

Comportamiento bibliométrico de la producción científica sobre trayectorias académicas

Exequiel Fontans-Álvarez¹
Natalia Aguirre-Ligüera²
José Fernández³

Resumen

El objetivo de este trabajo es caracterizar el dominio de la producción sobre trayectorias académicas en la base Web of Science mediante el uso combinado de análisis de estudios métricos de la información y análisis de redes sociales. El estudio de las trayectorias académicas de los investigadores da cuenta de los procesos involucrados en la generación de conocimiento: dónde se formó el investigador, con quién trabajó, a qué instituciones se adscribió, cómo adquirió sus habilidades y cómo desarrolló sus competencias y preferencias, entre otros elementos que resultan de interés para comprender su desempeño como investigador. Los principales resultados muestran dos períodos en la producción científica sobre la temática, con un crecimiento exponencial entre ambos. Se evidencia dispersión en las revistas que publican estos trabajos; sin embargo, un tercio de estos se publicaron en 35 títulos de revistas (6% del total de revistas). No se visualiza una comunidad de autores consolidada produciendo sobre el tema; la temática de los documentos se concentra en torno a tres tópicos: estudios de género, educación superior y ciencia, y otros temas transversales a estos. Se destaca que la aplicación de tres técnicas para el análisis otorga mayor robustez al trabajo.

Palabras clave: Trayectorias académicas; Análisis de *curriculum vitae*; Estudios métricos de la información.

Bibliometric behavior of scientific production on academic trajectories

Abstract

The objective of this work is to characterize the domain of the production on academic trajectories in the Web of Science base through the combined use of the metric information studies and social network analysis. The study of the academic

¹ Ayudante en el Departamento de Tratamiento y Transferencia de la Información, Instituto de Información de la Facultad de Información y Comunicación, Universidad de la República (Uruguay). Correo electrónico: exequiel.fontans@fic.edu.uy

² Profesora adjunta en el Departamento de Tratamiento y Transferencia de la Información, Instituto de Información de la Facultad de Información y Comunicación, Universidad de la República (Uruguay). Correo electrónico: natalia.aguirre@fic.edu.uy

³ Profesor agregado en el Departamento de Epistemología, Metodología e Historia, Instituto de Comunicación e Instituto de Información de la Facultad de Información y Comunicación, Universidad de la República (Uruguay). Correo electrónico: jose.fernandez@fic.edu.uy

trajectories of the researchers gives an account of the processes involved in the generation of knowledge: where the researcher was trained, with whom he worked, to which institutions he was assigned, how he acquired his skills, how he developed his skills and interests, among other elements, which are of interest to understand their research performance. The main results show two periods in the scientific production on the subject, with an exponential growth between both. Dispersion is evidenced in the journals that publish these works; however, a third of these were published in 35 journal titles (6 % of the total number of journals). A consolidated community of authors is not visualized producing on the subject; the subject of the documents is centered around three topics: gender studies, higher education and science and other topics that cross them. It is emphasized that the application of three techniques for the thematic analysis gives greater robustness to the work.

Keywords: Academic trajectories; Analysis of curriculum vitae; Metric information studies.

1. Introducción

Con el advenimiento de la sociedad de la información y el conocimiento, ha cobrado cada vez más relevancia el rol de las actividades científicas como motor del desarrollo de los países. Desde que ha sido necesaria la sistematización de las evaluaciones sobre la actividad científica, han surgido manuales que compilan y pretenden estandarizar indicadores consensuados por la comunidad internacional. De acuerdo con Ernesto Spinak:

Podemos considerar a la ciencia como un sistema de producción de información, en particular información en la forma de publicaciones, considerando publicación a cualquier «información registrada en formatos permanentes y disponibles para el uso común». Desde este punto de vista entonces, la ciencia puede verse como una empresa con insumos y resultados. La medición de esas dos categorías —insumos y resultados— son la base de los indicadores científicos (1998, p. 141).

En el mismo sentido, Rosa Sancho sostiene que «el proceso científico se puede considerar análogo a los modelos económicos coste-beneficio o inversión-resultado (“input-output”), susceptible, por tanto, de ser cuantificado» (1990, p. 79).

Esta metodología de medición de insumos y resultados es la que ha primado generalmente en los sistemas de evaluación y promoción de las actividades científicas. Sin embargo, algunos investigadores proponen un cambio de paradigma en la evaluación de la actividad científica, que implica moverse desde el paradigma del *output* o productividad hacia el paradigma de las capacidades (Bozeman, Dietz y Gaughan, 2001; Cañibano y Bozeman, 2009). Esta nueva aproximación se basa en la corriente de sociología de la ciencia e incorpora los aportes del capital humano y el capital social. En este sentido, Hernán Jaramillo, Carolina Lopera y María Carolina Albán (2008) plantean:

El capital humano y capital social no incluye solamente el conocimiento codificado, sino también el conocimiento tácito del investigador, sus habilidades y el «know-how». De este modo, la existencia de flujos de capital humano y social entre sectores y organizaciones son una variable determinante para comprender los procesos de difusión, transferencia de conocimiento y creación de redes sociales a partir de las carreras de los científicos (p. 2).

Esta perspectiva permite abrir el marco de análisis e integrar dimensiones que tienen que ver con las trayectorias académicas y potencialidades de reproducción y desarrollo del sistema científico.

Uno de los principales insumos para este tipo de análisis es el *curriculum vitae* (CV). Carolina Cañibano y Barry Bozeman (2009), en un artículo de revisión, identifican tres líneas de investigación en el uso de CV: las trayectorias académicas, la movilidad (geográfica y entre instituciones) y el mapeo de la capacidad colectiva. En este trabajo nos centraremos en la primera línea. El estudio de las trayectorias de las carreras académicas de los investigadores redundará en el análisis de los procesos que intervienen en la constitución de un investigador como tal, su formación, la selección de temas de trabajo, las

instituciones en las que desempeñó sus funciones y la participación en equipos, es decir, un conjunto de elementos que intervienen en la conformación de la experticia del investigador, en sus modos de hacer y en sus capacidades. Estos elementos cobran relevancia para explicar el comportamiento investigador del científico y comprender el sistema científico.

El objetivo de este trabajo es identificar el comportamiento bibliométrico de la producción científica sobre trayectorias académicas mediante la aplicación de indicadores y técnicas tales como los modelos matemáticos de Samuel Bradford y de George Zipf, el índice de estabilidad de Derek Price y el análisis de redes sociales (ARS), con el propósito de revelar las principales regularidades cuantitativas del referido tema de estudio.

2. Fuentes y metodología

2.1. Fuentes

La fuente utilizada para este estudio bibliométrico fue la plataforma Web of Science⁴ (WOS) de Clarivate Analytics, una de las más importantes bases de datos en línea que contiene información bibliográfica, científica, internacional y multidisciplinaria. Está conformada por un conjunto de servicios de suscripción, entre los que destaca la Web of Science Core Collection. Esta colección principal está integrada por tres bases de datos especializadas temáticamente en distintas áreas de conocimiento: Science Citation Index Expanded, que colecciona 8864 títulos de revista de las ciencias básicas, desde 1900 al presente; Social Sciences Citation Index, especializada en ciencias sociales, incluye artículos de 3245 revistas, desde 1900 al presente, y Arts & Humanities Citation Index, que indexa 1778 títulos de publicaciones periódicas de humanidades y artes, desde 1975 al presente. Además, brinda información bibliográfica sobre actas de congresos y está incorporando libros. Su registro más reciente es el Emerging Sources Citation Index, que incluye 5121 títulos de revistas, desde 2015 hasta el presente, que aún no ingresaron a la colección principal, pero que tienen una relevancia creciente. La base de datos de la WOS es una de las más usadas para estudios bibliométricos porque incluye el recuento de citas obtenidas por las publicaciones. Por su uso extendido, también ha sido evaluada desde distintos puntos de vista en la literatura del área. Algunos ejemplos: Ilva Gavel y Lars Iselid (2008) proporcionan datos cuantitativos sobre la cobertura de títulos de revista; Fernando Beigel y Maximiliano Salatino (2015) analizan el impacto que tiene la WOS en lo que ellos llaman el sistema académico mundial; Carmen López-Illescas, Félix Moya-Anegón y Henk Moed (2008) comparan la cobertura y el impacto de un área temática específica, oncología, en WOS y Scopus, y Juan Gorraiz y Christian Schloegl (2008) hacen lo mismo para el área farmacéutica, entre tantos otros trabajos.

Sin embargo, la utilización de la WOS para los estudios bibliométricos plantea algunos problemas vinculados a su cobertura temática, geográfica y lingüística. Estos sesgos privilegian a las ciencias básicas sobre las ciencias sociales —y más aún respecto a las humanidades—, las revistas de países como

⁴ Sitio web: <https://clarivate.com/products/web-of-science/>

EE. UU. y Reino Unido (entre otros países centrales) y las publicaciones en idioma inglés (Torres-Salinas, Delgado López-Cózar y Jiménez-Contreras, 2009; Zhou, Thijs y Glänzel, 2009, entre otros).

Paralelamente, la WOS posee como ventaja innegable que las revistas incluidas en sus bases de datos cumplen con criterios estrictos que garantizan ciertos estándares de calidad formal y de contenido. Las revistas son evaluadas en forma permanente, lo que determina su mantenimiento, exclusión o ingreso en los índices de citas. Los criterios son valorados por el Comité Editorial en forma combinada, no aislada, y se agrupan en cuatro apartados: normas básicas de publicación, contenido editorial, diversidad internacional y análisis de citas (Testa, 2016).

2.2. Estrategia de búsqueda

Es importante aclarar que para este trabajo usaremos el término *trayectorias académicas* por ser el de uso más corriente en español, aunque para la consulta en la colección principal de la WOS se construyó una estrategia de búsqueda con los términos usados en inglés según el siguiente detalle:

Tema: (*scientific careers*) OR Tema: (*academic careers*) OR Tema: (*careers in science*) OR Tema: (*careers in academic*) OR Tema: (*academic path*) OR Tema: (*scientific path*)

Los resultados se refinan por:

[excluyendo] Tipos de documento: (note or book chapter or biographical item or item about an individual or editorial material or letter or reprint or discussion or book review or news item)

Los resultados se descargan en una base de datos *ad hoc*. Se utilizan los campos: AU (autor), SO (título de la fuente), PY (año de publicación), JI (título de la fuente en formato ISO), ID (Key Word Plus) y WC (categorías WOS).

2.3. Procesamiento

Los datos son procesados mediante el uso combinado de técnicas bibliométricas y de ARS. Los indicadores bibliométricos obtenidos son los siguientes:

Indicadores de la actividad científica:

- Número de documentos y distribución temporal.
- Revistas de publicación, siguiendo la metodología propuesta por el modelo teórico de la dispersión de las publicaciones de Bradford (1948).
- Número de documentos por autor.

Indicadores de asociaciones temáticas:

- Análisis de palabras comunes o coocurrencia de palabras.

Para el estudio de las temáticas se optó por utilizar el campo ID, que contiene los Key Word Plus (KWP). Estos descriptores son asignados por Science Citation Index mediante un algoritmo que analiza los títulos de las referencias contenidas en los documentos (Cantos-Mateos *et al.*, 2014; Garfield, 1990;

Garfield y Sher, 1993). La ventaja sobre otros métodos como el de palabras clave asignadas por el autor es que son más específicos y aluden directamente al contenido conceptual de los documentos (Cantos-Mateos *et al.*, 2014). Sobre el conjunto de descriptores encontrados se aplicaron tres análisis: el modelo matemático de Zipf, la estabilidad de Price y el ARS, con distintos recortes para mejorar la visualización e interpretación de los resultados. El modelo matemático de Zipf se aplicó sobre todas las KWP. En la estabilidad de Price se consideraron solo las palabras presentes en cuatro años. Finalmente, para el ARS se consideraron las KWP presentes en cuatro años que, además, estuvieran en la primera zona de Zipf.

Para el cálculo del modelo matemático de Zipf (Zipf, 1949) se organizan los términos por zonas de acuerdo con su frecuencia de aparición. La primera zona la constituyen las palabras con una frecuencia igual o mayor a la raíz cuadrada del término más frecuente; lo mismo se aplica a las zonas subsiguientes. La última zona la constituyen los términos restantes (André *et al.*, 2014; Quoniam *et al.*, 2001).

En cuanto al análisis de la estabilidad de palabras entre períodos, fue propuesto por Price y Suha Gürsey (1975) para analizar la fugacidad y permanencia de los autores en un frente de investigación. Cobo, M. J. *et al.* (2012) adaptaron la metodología para el estudio de palabras clave. Este es el uso que se le da en el presente trabajo.

Finalmente, el ARS es un enfoque en las ciencias sociales que centra su atención en las relaciones entre actores, desplazándose del tradicional foco en los propios actores, por lo que es conocido también como análisis estructural (Freeman, 2012). No existe acuerdo en la literatura sobre si este enfoque constituye una técnica de abordaje o un nuevo paradigma dentro de las ciencias sociales (Molina González, 2001); lo que sí puede asegurarse es que constituye una herramienta cada vez más utilizada para el desarrollo de los estudios métricos de la información (Pinto y Gonzales-Aguilar, 2014).

Una red social es una representación gráfica (un grafo) compuesta por al menos dos componentes: los nodos o actores y las relaciones o vínculos existentes entre ellos (aristas). Los nodos pueden representar individuos, grupos, países, empresas, acontecimientos o palabras clave, entre otros. Las aristas representan los vínculos que se establecen entre los nodos según el abordaje del investigador, por ejemplo: amistad, afinidad, enemistad, coautoría o parentesco (Carrington, Scott y Wasserman, 2005; Molina González, 2001; Wasserman y Faust, 1994).

Los roles que desempeñan los actores en las redes pueden establecerse a través de la aplicación de la teoría de grafos, una rama de la matemática que aporta robustez al ARS. Según Stanley Wasserman y Katherine Faust (como se citó en Freeman, 2012), esta teoría tiene un vocabulario que puede ser utilizado para analizar muchas propiedades de las estructuras sociales, nos ofrece las operaciones matemáticas por las cuales esas propiedades pueden analizarse y medirse y nos permite probar teoremas sobre los grafos.

3. Resultados y discusión

3.1. Indicadores de la actividad científica

La estrategia de búsqueda lanzada recuperó 907 registros una vez aplicados los filtros. En la distribución temporal de los resultados pueden establecerse dos momentos: 1) período 1900-1989, en el que aparecen los primeros artículos sobre trayectorias académicas, y 2) período 1990-2016, integrado por tres décadas, en el que la producción crece exponencialmente, por lo que se mantiene la tendencia entre décadas (figura 1).

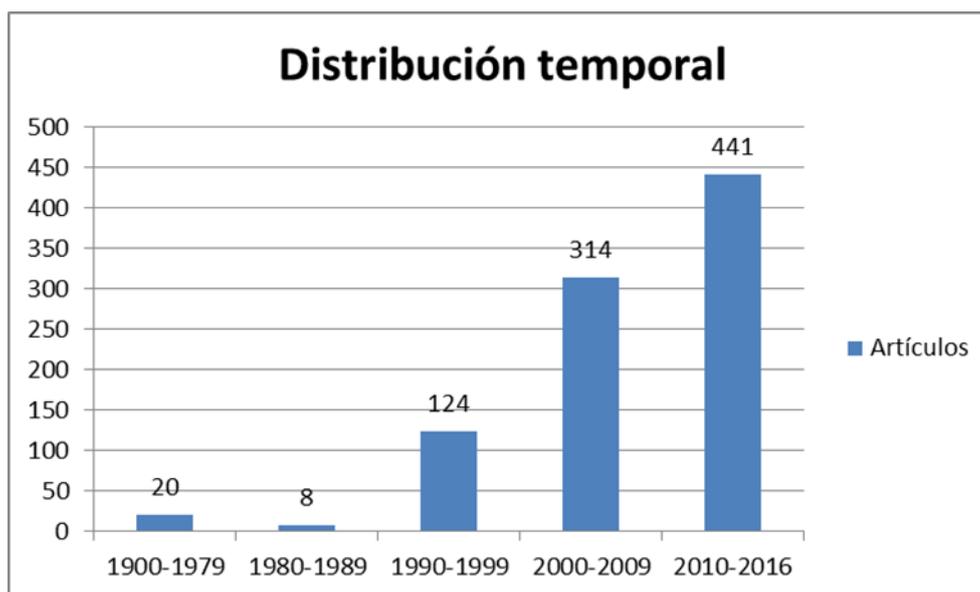


Figura 1. Distribución temporal de la producción
Fuente: WOS (elaboración propia)

Los 907 artículos del corpus se encuentran dispersos en 553 títulos de revista. A partir de la metodología de Bradford (1948) pueden identificarse tres zonas. El núcleo más productivo lo constituyen 35 títulos de revistas (6%) que concentran 277 artículos (31%) (tabla 1). Si se consideran la primera y la segunda zona en conjunto, el 22% de los títulos concentran la mitad de la producción (52%). Es importante señalar que en la primera zona quedaron comprendidas las revistas con cuatro artículos o más y en la segunda zona las revistas con tres y dos artículos, mientras que en la tercera zona todos los títulos de revista tienen un artículo.

Considerando las características de la distribución podría decirse que, por un lado, una parte considerable de la producción sobre la temática (31%) está concentrada en un pequeño núcleo de revistas (6%) y, por otro, la distribución tiene una larga cola, en la que cerca del 80% de los títulos solamente reportan un artículo sobre trayectorias académicas.

Paralelamente, poco más de la mitad (51%) de los artículos ($n=464$) se distribuyen en títulos de revistas clasificados en seis categorías temáticas (tabla 2).

En relación con la distribución de trabajos por autores, se conforma un corpus de 2370 autores, de los cuales 5 reportan 4 trabajos, 17 reportan 3 trabajos,

126 reportan 2 y los restantes 2222 son autores esporádicos, solo presentan un trabajo. Esta distribución podría interpretarse como la ausencia de una comunidad académica consolidada que esté trabajando en esta temática. Asimismo, la visualización que resulta de la colaboración autoral en el *software* Gephi muestra una red conformada por 550 comunidades e igual número de componentes conexos. Esto último refuerza la idea de una comunidad académica poco consolidada.

Tabla 1. Dispersión de la producción e identificación de los títulos de revista de la primera zona

Zona	Artículos			Títulos		
	Cantidad	Porcentaje	Acumulado	Cantidad	Porcentaje	Acumulado
1ra.	277	31%	31%	35	6%	6%
2da.	197	22%	52%	85	15%	22%
3ra.	433	48%	100%	433	78%	100%
	907			553		
Título			Títulos	Artículos	Total	
Acad. Med.			1	38	38	
J. Dent. Educ.			1	20	20	
Abstr. Pap. Am. Chem. Soc.			1	17	17	
High. Educ.			1	13	13	
Acad. Psych.			1	12	12	
Res. Policy; Scientometrics			2	11	22	
Int. J. Sci. Educ.			1	10	10	
Hum. Relat.			1	9	9	
Med. Educ.; Stud. High. Educ.			2	8	16	
Gend. Work. Organ.; J. Vocat. Behav.; PLoS One; Res. Evaluat.			4	7	28	
Career Dev. Int.; CBE-Life Sci. Educ.			2	6	12	
Am. J. Obstet. Gynecol.; Gend. Educ.; J. Invest. Med.; J. Am. Geriatr. Soc.; Proc. Natl. Acad. Sci. U. S. A.; Sci. Educ.; Soc. Stud. Sci.; Teach. Coll. Rec.			8	5	40	
Am. J. Med.; Best Pract. Res. Clin. Obstet. Gynaecol.; Isis; J. Res. Sci. Teach.; Mol. Biol. Cell; Neurology; Psychol. Sci.; Rev. High. Educ.; Sci. Technol. Hum. Values; Sociol. Educ.			10	4	40	
Total de títulos =	35			Total de artículos =	277	

Fuente: WOS (elaboración propia)

Tabla 2. Principales categorías de WOS

Categorías de Web of Science	Registros	% of 907	% Acum.
EDUCATION EDUCATIONAL RESEARCH	169	0,186	0,186
EDUCATION SCIENTIFIC DISCIPLINES	106	0,117	0,303
MANAGEMENT	58	0,064	0,367
HEALTH CARE SCIENCES SERVICES	57	0,063	0,430
MEDICINE GENERAL INTERNAL	39	0,043	0,473
SOCIOLOGY	35	0,039	0,512

Fuente: WOS (elaboración propia)

3.2. Indicadores de asociaciones temáticas

Del corpus fueron descartados 328 documentos que no presentaban KWP. Se identificaron luego de la normalización 1300 descriptores únicos (tabla 3).

Tabla 3. Presencia y normalización de las KWP en el corpus

Período completo (1928-2016)			
Documentos	Totales		907
	Indizados con KW+ [a]		579 (64,0%)
KW+	Totales [b]		2978
	Únicos	Sin normalización	1378
		Con normalización [c]	1300 (94,3%)
	Por documento [b/a]		5,1
	Frecuencia media de aparición [b/c]		2,3

Fuente: WOS (elaboración propia)

Como ya se mencionó, al análisis de las temáticas de los documentos se aplicaron tres enfoques: Zipf, Price y ARS, que requirieron algunos recortes (tabla 4). Zipf se aplicó sobre todas las KWP, mientras que para Price se consideraron solo las palabras presentes en cuatro años (476 artículos con 125 KWP) y, por lo tanto, el período se vio reducido a 1991-2016. Finalmente, para ARS se consideraron las KWP que además de estar presentes en cuatro años estuvieran en la primera zona de Zipf (476 artículos con 54 KWP).

Tabla 4. Resumen de los recortes sucesivos y análisis realizados en el corpus

Cantidad de registros	% sobre 907	Criterio de selección	Cantidad de KWP	Período temporal	Análisis aplicado
907	1,00		-	1928-2016	
579	0,64	Registros con KWp	1.300	1991-2016	Ley de Zipf
476	0,52	KWp presentes en 4 años	125	1991-2016	Estabilidad de Price; ARS
476	0,52	KWp presentes en la 1ra. Zona de Zipf	54	1991-2016	ARS (visualización)

Fuente: WOS (elaboración propia)

3.2.1. Modelo matemático de Zipf

Se identifican tres zonas de frecuencia de palabras clave, constituidas de la siguiente manera: la primera la integran todos los términos con una frecuencia igual o mayor a la raíz cuadrada del término más frecuente (*science*, frecuencia 84, raíz cuadrada 9,17), la segunda se calcula de la misma manera utilizando el término más frecuente que quedó fuera de la primera zona (*behavior*, frecuencia 8, raíz cuadrada 2,83) y la tercera la conforman el resto de los términos. El 4 % ($n = 55$) de los descriptores explican el 37 % de las ocurrencias; si se consideran las dos primeras zonas tenemos que el 25 % de los descriptores explican el 58 % de las ocurrencias (tabla 5).

Tabla 5. Distribución de la frecuencia de KWP en el corpus

Palabras	Porcentaje	Porc. Acum.	Frecuencia	Porcentaje	Porc. Acum.
55	4%	4%	1101	37%	37%
146	11%	15%	631	21%	58%
1099	85%	100%	1244	42%	100%

Porc. Acum. = Porcentaje acumulado

Total de palabras 1300

Frecuencia 2976

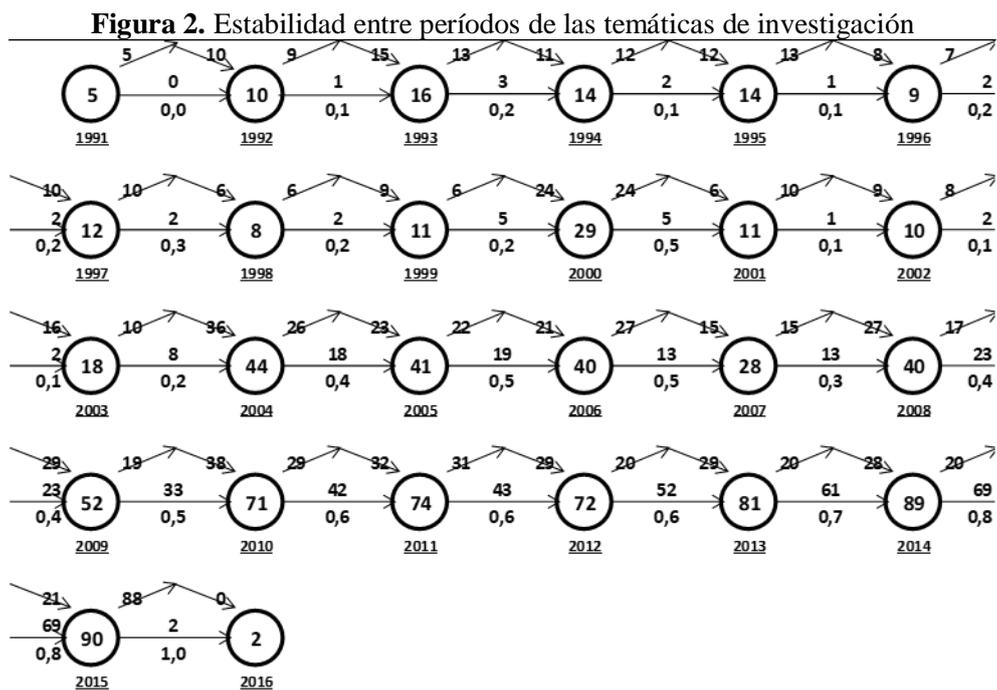
Fuente: WOS (elaboración propia)

3.2.2. Análisis de la estabilidad entre períodos

En la figura 2 se presenta el gráfico de solapamiento de las palabras clave que corresponden al período estudiado. Los círculos representan el período y el número de palabras clave asociadas (la unidad de análisis). La flecha horizontal representa el número de elementos compartido entre ambos períodos, el número en la parte superior de la flecha representa las palabras clave compartidas y el índice de estabilidad se representa debajo (el cociente de elementos compartidos con el período *a* sobre el total de elementos del período *b*). La flecha de entrada superior representa el número de elementos nuevos (que no estaban presentes en el período anterior) y la flecha de salida superior representa el número de elementos que no están en el período siguiente (Cobo *et al.*, 2012). Como ejemplo, veamos las entradas del año 2013: tenemos 81 palabras clave, de las cuales 52 (60%) estaban presentes en el período anterior (2012); los elementos nuevos, que no estaban presentes en el 2012, son 29. Con respecto a las salidas: 61 palabras clave (70%) están presentes en el 2014 y 29 no lo están.

El período estudiado puede dividirse en dos etapas: 1) 1990-2003, con 80 artículos y un bajo índice de estabilidad, un promedio de 0,2 para todo el período y un máximo de 0,5 para los años 2000-2001, y 2) 2004-2016, con 396 artículos y un índice promedio de 0,5 (sin considerar 2015-2016) que alcanzó en los dos últimos años el 0,8. Esto significa que del conjunto de KWP más frecuentes (que están presentes en cuatro años o más), a partir del año 2009, un 50% pasan de un período al siguiente.

Es importante señalar que de las 55 KWP pertenecientes a la primera zona de Zipf, 54 están presentes en cuatro períodos o más.



Fuente: WOS (elaboración propia)

3.2.3. Análisis de redes sociales

Para este análisis se construye la matriz de coocurrencias de 125×125 KWP y una matriz de atributos en la que se establecen para cada nodo la frecuencia de aparición en la totalidad del corpus y la pertenencia a la primera zona de Zipf. Finalmente, se realizan los grafos de coocurrencia de KWP con el *software* Gephi (Bastian, Heymann y Jacomy, 2009). Para la distribución se utiliza el algoritmo Fruchterman-Reingold (Fruchterman y Reingold, 1991) y Force Atlas 2 (Jacomy *et al.*, 2014); se identifican cuatro comunidades a través del algoritmo de detección de modularidad incorporado en Gephi (Blondel *et al.*, 2008; Lambiotte, Delvenne y Barahona, 2008). Para mejorar la visualización se descartan los nodos que no pertenecen a la primera zona de Zipf, lo que da como resultado un grafo de 54 nodos y 687 aristas (figura 3). En la tabla 6 se presentan las medidas de centralidad de las 20 palabras con mayor medida de vector propio. Esta medida da una orientación de la relevancia de un nodo en la red considerando la importancia de los nodos a los que está conectado. En la última columna de la tabla se indica a qué comunidad pertenece el nodo: 8 KWP pertenecen a *women*, 7 a *gender*, 4 a *science* y 1 a *model*.

Podría decirse que los estudios sobre trayectorias académicas se han concentrado en tres tópicos centrales: estudios de género (en especial la situación de las mujeres), la ciencia como actividad social y la educación superior. En un segundo nivel se identifican otros tópicos transversales a estos: desempeño, logros, rendimiento académico, productividad, impacto, gestión y evaluación. En Cathelijn Waaijer (2013) se llega a algunos tópicos similares partiendo del análisis de las editoriales de las revistas *Science* y *Nature* que trataban de la temática de las trayectorias académicas entre los años 2000 y 2012.

Figura 3. Grafo de KWP presentes en cuatro años o más y pertenecientes a la primera zona de Zipf. **Nota:** Los nodos representan KWP, el tamaño es proporcional al grado con peso, el grosor de las aristas es proporcional a las coocurrencias y el color identifica comunidades: violeta, *science*; verde, *women*; rojo, *gender*, y celeste, *model*.

Fuente: WOS (elaboración propia)

4. Conclusiones

La importancia de la investigación sobre trayectorias académicas cobra mayor relevancia a partir de la emergencia del paradigma de evaluación de la producción y reproducción de la actividad científica a partir de las capacidades del investigador.

La metodología aplicada para identificar las temáticas combina tres análisis: el modelo matemático de Zipf, la estabilidad de Price y el ARS. Es importante destacar que para Price no se tiene en cuenta la frecuencia de KWP en el período de tiempo y para Zipf, sí. Sin embargo, existe un solapamiento casi completo en el que los tópicos de la primera zona de Zipf están en el recorte de Price. En efecto, de 55 KWP solamente una no cumple con el criterio para el análisis de Price; podría interpretarse como una temática que tiene nueve o más artículos en el corpus, pero concentrados en menos de cuatro períodos.

Los trabajos sobre la temática de las trayectorias académicas han tenido un desarrollo importante a partir del 2004. Esto se ve reflejado en un aumento de la producción y una consolidación de un núcleo de tópicos de los que se ocupa.

El 30% de la producción se encuentra concentrada en 35 títulos de revista (6% del total de títulos). El 50% de los títulos se distribuyen entre seis categorías: dos refieren a educación, dos a salud, uno a gestión y otro a sociología, lo que da cuenta del abordaje de esta temática desde distintas disciplinas, es decir, de la convergencia de múltiples enfoques sectoriales que no se conectan entre sí.

Por otra parte, la ausencia de componente gigante en la red de autoría y la gran fragmentación de las comunidades, con más de 500, no permite identificar grupos de autores que se encarguen de la temática.

Se identificaron tres tópicos: estudios de género, educación superior y ciencia, y otros temas transversales a estos. Es de destacar que el 50% de los artículos del corpus tratan algunos de estos tópicos o temas transversales. Sin embargo, hay un 36% de los artículos ($n=328$) que fueron descartados por no tener KWP, lo cual representa, sin duda, una debilidad del trabajo.

La evolución de la producción sobre trayectorias académicas muestra cierta consistencia con el desarrollo de procesos socioeducativos que forman parte del contexto del problema de investigación: el desarrollo de los sistemas universitarios con el incremento del gasto en educación; la revolución científico-tecnológica; el surgimiento del enfoque de derechos, con cambios relevantes en las relaciones de género; los cambios de paradigma en la atención y gestión de los sistemas y servicios de salud; el desarrollo del *management* y su implementación paulatina en los sistemas universitarios y de gestión del conocimiento y, por último, el perfeccionamiento de la evaluación educativa y de

los sistemas educativos. Todas estas condiciones contextuales pueden ser factores que generen demanda de investigación sobre trayectorias académicas.

La evidencia encontrada plantea en principio una imagen de dispersión que no identifica una comunidad académica que claramente se ocupa del tema. Esta dispersión también puede indicar que un grupo importante de instituciones o comunidades académicas incursionan en algún momento en forma puntual y sectorial en la investigación de trayectorias académicas sin tener líneas de investigación específicas sostenidas en el tiempo.

Teniendo en cuenta que gran parte de los tópicos transversales se refieren mayoritariamente a enfoques vinculados a la gestión y a la evaluación, se puede pensar que hay un tema central que está ligado a la educación superior y la ciencia. El otro tópico relevante, género, se puede relacionar con el análisis de las trayectorias académicas como elemento central del enfoque de género, tanto como insumo para lograr acceso a derechos, como resultado de un proceso de cambio o, finalmente, como escenario en el que se pueden observar relaciones inequitativas de género. En estas áreas pueden incursionar desde sus perspectivas sectoriales la investigación en sistemas y servicios de salud, la investigación educativa y la sociología de las profesiones y de género, entre otras.

Referencias bibliográficas

- ANDRÉ, C.-S *et al.* (2014). Análise bibliométrica do periódico *Transinformação. El Profesional de la Información*, 23(4), 433-442.
- BASTIAN, M.; HEYMANN, S.; JACOMY, M. (2009). *Gephi: an open source software for exploring and manipulating networks*. Conferencia llevada a cabo en Third International AAAI Conference on Weblogs and Social Media, San José, California.
- BEIGEL, F.; SALATINO, M. (2015). Circuitos segmentados de consagración académica: las revistas de ciencias sociales y humanas en la Argentina. *Información, Cultura y Sociedad*, 32, 11-36.
- BLONDEL, V. D. *et al.* (2008). Fast unfolding of communities in large networks. *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), 1-12. doi:10.1088/1742-5468/2008/10/P10008
- BOZEMAN, B.; DIETZ, J.; GAUGHAN, M. (2001). Scientific and Technical Human Capital: an alternative model for research evaluation. *International Journal of Technology Management*, 22(7/8), 716-740. doi:10.1504/IJTM.2001.002988
- BRADFORD, S. C. (1948). *Documentation*. Londres: Crosby Lockwood and Son.
- CANTOS-MATEOS, G. *et al.* (2014). Estudio evolutivo de la investigación española con células madre: visualización e identificación de las principales líneas de investigación. *El Profesional de la Información*, 23(3), 259-271.
- CAÑIBANO, C.; BOZEMAN, B. (2009). Curriculum vitae method in science policy and research evaluation: the state-of-the-art. *Research Evaluation*, 18(2), 86-94. doi:10.3152/095820209X441754
- CARRINGTON, P. J.; SCOTT, J.; WASSERMAN, S. (Eds.). (2005). *Models and methods in Social Networks Analysis: structural analysis in the social sciences*. Cambridge: Cambridge University Press.
- COBO, M. J. *et al.* (2012). SciMAT: a new science mapping analysis software tool. *Journal of the American Society for Information Science and Technology*, 3(8), 1609-1630. doi:10.1002/asi.22688
- FREEMAN, L. C. (2012). *El desarrollo del análisis de redes sociales: un estudio de sociología de la ciencia*. Bloomington: Palibrio.
- FRUCHTERMAN, T. M. J.; REINGOLD, E. M. (1991). Graph Drawing by Force-directed Placement. *Software: Practice and Experience*, 21(11), 1129-1164. doi:10.1002/spe.4380211102
- GARFIELD, E. (1990). KeyWords Plus: ISI's Breakthrough Retrieval Method. Part 1. Expanding Your Searching Power on Current-Contents on Diskette. *Current Contents*, 1(32), 5-9.
- GARFIELD, E.; SHER, I. H. (1993). KeyWords Plus: Algorithmic Derivative Indexing. *Journal of the American Society for Information Science*, 44(5), 298-299. doi:10.1002/(sici)1097-4571(199306)44:5<298::aid-asi5>3.0.co;2-a

- GAVEL, Y.; ISELID, L. (2008). Web of Science and Scopus: a journal title overlap study. *Online Information Review*, 32(1), 8-21. doi:10.1108/14684520810865958
- GORRAIZ, J.; SCHLOEGL, C. (2008). A bibliometric analysis of pharmacology and pharmacy journals: Scopus versus Web of Science. *Journal of Information Science*, 34(5), 715-725. doi:10.1177/0165551507086991
- JACOMY, M. *et al.* ForceAtlas2, a continuous graph layout algorithm for handy network visualization designed for the Gephi software. *PloS One*, 9(6), e98679. doi:10.1371/journal.pone.0098679
- JARAMILLO, H.; LOPERA, C.; ALBÁN, M. C. (2008). Carreras académicas: utilización del CV para la modelación de carreras académicas y científicas. En *Serie Documentos Borradores de Investigación*, 96. Facultad de Economía, Universidad del Rosario. Recuperado de <http://repository.urosario.edu.co/bitstream/handle/10336/11286/4671.pdf>
- LAMBIOTTE, R.; DELVENNE, J.-C.; BARAHONA, M. (2008). Laplacian Dynamics and Multiscale Modular Structure in Networks. *IEEE Transactions on Network Science and Engineering*, 1(2), 76-90. doi:10.1109/TNSE.2015.2391998
- LÓPEZ-ILLESCAS, C.; MOYA-ANEGÓN, F.; MOED, H. F. (2008). Coverage and citation impact of oncological journals in the Web of Science and Scopus. *Journal of Informetrics*, 2(4), 304-316. doi:10.1016/j.joi.2008.08.001
- MOLINA GONZÁLEZ, J. L. (2001). *El análisis de redes sociales: una introducción*. Barcelona: Bellaterra.
- PINTO, A. L.; GONZALES-AGUILAR, A. (2014). Visibilidad de los estudios en análisis de redes sociales en América del Sur: su evolución y métricas de 1990-2013. *Transinformação*, 26(3), 253-267. doi:10.1590/0103-3786201400030003
- PRICE, D. J. DE S.; GÜRSEY, S. (1975). Studies in scientometrics I: transience and continuance in scientific authorship. *Ciência da Informação*, 4(1), 27-40.
- QUONIAM *et al.* (2001). Inteligência obtida pela aplicação de data mining em base de teses francesas sobre o Brasil. *Ciência da Informação*, 30(2), 20-28. doi:10.1590/S0100-19652001000200004
- SANCHO, R. (1990). Indicadores bibliométricos utilizados en la evaluación de la ciencia y la tecnología. Revisión bibliográfica. *Revista Española de Documentación Científica*, 13(3-4), 842-865.
- SPINAK, E. (1998). Indicadores cienciométricos. *Ciência da Informação*, 27(2), 141-148.
- TESTA, J. (2016). Journal Selection Process. Recuperado de http://wokinfo.com/media/pdf/journal_selection_essay-es.pdf
- TORRES-SALINAS, D.; DELGADO LÓPEZ-CÓZAR, E.; JIMÉNEZ-CONTRERAS, E. (2009). Análisis de la producción de la Universidad de Navarra en revistas de ciencias sociales y humanidades empleando ranking de revistas españolas y la Web of Science. *Revista Española de Documentación Científica*, 32(1), 22-39. doi:10.3989/redc.2009.1.673

- WAAIJER, C. (2013). Careers in science: policy issues according to *Nature* and *Science* editorials. *Scientometrics*, 96(2), 485-495. doi:10.1007/s11192-013-0958-z
- WASSERMAN, S.; FAUST, K. (1994). *Social network analysis: Methods and applications*. Cambridge: Cambridge University Press.
- ZHOU, P.; THIJS, B.; GLÄNZEL, W. (2009). Is China also becoming a giant in social sciences? *Scientometrics*, 79(3), 593-621. doi:10.1007/s11192-007-2068-x
- ZIPF, G. K. (1949). *Human behaviour and the principle of least effort. An introduction to Human Ecology*. Cambridge: Adisson-Wesley Press.

Anexo. Medidas centrales de las 44 KWP.

Label	Cant.	Degree	Deg. N.	Clos. N.	Bet. N.
SCIENCE	42	53	0,663	0,752	0,060
PERFORMANCE	22	42	0,525	0,681	0,032
SEX_DIFFERENCES	16	37	0,463	0,653	0,024
IMPACT	18	32	0,400	0,627	0,020
SCIENTIFIC_PRODUCTIVITY	10	25	0,313	0,594	0,007
RESEARCH_PRODUCTIVITY	8	23	0,288	0,585	0,008
PATTERNS	8	21	0,263	0,572	0,005
PUBLICATION_PRODUCTIVITY	8	21	0,263	0,572	0,004
EDUCATION	20	34	0,425	0,637	0,013
PERCEPTIONS	18	34	0,425	0,637	0,023
STUDENTS	22	31	0,388	0,622	0,018
SCHOOLS	15	29	0,363	0,608	0,011
UNITED_STATES	11	27	0,338	0,603	0,010
ACHIEVEMENT	11	22	0,275	0,577	0,005
CHOICE	13	21	0,263	0,577	0,008
GENDER_DIFFERENCES	7	12	0,150	0,534	0,002
UNIVERSITY	26	44	0,550	0,693	0,024
ACADEMIC_CAREERS	25	42	0,525	0,681	0,024
WORK	20	38	0,475	0,658	0,020
MODEL	18	31	0,388	0,622	0,018
MANAGEMENT	14	31	0,388	0,622	0,009
SATISFACTION	11	28	0,350	0,608	0,013
MOBILITY	10	27	0,338	0,598	0,005
ORGANIZATIONS	10	25	0,313	0,594	0,008
OUTCOMES	9	22	0,275	0,581	0,002
KNOWLEDGE	8	17	0,213	0,556	0,003
DISCRIMINATION	7	16	0,200	0,556	0,006
FUTURE	11	15	0,188	0,541	0,002
HIGHER_EDUCATION	9	14	0,175	0,534	0,003
WOMEN	52	62	0,775	0,823	0,088
FACULTY	33	54	0,675	0,760	0,064
CAREER	26	44	0,550	0,693	0,028
GENDER	24	40	0,500	0,669	0,017
SUCCESS	13	39	0,488	0,664	0,021
PERSPECTIVE	15	37	0,463	0,653	0,015
PRODUCTIVITY	17	37	0,463	0,648	0,020
PROGRAM	17	36	0,450	0,648	0,020
ADVANCEMENT	7	27	0,338	0,598	0,005
DEPARTMENTS	8	26	0,325	0,594	0,008
ACADEMIC_MEDICINE	14	26	0,325	0,598	0,008
MEDICINE	17	26	0,325	0,598	0,007
SCIENTISTS	10	22	0,275	0,577	0,005
PHYSICIANS	13	18	0,225	0,564	0,002
PSYCHOLOGY	7	12	0,150	0,530	0,003

Cant. = Frecuencia de aparición en el corpus

Deg. N. = Degree normalizado

Clos. N. = Closeness normalizado

Bet. N. = Betweenness Normalizado

Fuente: WOS (elaboración propia).