autor y el quinto como el segundo (por ser miembros del cuerpo académico).

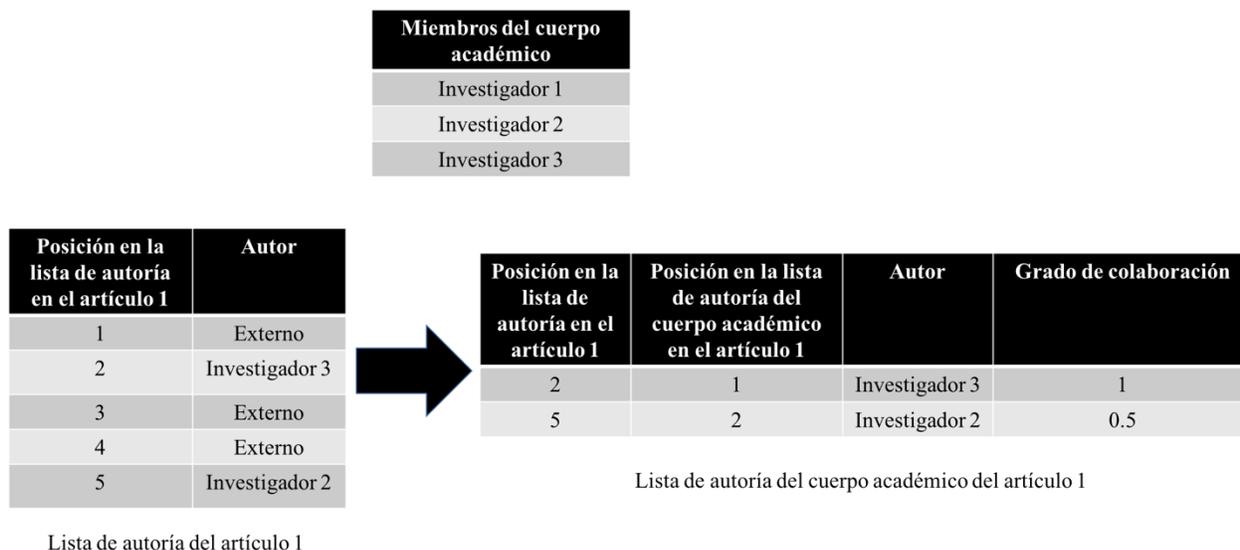
4) Calcular el grado y porcentaje de contribución por integrante en las publicaciones.

A la lista generada en el paso 3 se le denominará *Lista de autoría del cuerpo académico*. El grado de colaboración se obtiene a través de la siguiente fórmula:

$$\text{Grado de colaboración} = 2 * 2^{-p}$$

Donde  $p$  es la posición del miembro del cuerpo académico en la lista de autoría del cuerpo académico en la publicación. Lo que se busca con el grado de colaboración es que su valor para el primer autor del cuerpo académico sea 1; del segundo autor, 0.5; del tercer autor, 0.25, y así sucesivamente. En la figura 2 se presenta un cuerpo académico junto con sus integrantes a modo de ejemplo, así como un artículo denominado *Artículo 1* y la lista de autoría que da lugar a la Lista de autoría del cuerpo académico, en la que se ha eliminado a los autores no pertenecientes al cuerpo académico. También se presenta el grado de colaboración de los miembros participantes en el Artículo 1.

**Figura 2.** Ejemplo de una publicación perteneciente a un cuerpo académico



Fuente: Elaboración propia

Por otra parte, el grado de contribución de cada miembro del cuerpo académico se calcula a través de la siguiente fórmula:

$$GCT = \sum_{i=1}^n \text{Grado de colaboración}_i$$

Donde:

- $n$  es el total de publicaciones donde participa el miembro en cuestión, y
- $GCT$  es la suma de los pesos en las aristas que conectan al vértice que representa al miembro del cuerpo académico, y representa el grado de colaboración total de un integrante al cuerpo académico.

A partir del valor de  $GCT$  se puede obtener el porcentaje de contribución de un investigador dentro de su cuerpo académico siguiendo la siguiente fórmula:

$$PC = \frac{B * GCT}{t}$$

Donde:

- $PC$  es el porcentaje de contribución del miembro del cuerpo académico (porcentaje);
- $B = 100$ , haciendo referencia al 100 % de las publicaciones del cuerpo académico;

- $t$  es el número total de publicaciones del cuerpo académico, y
- $GCT$  es el grado de colaboración total del investigador al cuerpo académico.

### Visualización

Con los datos resultantes, producto de las dos etapas anteriores, se realizan tablas en NodeXL para generar grafos y llevar a cabo el posterior análisis de los aspectos relevantes sobre el cuerpo académico estudiado. Cada grafo  $G$  se representa como:

- Grafo  $G(V, E)$ ;
- $V$  es un conjunto de vértices;
- $E$  es un conjunto de aristas;
- $VI$  es subconjunto de  $V$  y  $VI = \{\text{Vértices que representan a los investigadores, miembros del cuerpo académico}\}$ , y
- $VP$  es un subconjunto de  $V$  y  $VP = \{\text{Vértices que representan a las publicaciones del cuerpo académico}\}$ .

Asimismo, el grafo se construye manteniendo las siguientes características:

- El grafo no es dirigido.
- El grafo es ponderado, y los pesos de las aristas son los valores del grado de colaboración de cada investigador del cuerpo académico involucrado en cada publicación.
- El grafo es etiquetado, y los vértices pertenecientes a  $VI$  se etiquetan con el prefijo  $I$ , posteriormente se coloca el número identificador del investigador. Los vértices pertenecientes a  $VP$  se etiquetan con el prefijo  $P$  y consecutivamente se coloca el número identificador de la publicación.

### Caso de estudio

El estudio que se realizó fue el análisis de la producción de tres cuerpos académicos. Cada uno pertenece a un grupo distinto: CAEF, CAEC y CAC. Los tres cuerpos académicos que se analizaron pertenecen a la Universidad Autónoma de Yucatán (UADY), una importante institución pública del sureste de México que pertenece a la categoría de universidades estatales y afines. Cada cuerpo académico clasificado pertenece a una línea de investigación diferente.

## Cuerpo académico en formación

El CAEF analizado se enfoca en la diversidad y conservación de recursos vegetales tropicales, con un impacto en el área de ciencias naturales y exactas, específicamente la disciplina de la conservación y manejo de recursos naturales. En la tabla 1 se presenta el grado de los integrantes del cuerpo académico junto con su respectiva producción registrada en Scopus.

**Tabla 1.** Datos de los integrantes del CAEF analizado en el caso de estudio

Miembro	Grado	Número de publicaciones en colaboración	Número de publicaciones individuales
Investigador 1	Doctor	16	0
Investigador 2	Doctor	16	0
Investigador 3	Doctor	6	0

Fuente: Elaboración propia

En la tabla 2 se presentan las publicaciones que se relacionan con el CAEF, posterior a la depuración de las publicaciones en colaboración externas. En este caso, no hay publicaciones individuales para omitir.

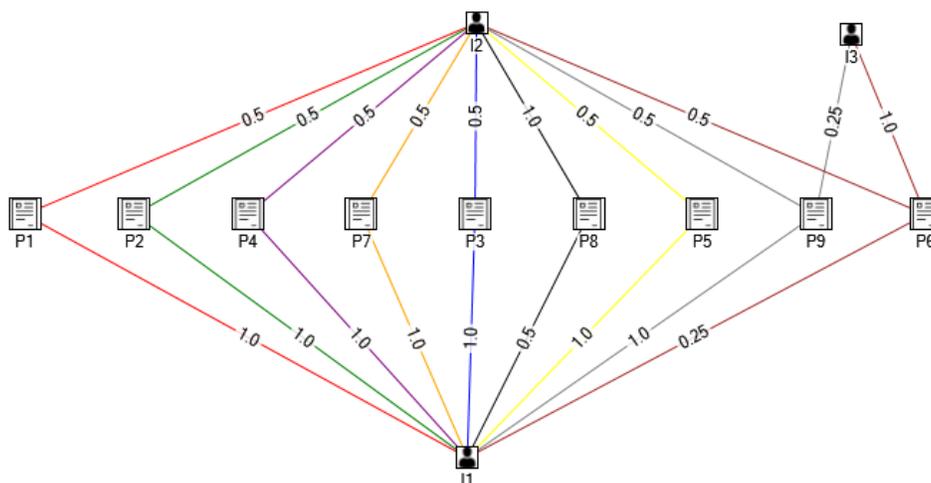
**Tabla 2.** Publicaciones pertenecientes al CAEF con base en Scopus

Publicaciones	Autores (en orden)
Publicación 1 (Artículo)	Investigador 1 (1), Investigador 2 (0.5)
Publicación 2 (Artículo)	Investigador 1 (1), Investigador 2 (0.5)
Publicación 3 (Artículo)	Investigador 1 (1), Investigador 2 (0.5)
Publicación 4 (Artículo)	Investigador 1 (1), Investigador 2 (0.5)
Publicación 5 (Artículo)	Investigador 1 (1), Investigador 2 (0.5)
Publicación 6 (Artículo)	Investigador 3 (1), Investigador 2 (0.5), Investigador 1 (0.25)
Publicación 7 (Artículo)	Investigador 1 (1), Investigador 2 (0.5)
Publicación 8 (Artículo en prensa)	Investigador 2 (1), Investigador 1 (0.5)
Publicación 9 (Artículo)	Investigador 1 (1), Investigador 2 (0.5), Investigador 3 (0.25)

Fuente: Elaboración propia

Utilizando como fundamento los resultados obtenidos en la tabla 2, se creó una tabla para ser procesada por NodeXL y construir el grafo de colaboración, el cual se expone en la figura 3.

**Figura 3.** Grafo de colaboración entre los miembros del CAEF



Elaboración propia

### Resultados de la evaluación del CAEF

A partir de la información de la tabla 2 y del grafo de la figura 2, se efectúa el análisis de la producción científica del CAEF de acuerdo con el tipo de publicación (tabla 3) y el nivel de la colaboración entre los miembros, así como su aporte en las publicaciones que es reflejado en el grafo obtenido. Para realizar la evaluación de la colaboración entre los miembros del CAEF y su aporte en las publicaciones, se consideró la participación de los integrantes, además del grado (exponiendo la suma total del grado de colaboración en cada publicación) y el porcentaje de colaboración en las publicaciones del cuerpo académico de cada uno de ellos (tabla 4).

**Tabla 3.** Evaluación de la producción científica del CAEF analizado

Tipo de publicación	Total
Artículo en prensa	1
Artículo	8
Producción total	9

Fuente: Elaboración propia

**Tabla 4.** Evaluación de la colaboración entre los miembros del CAEF y su aporte en las publicaciones con base en el grafo obtenido

Participación de los integrantes (grado de cada vértice) (en orden descendente)	$g(I1) = 9$ (El Investigador 1 participa en nueve publicaciones) $g(I2) = 9$ (El Investigador 2 participa en nueve publicaciones) $g(I3) = 2$ (El Investigador 3 participa en dos publicaciones)
Grado de colaboración total de los investigadores en las publicaciones del cuerpo académico ( $GCT$ ) (en orden descendente)	Investigador 1: $1 + 1 + 1 + 1 + 1 + 0.25 + 1 + 0.5 + 1 = 7.75$ Investigador 2: $0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 + 1 + 0.5 = 5$ Investigador 3: $1 + 0.25 = 1.25$
Porcentaje de contribución en las publicaciones del cuerpo académico ( $PC$ ) (en orden descendente)	Investigador 1: $\frac{100 \cdot 7.75}{9} = 86.11 \%$ Investigador 2: $\frac{100 \cdot 5}{9} = 55.5 \%$ Investigador 3: $\frac{100 \cdot 1.25}{9} = 13.88 \%$

Fuente: Elaboración propia

La colaboración más común que se pueden inferir del grafo de la figura 2 y de la información de la tabla 4 es entre el Investigador 1 y el Investigador 2 con nueve publicaciones en común.

## Cuerpo académico en consolidación

El CAEC analizado se enfoca en la investigación y el desarrollo tecnológico orientado a la generación de conocimiento para la integración de aplicaciones en sistemas de cómputo distribuido o paralelo, con un impacto en el área de ingeniería y tecnología en la disciplina de ciencias de la computación. En la tabla 5 se exhibe el grado de sus integrantes con su respectiva producción científica registrada en Scopus.

**Tabla 5.** Datos de los integrantes del CAEC analizado en el caso de estudio

Miembro	Grado	Número de publicaciones en colaboración	Número de publicaciones individuales
Investigador 4	Doctor	3	0
Investigador 5	Maestro	16	0
Investigador 6	Maestro	10	0
Investigador 7	Doctor	13	0
Investigador 8	Doctor	7	0

Fuente: Elaboración propia

Una vez que se realizó la depuración de las publicaciones registradas en Scopus que tienen por lo menos dos miembros del CAEC, se obtienen las publicaciones y los autores correspondientes (tabla 6).

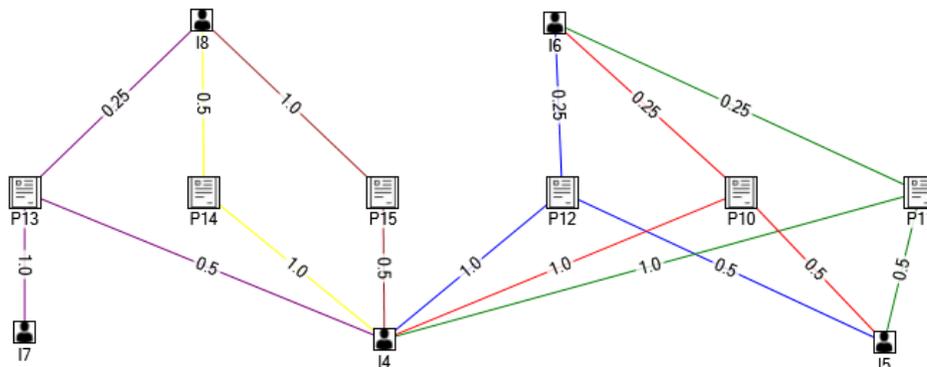
**Tabla 6.** Publicaciones pertenecientes al CAEC con base en Scopus

Publicaciones	Autores (en orden)
Publicación 10 (Artículo de conferencia)	Investigador 4 (1), Investigador 5 (0.5), Investigador 6 (0.25) y 2 externos
Publicación 11 (Artículo de conferencia)	Investigador 4 (1), Investigador 5 (0.5), Investigador 6 (0.25) y 2 externos
Publicación 12 (Artículo de conferencia)	Investigador 4 (1), Investigador 5 (0.5), Investigador 6 (0.25) y 2 externos
Publicación 13 (Artículo)	Investigador 7 (1), Investigador 4 (0.5), Investigador 8 (0.25) y 1 externo
Publicación 14 (Artículo)	Investigador 4 (1), Investigador 8 (0.5) y 2 externos
Publicación 15 (Artículo de conferencia)	Investigador 8 (1), Investigador 4 (0.5) y 1 externo

Fuente: Elaboración propia

Utilizando los resultados obtenidos en la tabla 6, se creó una tabla para ser procesada por NodeXL y construir el grafo de colaboración, el cual se expone en la figura 4.

**Figura 4.** Grafo de colaboración entre los miembros del CAEC



Fuente: Elaboración propia

### Resultados de la evaluación del CAEC

Tomando como referencia la tabla 6 y el grafo de la figura 4, se evalúa si el CAEC cumple con los requisitos establecidos por el Prodep relacionados con la producción científica (tabla 7). El análisis de la producción científica del CAEC analizado por tipo de publicación (tabla 8) y

la evaluación de la colaboración entre los miembros y su aporte en las publicaciones reflejado en el grafo obtenido se presentan en la tabla 9.

**Tabla 7.** Evaluación de las características generales del CAEC con base en lo establecido por Prodep

<b>Característica</b>	<b>Cumple</b>
Cuentan con productos académicos con reconocimiento debido a su buena calidad, derivados de las líneas de investigación consolidadas.	Sí
Participan de forma conjunta en líneas de investigación o aplicación innovadora del conocimiento.	Sí

Fuente: Elaboración propia

**Tabla 8.** Evaluación de la producción científica del CAEC analizado

<b>Tipo de publicación</b>	<b>Total</b>
Artículo de conferencia	4
Artículo	2
Producción total	6

Fuente: Elaboración propia

**Tabla 9.** Evaluación de la colaboración entre los miembros del CAEC y su aporte en las publicaciones con base en el grafo obtenido

Participación de los integrantes (grado de cada vértice) (en orden descendente)	$g(I4) = 6$ (El Investigador 4 participa en seis publicaciones) $g(I5) = 3$ (El Investigador 5 participa en tres publicaciones) $g(I6) = 3$ (El Investigador 6 participa en tres publicaciones) $g(I8) = 3$ (El investigador 8 participa en tres publicaciones) $g(I7) = 1$ (El investigador 7 participa en una publicación)
Grado de colaboración total de los investigadores en las publicaciones del cuerpo académico ( <i>GCT</i> ) (en orden descendente)	Investigador 4: $1 + 1 + 1 + 0.5 + 1 + 0.5 = 5$ Investigador 8: $0.25 + 0.5 + 1 = 1.75$ Investigador 5: $0.5 + 0.5 + 0.5 = 1.5$ Investigador 7: 1 Investigador 6: $0.25 + 0.25 + 0.25 = 0.75$
Porcentaje de contribución en las publicaciones del cuerpo académico ( <i>PC</i> ) (en orden descendente)	Investigador 4: $\frac{100 \cdot 5}{6} = 83.3 \%$ Investigador 8: $\frac{100 \cdot 1.75}{6} = 29.1 \%$ Investigador 5: $\frac{100 \cdot 1.5}{6} = 25 \%$ Investigador 7: $\frac{100 \cdot 1}{6} = 16.6 \%$ Investigador 6: $\frac{100 \cdot 0.75}{6} = 12.5 \%$

Fuente: Elaboración propia

Las colaboraciones más representativas observadas en el grafo (figura 4) y en la tabla 9 son:

- Grupo 1: Investigador 4, Investigador 5, Investigador 6 (tres publicaciones).
- Grupo 2: Investigador 4, Investigador 8 (tres publicaciones).

### Cuerpo académico consolidado

El CAC analizado se enfoca en el estudio de los parásitos, bacterias y virus que son causa de enfermedad en el ser humano, por lo que su trabajo tiene un mayor impacto en el área de ciencias de la salud, en la disciplina de biomedicina. En la tabla 10 se da a conocer el grado de sus integrantes, así como su producción científica registrada en Scopus.

**Tabla 10.** Datos de los integrantes del CAC analizado en el caso de estudio

Miembro	Grado	Número de publicaciones en colaboración	Número de publicaciones individuales
Investigador 9	Maestro	37	0
Investigador 10	Doctor	60	0
Investigador 11	Doctor	18	0
Investigador 12	Doctor	9	0
Investigador 13	Doctor	12	0
Investigador 14	Doctor	19	0
Investigador 15	Doctor	7	0
Investigador 16	Doctor	32	0

Fuente: Elaboración propia

Una vez que se realizó la depuración de las publicaciones registradas en Scopus que tienen por lo menos dos miembros del CAC, se obtienen las publicaciones y los autores correspondientes (tabla 11).

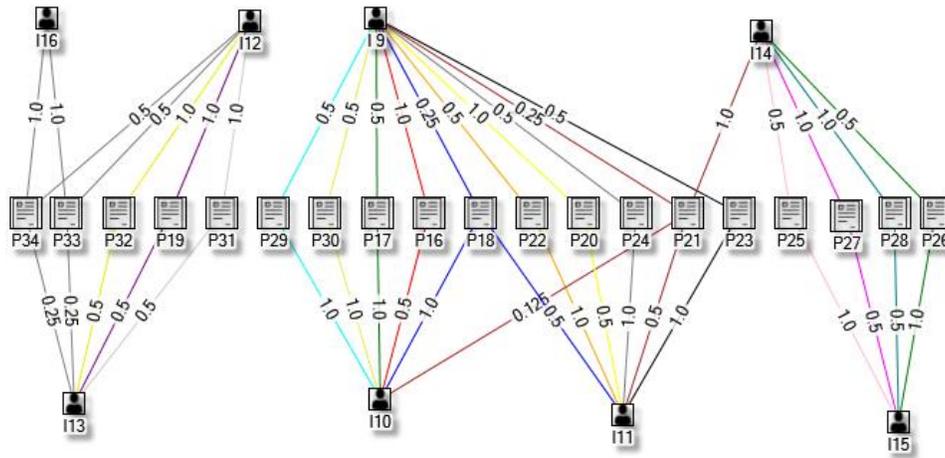
**Tabla 11.** Publicaciones pertenecientes al CAC con base en Scopus

<b>Publicaciones</b>	<b>Autores (en orden)</b>
Publicación 16 (Artículo)	Investigador 9 (1), Investigador 10 (0.5) y 11 externos
Publicación 17 (Artículo)	Investigador 10 (1), Investigador 9 (0.5) y 5 externos
Publicación 18 (Artículo)	Investigador 10 (1), Investigador 11 (0.5), Investigador 9 (0.25) y 7 externos
Publicación 19 (Nota)	Investigador 12 (1), Investigador 13 (0.5)
Publicación 20 (Artículo)	Investigador 9 (1), Investigador 11 (0.5) y 10 externos
Publicación 21 (Artículo)	Investigador 14 (1), Investigador 11 (0.5), Investigador 9 (0.25), Investigador 10 (0.125) y 9 externos
Publicación 22 (Artículo)	Investigador 11 (1), Investigador 9 (0.5) y 7 externos
Publicación 23 (Artículo)	Investigador 11 (1), Investigador 9 (0.5)
Publicación 24 (Artículo)	Investigador 11 (1), Investigador 9 (0.5)
Publicación 25 (Artículo)	Investigador 15 (1), investigador 14 (0.5)
Publicación 26 (Artículo)	Investigador 15 (1), investigador 14 (0.5)
Publicación 27 (Artículo)	Investigador 14 (1), investigador 15 (0.5)
Publicación 28 (Artículo)	Investigador 14 (1), investigador 15 (0.5)
Publicación 29 (Artículo)	Investigador 10 (1), Investigador 9 (0.5)
Publicación 30 (Artículo)	Investigador 10 (1), Investigador 9 (0.5)
Publicación 31 (Artículo)	Investigador 12 (1), Investigador 13 (0.5)
Publicación 32 (Artículo)	Investigador 12 (1), Investigador 13 (0.5)
Publicación 33 (Artículo)	Investigador 16 (1), Investigador 12 (0.5), Investigador 13 (0.25)
Publicación 34 (Artículo)	Investigador 16 (1), Investigador 12 (0.5), Investigador 13 (0.25)

Fuente: Elaboración propia

Utilizando los resultados obtenidos en la tabla 11, se creó una tabla para ser procesada por NodeXL y construir el grafo de colaboración, el cual se expone en la figura 5.

**Figura 5.** Grafo de colaboración entre los miembros del CAC



Fuente: Elaboración propia

### Resultados de la evaluación del CAC

Tomando como referencia la tabla 11 y el grafo de la figura 5, y de igual manera que en la evaluación anterior, se verifica si el CAC cumple con los requisitos establecidos por el Prodep relacionados con la producción científica (tabla 12). El análisis de la producción científica del CAC analizado por tipo de publicación (tabla 13) y la evaluación de la colaboración entre los miembros y su aporte en las publicaciones reflejado en el grafo obtenido se presentan en la tabla 14.

**Tabla 12.** Evaluación de las características generales del CAC con base en lo establecido por Prodep

Característica	Cumple
Alto compromiso con la institución por medio de la colaboración y producción científica y académica.	Sí
Demuestran una intensa actividad académica manifiesta en congresos, seminarios, mesas, talleres, etc., de forma regular y frecuente.	Sí
Sostienen una intensa participación en redes de intercambio académico.	Sí

Fuente: Elaboración propia

**Tabla 13.** Evaluación de la producción científica del CAC analizado

<b>Tipo de publicación</b>	<b>Total</b>
Artículo	18
Nota	1
Producción total	19

Fuente: Elaboración propia

**Tabla 14.** Evaluación de la colaboración entre los miembros del CAC y su aporte en las publicaciones con base en el grafo obtenido

Participación de los integrantes (grado de cada vértice) (en orden descendente)	$g(I9) = 10$ (El Investigador 9 participa en 10 publicaciones) $g(I10) = 6$ (El Investigador 10 participa en seis publicaciones) $g(I11) = 6$ (El investigador 11 participa en seis publicaciones) $g(I12) = 5$ (El investigador 12 participa en cinco publicaciones) $g(I13) = 5$ (El investigador 13 participa en cinco publicaciones) $g(I14) = 5$ (El investigador 14 participa en cinco publicaciones) $g(I15) = 4$ (El investigador 15 participa en cuatro publicaciones) $g(I16) = 2$ (El investigador 16 participa en dos publicaciones)
Grado de colaboración total de los investigadores en las publicaciones del cuerpo académico ( <i>GCT</i> ) (en orden descendente)	Investigador 9: $1 + 0.5 + 0.25 + 1 + 0.25 + 0.5 + 0.5 + 0.5 + 0.5 + 0.5 = 5.5$ Investigador 10: $0.5 + 1 + 1 + 0.125 + 1 + 1 = 4.625$ Investigador 11: $0.5 + 0.5 + 0.5 + 1 + 1 + 1 = 4.5$ Investigador 12: $1 + 1 + 1 + 0.5 + 0.5 = 4$ Investigador 14: $1 + 0.5 + 0.5 + 1 + 1 = 4$ Investigador 15: $1 + 1 + 0.5 + 0.5 = 3$ Investigador 13: $0.5 + 0.5 + 0.5 + 0.25 + 0.25 = 2$ Investigador 16: $1 + 1 = 2$
Porcentaje de contribución en las publicaciones del cuerpo académico ( <i>PC</i> ) (en orden descendente)	Investigador 9: $\frac{100 \cdot 5.5}{19} = 28.94 \%$ Investigador 10: $\frac{100 \cdot 4.625}{19} = 24.34 \%$ Investigador 11: $\frac{100 \cdot 4.5}{19} = 23.68 \%$ Investigador 12: $\frac{100 \cdot 4}{19} = 21.05 \%$ Investigador 14: $\frac{100 \cdot 4}{19} = 21.05 \%$ Investigador 15: $\frac{100 \cdot 3}{19} = 15.78 \%$ Investigador 13: $\frac{100 \cdot 2}{19} = 10.52 \%$ Investigador 16: $\frac{100 \cdot 2}{19} = 10.52 \%$

Fuente: Elaboración propia

Las colaboraciones más representativas observadas en el grafo (figura 5) y en la tabla 14 son:

- Grupo 1: Investigador 9, Investigador 10 (seis publicaciones).
- Grupo 2: Investigador 9, Investigador 11 (seis publicaciones).
- Grupo 3: Investigador 12, Investigador 13 (cinco publicaciones).
- Grupo 4: Investigador 14, Investigador 15 (cuatro publicaciones).

## Discusión

Para recopilar la producción científica fue necesario realizar una búsqueda en Scopus con el nombre y apellido de cada miembro. El número de consultas realizadas por cada cuerpo académico depende del número de miembros; cada consulta se realizó para cada uno de los perfiles de Scopus de los integrantes del cuerpo académico en cuestión. Sin embargo, existen investigadores que tienen más de un perfil en Scopus, debido a que han registrado su nombre en formatos diferentes. Por ejemplo, en algunas publicaciones se encuentra registrado un nombre y solo el apellido paterno, en otras la inicial de su nombre y los dos apellidos, así como otros formatos. Entre mayor número de perfiles de los miembros del cuerpo académico existan, mayor es el número de consultas porque, en teoría, cada perfil es un investigador diferente, lo cual no refleja la realidad. Entre más consultas, mayor es el tiempo de la recopilación de datos, el cual se podría evitar si cada investigador tuviera un solo formato para escribir su nombre y, así, se generara un único perfil de Scopus. Incluso un usuario puede no considerar los perfiles adicionales, sino que únicamente el que tiene registrado un mayor número de publicaciones, ya que puede existir la creencia de que es el más completo, cuando la realidad es que cada perfil de Scopus tiene publicaciones diferentes.

Por otra parte, la teoría de grafos, a pesar de ser una técnica aparentemente sencilla para la visualización de datos, fue útil debido a los conceptos en los que se fundamenta, e incluso hizo equivalentes dichos conceptos con aspectos del tipo de cuerpo académico. Aunado a ello, a simple vista se pudo observar que el CAEF tiene un mayor número de publicaciones que el CAEC. Pero esto puede ser dependiente del año de registro del cuerpo académico y de la línea de investigación a la que corresponde. El CAC fue el que tuvo un mayor número de publicaciones, lo cual permite apreciar la colaboración constante y el impacto de las publicaciones en el ámbito científico. Por otra parte, realizar la metodología propuesta de forma manual es menos eficiente a comparación

de una solución tecnológica que pueda aprovechar el servicio web (API) de Scopus, y generar una tabla para que, a partir de ella, sea posible obtener el grafo de manera automática. Dicha tabla debe cumplir con el formato aceptado por la herramienta a utilizar.

## Conclusiones

La teoría de grafos puede utilizarse para analizar y representar datos, así como para la toma de decisiones basada en sus resultados, que depende de la institución a la que pertenecen los cuerpos académicos o alguna que los evalúa. En este caso, resultó útil para conocer la producción de tres tipos de cuerpos académicos siguiendo una metodología sencilla, basada en la extracción del conocimiento. Se observó que los tres cuerpos académicos cumplen con lo que Prodep les solicita, además de localizar a los elementos fuertes o centrales de los cuerpos académicos respecto a su participación y nivel de contribución, a los elementos que menor participación tienen e incluso los que tienen un nivel de contribución considerablemente alejado del más alto pero que aportan en la mayoría de las publicaciones. Como trabajo futuro, se plantea realizar una solución tecnológica que permita realizar de forma automática el proceso manual que se realizó en este caso de estudio, y con ello hacer el trabajo más eficiente, en menor tiempo y con mayor precisión, e incluso obtener datos adicionales y relevantes con base en *software* especializado en grafos como Gephi.

## Agradecimientos

Este trabajo ha sido desarrollado gracias al apoyo del Consejo Nacional de Ciencia y Tecnología (Conacyt) a través de la beca con número (CVU/Becario): 853088/630948.

## Referencias

- Álvarez, M. y Parra, J. (2013). *Teoría de grafos*. Bío-Bío, Chile: Universidad del Bío-Bío.
- Aparicio, M. and Costa, C. J. (2015). Data visualization. *Communication Design Quarterly Review*, 3(1), 7–11. Retrieved from <https://doi.org/10.1145/2721882.2721883>.
- California State University System [CSU]. (2018). Merlot. Retrieved from <https://www.merlot.org/merlot/index.htm>.
- Chen, C., Härdle, W.K. and Unwin, A. (eds.) (2007). *Handbook of Data Visualization*. Alemania: Springer Science & Business Media. Retrieved from <https://doi.org/10.1007/978-3-540-33037-0>.
- Cornell University. (2018). arXiv.org e-Print archive. Retrieved from <https://arxiv.org/>.
- dblp team [dblp]. (2018). dblp: computer science bibliography. Retrieved from <https://dblp.uni-trier.de/>.
- Dirección General de Educación Superior Universitaria [Dgesu]. (2018). Dirección General Educación Superior Universitaria | Inicio. México: Dirección General de Educación Superior Universitaria. Recuperado de <http://www.dgesu.ses.sep.gob.mx/PRODEP.htm>.
- Elsevier. (2017). Elsevier Scopus APIs. Retrieved from [https://dev.elsevier.com/sc\\_apis.html](https://dev.elsevier.com/sc_apis.html).
- Elsevier. (2018). Scopus, la mayor base de datos de bibliografía revisada por pares. Recuperado de <https://www.elsevier.com/es-mx/solutions/scopus>.
- Gephi Consortium. (2018). Gephi - The Open Graph Viz Platform. Retrieved from <https://gephi.org/>.
- Giatsidis, C., Berberich, K., Thilikos, D. M. and Vazirgiannis, M. (2012). Visual Exploration of Collaboration Networks based on Graph Degeneracy. Paper presented at the 18<sup>th</sup> ACM SIGKDD International Conference on Knowledge Discovery and Data Mining. Beijing, August 12-16, 2012.
- González, G. y Gómez, J. (2014). La colaboración científica: principales líneas de investigación y retos de futuro. *Revista Española de Documentación Científica*, 37(4), 1-15.
- Google. (2018). About Fusion Tables - Fusion Tables Help. Retrieved from <https://support.google.com/fusiontables/answer/2571232?hl=en>.
- Grus, J. (2015). *Data Science from Scratch* (1<sup>st</sup> ed.). United States: O'Reilly Media, Inc.
- Huamaní, Ch., González, A. G., Curioso, W. H. y Pacheco, J. (2012). Redes de colaboración y producción científica sudamericana en medicina clínica, ISI Current Contents 2000-2009. *Revista médica de Chile*, 140(4), 466-475.

- López, S. (2010). Cuerpos Académicos: Factores de Integración y Producción de Conocimiento. *Revista de la Educación Superior*, 39(155), 7–26.
- Newman, M. E. J. (2001a). Scientific collaboration networks. I. Network construction and fundamental results. *Physical Review E*, 64(1).
- Newman, M. E. J. (2001b). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E*, 64(1).
- Prieto, M., Menéndez, V. y Zapata, A. (2010). Data Mining Learning Objects. En Romero, C., Ventura, S. y Pechenizkiy, M. (coords.), *Handbook of Educational Data Mining* (pp. 315-342). United States: CRC Press.
- Sallán, J. M., Suñé, A., Fernández V. y Fonollosa, J. B. (2002). *Métodos cuantitativos en organización industrial I*. España: Edicions UPC.
- Scopus. (2019). Search for an author profile. Retrieved from <https://www.scopus.com/freelookup/form/author.uri>.
- Staudt, C. L. (2011). *Analysis of Scientific Collaboration Networks: Social Factors, Evolution, and Topical Clustering*. (diploma thesis). Karlsruhe Institute of Technology, Germany.
- Social Media Research Foundation. (2018). NodeXL: Network Overview, Discovery and Exploration for Excel. Retrieved from <https://archive.codeplex.com/?p=nodexl>.
- Trudeau, R. J. (2017). *Introduction to Graph Theory*. Stanford, United States: Stanford University Press.

<b>Rol de Contribución</b>	<b>Autor (es)</b>
Conceptualización	Víctor Hugo Menéndez Domínguez
Metodología	María Enriqueta Castellanos Bolaños
Software	Jared David Tadeo Guerrero Sosa
Validación	María Enriqueta Castellanos Bolaños
Análisis Formal	Jared David Tadeo Guerrero Sosa
Investigación	Esmeralda Zurita Gallegos
Recursos	María Enriqueta Castellanos Bolaños
Curación de datos	Esmeralda Zurita Gallegos
Escritura - Preparación del borrador original	Jared David Tadeo Guerrero Sosa
Escritura - Revisión y edición	Víctor Hugo Menéndez Domínguez
Visualización	María Enriqueta Castellanos Bolaños
Supervisión	Víctor Hugo Menéndez Domínguez
Administración de Proyectos	Víctor Hugo Menéndez Domínguez
Adquisición de fondos	Víctor Hugo Menéndez Domínguez