

The logo for the journal EMPIRIA, featuring the word in a bold, serif font inside a rectangular border.

EMPIRIA. Revista de Metodología de las Ciencias Sociales

ISSN: 1139-5737

ISSN: 2174-0682

empiria@poli.uned.es

Universidad Nacional de Educación a Distancia

España

Martínez-Urbe, Luis

La sociología a través de sus publicaciones en revistas de impacto mediante el uso de big data

EMPIRIA. Revista de Metodología de las Ciencias Sociales, núm. 53, 2022, Enero-, pp. 53-88

Universidad Nacional de Educación a Distancia

Madrid, España

DOI: <https://doi.org/10.5944/empiria.53.2022.32612>

Disponible en: <https://www.redalyc.org/articulo.oa?id=297170952004>

- Cómo citar el artículo
- Número completo
- Más información del artículo
- Página de la revista en [redalyc.org](https://www.redalyc.org)

[redalyc.org](https://www.redalyc.org)

Sistema de Información Científica Redalyc

Red de Revistas Científicas de América Latina y el Caribe, España y Portugal

Proyecto académico sin fines de lucro, desarrollado bajo la iniciativa de acceso abierto

La sociología a través de sus publicaciones en revistas de impacto mediante el uso de big data

Sociology through its publications in impact factor journals using big data

LUIS MARTÍNEZ-URIBE

Fundación Juan March
Universidad de Salamanca
lmartinez@march.es (ESPAÑA)
ORCID <https://orcid.org/0000-0002-7795-3972>

Recibido: 14.09. 2020

Aceptado: 23.11.2021

RESUMEN

Al igual que las demás disciplinas científicas, la sociología se puede concebir como un sistema social compuesto de complejas relaciones entre actores que incluyen a investigadores, instituciones, revistas y editoriales. Actualmente, el fenómeno del big data ofrece la posibilidad de usar grandes colecciones de datos que permiten la representación de los vínculos sociales que se dan en la ciencia. En concreto, a través de las grandes fuentes de datos bibliométricas, big scholarly data, la sociología tiene a su alcance ingentes cantidades de datos para describir y estudiar con detalle la evolución de las disciplinas científicas.

En este artículo describimos la sociología de los últimos años a través de las publicaciones en las revistas de impacto. Para hacer esto, se emplean datos de revistas de sociología del Journal Citation Reports ampliados con la información de los artículos del Microsoft Academic Graph. Realizamos un análisis descriptivo de las revistas, sus países de origen, lenguas, editoriales y décadas de aparición e impacto. A continuación, evaluamos la evolución temporal del número de artículos y citas, así como la coautoría y el género de los autores. Tras esto, establecemos cuatro grupos de tipos de revistas y estudiamos sus diferencias en las dimensiones anteriores mediante contrastes de hipótesis. Finalmente, representamos las relaciones entre autores y revistas usando una red de afiliación que nos permite detectar grupos de revistas que forman interesantes comunidades temáticas y geográficas.

La novedad del trabajo consiste en haber utilizado una fuente de datos de las denominadas big scholarly data con más de 300 millones de publicaciones y

aportar varias estrategias que permiten seleccionar los datos de interés entre los millones de publicaciones y reducir su dimensionalidad para representarlos en forma de red. Los resultados corroboran análisis de estudios previos y presentan la sociología como una disciplina dominada por revistas generalistas anglosajonas que abarca una amplia variedad de temáticas y con enfoques metodológicos diversos que dependen del ámbito geográfico en el que se desarrollan. Unas pocas revistas dominan las citas, mientras que los artículos de revistas metodológicas tienen un grado mayor de coautoría y las revistas temáticas un menor sesgo de género.

PALABRAS CLAVE

Sociología del conocimiento, sociología, grandes datos, datos masivos, redes de afiliación, bibliometría, métodos de investigación, Microsoft Academic Graph, análisis de coincidencias.

ABSTRACT

Like other scientific disciplines, sociology can be observed as a social system made up of researchers, institutions, journals and publishers. These relationships are established via conceptual communications which form networks that establish the way in which disciplines are organized. At present, the big data phenomena offers the capacity to use large data collections to analyse social processes. Big scholarly data sources offer sociology immense quantities of data useful to describe and study the evolution of scientific disciplines in detail.

In this article we characterised the last thirty years of sociology through its publications in impact factor journals. To do this, we use data about the sociology journals from Journal Citation reports augmented with article information from Microsoft Academic Graph. The analysis starts by describing the journals, countries of origin, languages, publishers, the decades in which they appeared and their impact factor. After this, we evaluate the evolution of numbers of articles and citations as well as co-authorship and gender proportion. Subsequently, we establish four groups of journal types and study their differences in the previous dimensions using hypothesis tests. Finally, we represent the relationships between authors and journals using an affiliation network that allows us to detect groups of journals that form interesting thematic and geographic communities.

The novelty of the work consists in having used a data source of the so-called big scholarly data with more than 300 million publications. The paper also provides several strategies to select the data of interest among the millions of publications to reduce their dimensionality in order to represent them in the form of a network. The results show a discipline dominated by Anglo-Saxon countries and large publishing conglomerates. The most prominent journals dominate citations whilst methodological journals have a higher degree of co-authorship and

thematic journals have the lowest gender bias. The affiliation network between authors and journals contains two large groups, one formed by the pioneering American journals together with quantitative methodological journals and another one made up of English and qualitative methodological journals.

KEYWORDS

Sociology of knowledge, sociology, big data, affiliation networks, bibliometrics, research methods, Microsoft Academic Graph, coincidence analysis.

1. INTRODUCCIÓN

La sociología como disciplina científica surge en distintas comunidades nacionales en distintos momentos en el tiempo y ha evolucionado en una variedad amplia de temáticas y métodos para resolver el extenso espectro de problemas que se plantea. La disciplina ha sido acusada de estar dominada por los países anglosajones, fragmentada en temáticas volátiles, con falta de cohesión metodológica y sesgo de género. Estudios bibliométricos en los últimos años se han encargado de mapear y analizar la evolución de las disciplinas a partir de la información de sus publicaciones. En sociología estos estudios se han centrado en medir la producción y el impacto de ciertas comunidades y subdisciplinas, pero no de la disciplina en su conjunto. La reciente aparición del fenómeno *big data* ha generado nuevas fuentes de datos que aportan nuevas oportunidades y retos metodológicos. En el terreno de bibliometría este es el caso que nos plantean los datos del Microsoft Academic Graph, una fuente de datos reciente, denominada *big scholarly data*, que cuenta con 230 millones de publicaciones con información enriquecida mediante técnicas de inteligencia artificial. ¿Hasta qué punto estos datos corroboran algunos de estos aspectos de la sociología? ¿Cuáles son los retos metodológicos que nos plantea su utilización?

El objetivo de este artículo consiste en caracterizar la disciplina sociológica a través de sus publicaciones en revistas de impacto utilizando *big scholarly data*, nuevos métodos y estrategias que nos permiten reducir la dimensionalidad para explorar grandes cantidades de datos.

En este artículo se concibe la sociología como una disciplina en la que se entretienen complejos sistemas de relaciones sociales entre investigadores, instituciones, asociaciones y editoriales. Partiendo de la premisa de que esas relaciones y su evolución en el tiempo se conforman en forma de redes, se da cuenta de la sociología de los últimos años a través de sus revistas de impacto. Para ello, se emplean datos de revistas de sociología del *Journal Citation Reports* ampliados con la información de los artículos del *Microsoft Academic Graph*.

El artículo se estructura de la siguiente forma: comienza con una breve introducción de la sociología como disciplina científica resaltando el corpus teórico que la conceptualiza como un sistema social dominado por sus comunicaciones.

Se presenta a continuación el fenómeno *big bata* con sus expectativas y retos para después enfocarse en los grandes datos bibliométricos con su capacidad para analizar las disciplinas científicas. Las secciones de datos y metodologías describen las fuentes de datos que se emplean y los métodos estadísticos que se aplican. Tras esto, se exploran los datos obtenidos a través de análisis descriptivos, test de hipótesis y análisis de redes sociales.

2. LA SOCIOLOGÍA Y SU CONCEPTUALIZACIÓN COMO DISCIPLINA

La definición de la sociología a través de la identificación de sus dominios y métodos ha sido un tema recurrente desde sus inicios. Emile Durkheim (1982) la definía como la ciencia de los hechos sociales y de esta forma la diferenciaba del resto de disciplinas científicas. Por otro lado, Albion W. Small (1906), fundador del primer departamento de sociología en Estados Unidos, leía un artículo ante el *Sociology Club* de la Universidad de Chicago donde consideraba una pérdida de tiempo intentar definir la disciplina.

La sociología como disciplina ha sido acusada de fragmentarse con nuevas áreas de interés que surgen constantemente y que amenazan su estatus y unidad (Moody y Light, 2006, O'Reilly, 2009). Esta misma preocupación la recoge Smelser al destacar la falta de cohesión en el nivel conceptual:

“...sociology, by comparison with some other sciences, lacks a single, accepted conceptual framework. The field is difficult to distinguish from other because it contains a diversity of frameworks, some of which it shares with other fields such as psychology and social anthropology. If anything, then, sociology is too comprehensive, diffuse, soft in the center, and fuzzy around the edges.” (Smelser, 2014)

Otras perspectivas la describen como una disciplina que se origina de modo diverso en distintas comunidades nacionales, que con el tiempo pasan a formar redes supranacionales que terminan siendo globales (Vanderstraeten, 2010). Distintos estudios revelan el dominio de prominentes revistas generalistas anglosajonas (Moody y Light, 2006), una división marcada entre la tradición americana más cuantitativa y la británica más teórica (Zougiris, 2018), una coautoría más habitual en los trabajos cuantitativos (Moody, 2004) y un cierto sesgo de género entre los autores que publican (Grant y Ward, 1991).

3. MARCO TEÓRICO: LAS DISCIPLINAS COMO SISTEMAS SOCIALES CONSTITUIDOS POR REDES DE PUBLICACIONES

Este artículo se encuadra en un marco teórico que concibe la producción científica desde el estudio de los procesos de interacción social, de su entorno y de su evolución en el tiempo. La imaginación social (Mills, 1959) cobra entonc-

es un papel esencial para apreciar el escenario en el que estos procesos actúan. Estudios en la sociología del conocimiento sugieren que el conjunto de ideas que uno considera verdaderas depende en gran medida del grupo al que pertenezcas. Así científicos que pertenecen a redes de colaboración comparten ideas, utilizan metodologías similares y se influyen unos a otros (Moody, 2004).

Las disciplinas científicas pueden observarse como sistemas sociales cuya comunicación es fundamental para entender los mecanismos que las conforman. Estos dominios científicos son dependientes de comunicaciones conceptuales observadas por terceros (Stichweh, 2008). Los artículos académicos son un buen ejemplo de comunicaciones, suponen prestigio, establecen reglas de recompensa y reclaman autoridad (Stinchcombe, 1984). De manera discontinua, estas disciplinas se forman y evolucionan en el tiempo a través de un compendio compartido de teorías, métodos y problemas que han de ser solucionados (Khun, 1962:11). La ciencia es pues un ámbito social donde están presentes el interés, las relaciones de poder, cuando no la ostentación de prestigio (Bourdieu, 2004: 29).

La actual comunicación científica se engloba en la sociedad red definida por Castells como “*la nueva estructura social de la Era de la Información, basada en redes de producción, poder y experiencia*” (Castells, 1998:350). Estas redes se convierten en interesantes laboratorios para entender e interpretar el proceso de producción científica (Latour y Woolgar, 1987). De este modo, las disciplinas pueden estudiarse como redes de publicaciones en las que se van construyendo temas y procedimientos aceptados por la comunidad científica (Luhmann, 1995).

Moody (2004) utiliza la idea de estructura de red y propone tres formas posibles de redes de colaboración en sociología. Una primera estructura es la que se ve afectada por no disponer de una teoría unificada de la disciplina, esto hace que la red esté compuesta por múltiples especialidades desconectadas y con subredes altamente agrupadas. La segunda estructura está determinada por la idea de que la producción científica depende de unos pocos científicos, *scientific stars*, cuyo trabajo determina el curso de la disciplina y genera redes en forma de estrella. Finalmente, las colaboraciones con límites teóricos más permeables generan redes de gran alcance y cohesionadas estructuralmente.

4. **BIG SCHOLARLY DATA, UNA FUENTE PARA EL ANÁLISIS DE DISCIPLINAS**

El primer uso documentado del término *big data* aparece en un artículo de científicos de la NASA en 1997 describiendo un problema de visualización de datos debido a conjuntos de datos tan grandes que ponen a prueba la capacidad de la memoria principal, el disco local e incluso el disco remoto. A esto lo llamaron el problema del *big data* (Press, 2014). El *big data* se puede concebir como una combinación de nuevas fuentes de datos de gran tamaño con las infraestructuras tecnológicas y los métodos para su tratamiento. El fenómeno del *big data* trae consigo beneficios para la investigación social como la capacidad

de usar datos sobre fenómenos sociales no disponibles anteriormente, poseer información sobre poblaciones completas, recabar datos de modo inconsciente para la persona observada (*unobtrusive data*) y analizar información de procesos sociales en tiempo real (Boyd y Crawford, 2012, Espeland y Stevens, 2008, Manovich, 2015, Martinho, 2018, McFarland et al., 2016, Moretti, 2000, Tinati et al., 2014). Como contrapeso al entusiasmo suscitado por estos nuevos recursos, no faltan los análisis críticos que reflexionan sobre sus limitaciones y reabren el debate ya acontecido en otros dominios científicos sobre la investigación centrada en datos versus la investigación basada en hipótesis (Carroll, 2009, Neresini, 2017). Ya en los años 80 los estudios interpretativos sobre la construcción social de la tecnología consideran los datos y la tecnología como el resultado de un proceso de construcción social y advertían del peligro de las perspectivas positivistas donde los datos son considerados objetivos (Pinch y Bijker, 1984). La literatura más reciente nos alerta de los riesgos de los métodos y teoría que acompañan al Big Data ya que en ocasiones los datos solo capturan información de ciertas actividades, pueden ser incorrectos, parciales o no contar con información sobre su procedencia (Giardullo, 2016, Halford y Savage, 2017, McFarland y McFarland, 2015).

En las ciencias sociales nuevas subdisciplinas como *social data science* o *computational social science* están asumiendo este reto, abordándolo de manera interdisciplinar junto con matemáticos e ingenieros (Blok et al., 2017, Burrows y Savage, 2014, Lazer et al., 2009, Savage y Burrows, 2007). En este contexto, los sociólogos han de contribuir con su conocimiento teórico a interpretar cómo se estructura lo social (Tubaro, 2014). Los datos por si mismos no son suficientes (Grimmer, 2015) y el trabajo con grandes datos debe seguir siendo una operación teórica, pues la interpretación es crucial para analizar la realidad social (Boyd y Crawford, 2012).

La sociología ha utilizado métodos cuantitativos desde sus inicios aportando rigor a la disciplina. Asimismo, en muchos casos la sociología ha contribuido de manera importante al desarrollo de métodos estadísticos (Clogg, 1992). La revisión de Raftery (2000) del uso de estadísticas en sociología presenta tres generaciones de métodos estadísticos en sociología: una primera generación, a partir de la Segunda Guerra Mundial, centrada en las tabulaciones cruzadas de encuestas y censos de pocas variables; una segunda generación desde 1960 que trata con datos de encuesta con muchas variables, y una tercera generación que comienza en los años 80 donde los formatos varían e incluyen datos textuales, redes sociales o datos espaciales. Actualmente, también se utilizan técnicas como la visualización de datos (Healy y Moody, 2014), metodologías de procesamiento del lenguaje natural (PLN) (Evans y Aceves, 2016), minería de datos o modelos de aprendizaje automático (*machine learning ML*) (Amaturo y Punziano, 2017, Frank et al., 2019).

Por otro lado, la bibliometría ha acuñado el término *big scholarly data* (Xia et al., 2017), para referirse a aquellas fuentes académicas de datos que han sufrido un crecimiento exponencial en los últimos años. Bases de datos, algunas de pago como *Web of Knowledge* y *Scopus* o gratuitas como *Google Scholar*

y *Microsoft Academic Graph*, acumulan una inmensa cantidad de información académica sobre autores, citas, artículos de revistas, actas de congresos, tesis y libros. Estos datos proveen indicadores que se analizan generalmente en investigaciones bibliométricas, para medir el impacto y la productividad científica o evaluar las redes académicas de comunicación e investigación.

Además, en los últimos años, según describen Su & Lee (2010), existe una multitud de estudios que mapean y estudian la evolución de dominios científicos a partir de los datos y metadatos de sus publicaciones científicas. Existe una amplia tradición de estudio de disciplinas científicas en la sociología de la ciencia. Estos estudios entienden la comunicación científica a través de revistas especializadas como una forma de organización y control de las disciplinas constituidas como construcciones sociales (Vanderstraeten, 2010). Sin embargo, Gupta & Battacharya (2004) argumentan que un nuevo enfoque de estudio surge cuando comienzan los estudios cuantitativos con información de publicaciones, ya que estas no solo revelan la estrategia científica de sus autores, sino que también proporcionan información sobre las dinámicas compartidas por la disciplina a la que pertenecen.

Los estudios bibliométricos de la sociología se han centrado hasta el momento en la producción sociológica y en los patrones de coautoría de países concretos como Australia, Países Bajos, Italia, Francia o España (De Haan, 1997, Gantman y Dabós, 2018, Phelan, 2000, Riviera, 2015, Vanderstraeten, 2010), en subdisciplinas como la sociología médica o la desigualdad de riqueza, (Korom, 2019, Seale, 2008), en las élites académicas en sociología (Korom, 2020), en la división de escuelas metodológicas (Oromaner, 1981, Schwemmer y Wieczorek, 2019), en la diferencia de citas entre los artículos teóricos y metodológicos (Perritz, 1983), así como en los efectos de los rankings, que favorecen ciertos tipos de culturas de investigación en los departamentos de sociología (Moksony et al., 2014).

Sin embargo, tales perspectivas dejan aún sin describir o explicar cuestiones tales como las diferencias entre diversos países, o entre diversas subdisciplinas temáticas y metodológicas, ni abordan la evolución de los patrones de género y colaboración a través de la coautoría, temas que serán objeto de revisión en las próximas páginas.

5. DATOS: JOURNAL CITATION REPORTS Y MICROSOFT ACADEMIC GRAPH

Para disponer de datos que permitan evaluar la producción científica en sociología se han seleccionado las revistas de impacto utilizando la base de datos de *Web of Science Journal Citation Reports* (JCR) (2020)¹, fuente comúnmente

¹ JCR incluye datos de citas extraídos de aproximadamente 12,000 revistas académicas y técnicas y actas de congresos de más de 3,300 editoriales en más de 60 países. Cubre casi todas las especialidades en ciencia, tecnología y ciencias sociales y permite la evaluación y comparación

reconocida para el establecimiento de rankings de las revistas académicas en función de su impacto en la producción científica. En la fecha de consulta de los datos, JCR contenía información acerca de 165 revistas en la categoría de sociología entre 1997 a 2018. Esta información se enriquece añadiendo las editoriales, los idiomas de publicación, los países de origen y la década en la que aparecen cada una de las revistas. Jacobs (2016) identifica limitaciones en los datos JCR de sociología al no ser del todo exhaustivos. En la categoría de sociología se echan en falta importantes revistas como *American Sociologist*, *Context o Work, Family and Community*. Además, algunas de las revistas presentes, en particular *Annals of Tourism Research* y *Cornell Hospitality Quarterly*, cuya temática es de dudosa inclusión en esta categoría. Estas dos últimas las eliminamos de nuestra selección. Además, los resultados de analizar estos datos pueden verse condicionados por el sesgo anglosajón en la cobertura de las revistas JCR enfocadas en publicaciones americanas, inglesas y de países bajos debido a la fuerte presencia de importantes editoriales comerciales (Rodríguez-Yunta, 2009).

La información de los artículos y autores de estas revistas se incorporaron desde los datos de Microsoft *Academic Graph* (MAG) (Sinha et al., 2015), fuente que se distribuye libremente con una licencia de datos abierta y contiene información de 230 millones de publicaciones incluyendo artículos de revistas y actas de congresos, casi 240 millones de autores, 50.000 revistas, 4.500 conferencias y 25.500 instituciones. La base de datos de MAG es el resultado de procesos de captura de datos que mezclan información indizada por el buscador *Bing* junto con las fuentes de sindicación (rss) de los editores. Estos datos son enriquecidos mediante procesos automáticos que utilizan inteligencia artificial, en concreto de procesamiento del lenguaje natural, que ayudan a detectar y desambiguar entidades y sus relaciones (autores, afiliaciones, revistas) además de identificar conceptos que definen cada una de las publicaciones y que posteriormente se organizan en una taxonomía (Wang et al., 2019). Estos datos tienen limitaciones al no incluir otros tipos de comunicación científica. La sociología, al igual que otras disciplinas de ciencias sociales, utiliza otros canales de difusión además de los artículos de revistas como son los libros o seminarios (Clemens et al., 1995).

Los datos de JCR se extrajeron de su plataforma online. Para el acceso y la consulta de los datos de MAG se desplegó una infraestructura de *Big Data* en la nube de Microsoft Azure. Los componentes necesarios incluían un almacenamiento de ficheros de texto plano con los datos y un motor de analítica con el que definir el esquema de los ficheros y realizar consultas en lenguaje U-SQL para obtener la selección de datos necesaria². Además, los datos de los autores de las publicaciones se enriquecieron añadiendo el género a través del nombre del autor utilizando Gender-API³, un servicio online que utiliza diversas fuentes

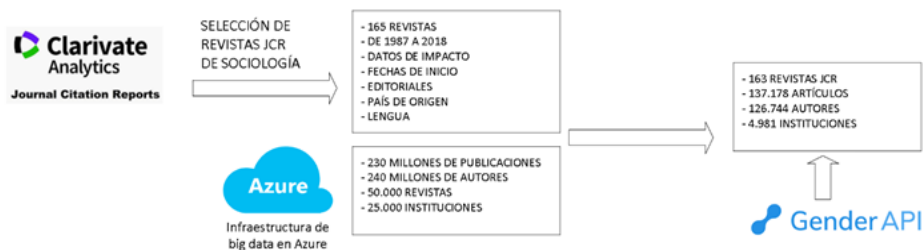
de revistas para identificar las revistas mejor clasificadas y de mayor impacto en un campo en particular.

² Instrucciones para desplegar la infraestructura necesaria en la nube están disponibles en la siguiente dirección: <https://docs.microsoft.com/en-us/academic-services/graph/get-started-set-up-provisioning>

³ Gender API es una plataforma web capaz de determinar el género a partir del nombre

y clasifica los nombres con un grado por género con un grado de acierto alto (Santamaría y Mihaljević, 2018). La figura 1 muestra esquemáticamente todo el proceso anterior de captura, selección y enriquecimiento de datos.

Figura 1. Proceso de captura y enriquecimiento de datos



6. METODOLOGÍA

Toda la manipulación y análisis de datos para este artículo se ha realizado utilizando el lenguaje de programación estadística R. En la siguiente sección se presentará un análisis descriptivo de las revistas JCR de sociología por países, idiomas de publicación, editoriales, década de aparición e impacto. Posteriormente, se mostrarán las revistas a través de sus artículos describiendo la evolución anual del número de artículos y citas. Estos análisis nos permiten empezar a identificar patrones en las tipologías de las revistas. Tras esto, se seleccionarán algunas de las revistas y se organizarán en cuatro grupos para los que se estudiarán sus diferencias en las dimensiones de citas, coautoría y proporción de género aplicando contrastes de hipótesis.

En la última sección de análisis representamos las relaciones entre revistas y autores a través de una red de afiliación. Las redes de afiliación son redes bipartitas o bimodales, es decir, aquellas que tienen dos tipos de nodos distintos y las conexiones solo se producen entre elementos que pertenecen a tipos diferentes:

Una red de afiliación $G=(U,V,E)$

donde $\forall u_1, u_2 \in U, \forall v_1, v_2 \in V$ no existen aristas e tales que $e=(u_1, u_2)$ ni $e=(v_1, v_2)$

Las redes bimodales pueden proyectarse convirtiéndose en redes de modo uno. Esta reducción de dimensionalidad permite centrarse en las relaciones de un tipo de nodo. Las redes de afiliación son especialmente útiles para establecer la relación de pertenencia entre actores y grupos. En nuestro caso, la red de afiliación tiene por actores a los autores de las publicaciones y los grupos son las revistas JCR donde publican. A través de esta red podremos detectar aquellos grupos de revistas con comunidades comunes de autores que publican en ellas.

El marco estadístico metodológico que se emplea para generar la red de afiliación es el análisis de coincidencias (Escobar, 2015, Escobar y Tejero, 2018) que es aplicado usando el paquete de R *netCoin* (Escobar y Martínez-Uribe, 2020). Este marco tiene por objetivo principal detectar el tipo de personas, eventos, atributos, etc. que tienden a aparecer de manera simultánea en un número limitado de espacios.

Partimos de N espacios delimitados denominados escenarios, en cada escenario hay un conjunto de tamaño M de variables aleatorias denominadas eventos $X_j, j=1, \dots, M$. $X_j = 1$ si el evento j ocurre y $X_j = 0$ si no ocurre. Dos eventos son coincidentes si ocurren en el mismo escenario. A partir de esta misma información, también podría decirse que dos escenarios son semejantes si concurren en ellos los mismos eventos de modo no aleatorio.

El conjunto de los escenarios y los eventos forman una matriz binaria de incidencias $X=(x_{ij})$ de dimensiones $N \times M$ con los escenarios en las filas y los eventos en las columnas. Esta matriz es binaria con sus elementos x_{ij} iguales a 0 o 1 indicando si el evento X_j ocurre en el escenario i -ésimo.

Con la matriz de incidencias puede obtenerse la matriz simétrica de coincidencias C de tamaño $M \times M$ a través de la operación $C = X^T X$, donde X^T es la matriz transpuesta de X . Cada elemento c_{ij} representa el número de escenarios en los que X_i y X_j tienen el valor 1, es decir, coinciden. Esta matriz de incidencias nos permite obtener las siguiente tres métricas probabilísticas:

La primera es la probabilidad de que suceda un evento X_i que se obtiene dividiendo el número de veces que sucede el evento entre el número de escenarios:

$$P(X_i) = \frac{c_{ii}}{N}$$

También disponemos de la probabilidad conjunta de dos eventos X_i y X_j , denotada por $P(X_i \cap X_j)$, que viene dada por la frecuencia de ocurrencia de los dos eventos en el mismo escenario dividido entre el número de escenarios:

$$P(X_i \cap X_j) = \frac{c_{ij}}{N}$$

Finalmente, tenemos la probabilidad condicional de que sucedan dos eventos X_i y X_j , denotada por $P(X_i | X_j)$, que expresa la probabilidad de que ocurra un evento cuando un segundo evento ya ha ocurrido y se obtiene dividiendo las probabilidades de cada evento:

$$P(X_i | X_j) = \frac{P(X_i \cap X_j)}{P(X_j)} = \frac{c_{ij}}{c_{jj}}$$

De esta forma podemos hablar de que dos eventos coinciden en probabilidad basándonos en el concepto de eventos independientes. Dos eventos X_i y X_j , son independientes cuando:

$$P(X_i | X_j) = P(X_i) \leftrightarrow \frac{c_{ij}}{c_{jj}} = \frac{c_{ii}}{N}$$

Lo cual se cumple si la frecuencia c_{ij} correspondiente a la probabilidad conjunta de X_i y X_j viene dada por:

$$c_{ij}^* = \frac{c_{ii}c_{jj}}{N} \text{ donde } c_{ij}^* \text{ es la frecuencia esperada}$$

Así diremos que dos eventos X_i y X_j son coincidentes en probabilidad (tienen grado de dependencia) si:

$$c_{ij} > \frac{c_{ii}c_{jj}}{N} = c_{ij}^*$$

La diferencia entre c_{ij} y c_{ij}^* toma una distribución normal con el siguiente error estándar (Haberman, 1973):

$$\sqrt{\left(\left(1 - \frac{c_{ii}}{N}\right) - \left(1 - \frac{c_{jj}}{N}\right)\right)}$$

Este error puede utilizarse para estandarizar la diferencia entre el valor empírico de eventos coincidentes c_{ij} y la frecuencia esperada c_{ij}^* suponiendo que son independientes y obtener así el residuo de Haberman ($r_{jk} \sim N(0,1)$).

$$r_{ij} = \frac{c_{ij} - c_{ij}^*}{\left(\left(1 - \frac{c_{ii}}{N}\right) - \left(1 - \frac{c_{jj}}{N}\right)\right)^{\frac{1}{2}}}$$

Así se conforma la matriz de adyacencias A de dimensiones $M \times M$ donde dos eventos X_i y X_j se dice que son coincidentes si cumplen la siguiente norma:

$$A[i, j] = 1 \leftrightarrow P(r_{ij} \leq 0) < c \wedge i \neq j \text{ donde } c \text{ es el nivel de significación}$$

El valor de c debe ser la probabilidad de concluir que una cierta hipótesis es falsa cuando resulta que es cierta (Error de tipo I) con c supuesto pequeño. Por ello la hipótesis debe ser que ($r_{ij} > 0$), es decir la hipótesis es que $c_{ij} > c_{ij}^*$, lo cual equivale a que X_i y X_j son coincidentes en probabilidad. Así $A[i, j] = 1$ cuando la probabilidad de que r_{ij} sea menor o igual a cero es pequeña. Es decir que lo muy probable es que $c_{ij} > c_{ij}^*$.

Alternativamente, se podría obtener la matriz S de similitudes de escenarios mediante la fórmula $S = XX^T$, en cuyo caso aparecería los elementos s_{pq} , que indicaría cuántos eventos iguales comparten los escenarios p y q . De modo análogo, podrían calcularse el residuo de Haberman y otra matriz de adyacencias A de dimensiones $N \times N$ a fin de detectar cuando hay una similitud entre escenarios.

A partir de la matriz de adyacencias A se elabora una red en la que los eventos son los nodos y sus vínculos los valores de los residuos de Haberman. En este artículo, los escenarios son las revistas JCR de sociología y los autores

que publican los artículos en ellas sería los eventos. De esta manera la matriz de adyacencias *A* indica qué revistas son semejantes en los autores que en ella publican usando el residuo de Haberman para indicar la fuerza de la relación.

7. ANÁLISIS DESCRIPTIVO DE LAS REVISTAS JCR DE SOCIOLOGÍA

Las revistas JCR de sociología conforman un conjunto de publicaciones especializadas controladas por editoriales en múltiples idiomas. Las tablas a continuación (tabla 1) muestran la distribución de las revistas por países, idiomas y editoriales. De las 163 revistas 67 son de Estados Unidos y 54 de Reino Unido. El resto de las revistas provienen de otros 21 países, la mayor parte del continente europeo, aunque también de Asia, América y Oceanía.

El 90% de las revistas publican en inglés, aunque también las hay con artículos en distintos idiomas simultáneamente. A nivel editorial, dominan los grandes conglomerados editoriales internacionales como SAGE, Taylor & Francis, Elsevier, Blackwell o Routledge. A estos los acompañan, con menor representación, editoriales universitarias como las de Oxford, Cambridge y Chicago.

Tabla 1. Distribución de revistas por países de origen, idioma y por editoriales

<i>País</i>	<i>Revistas</i>	<i>%</i>	<i>Idioma</i>	<i>Revistas</i>	<i>%</i>
Estados Unidos	67	41.1%	Inglés	147	90.2%
Reino Unido	54	33.1%	Alemán	5	3.07%
Países bajos	10	6.13%	Frances	3	1.84%
Alemania	7	4.29%	Español	3	1.84%
Canadá	3	1.84%	Otros	5	3.07%
Otros	22	13.5%			

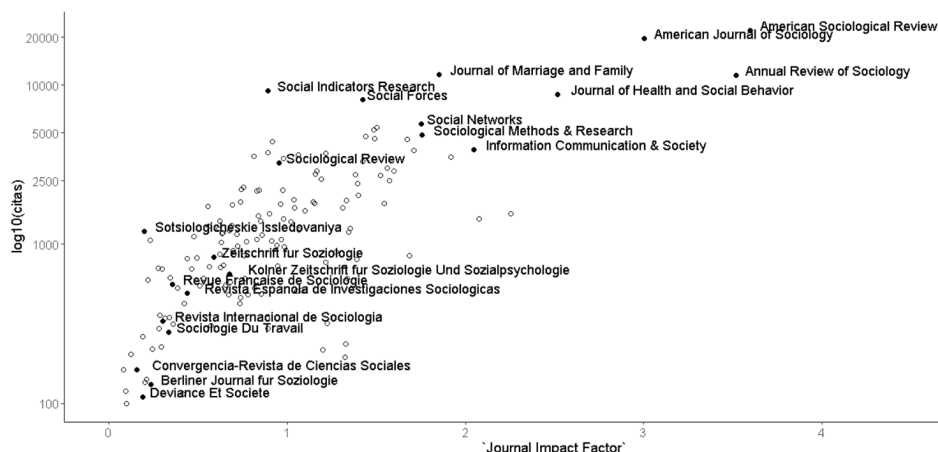
<i>Editoriales</i>	<i>Revistas</i>	<i>%</i>
SAGE	40	24.5%
Taylor & Francis	13	7.98%
Blackwell Publishing Inc	12	7.36%
Elsevier Ltd.	12	7.36%
Routledge	6	3.68%
Wiley	6	3.68%
Oxford University Press	5	3.07%
Cambridge University Press	4	2.45%
Kluwer Academic Press	4	2.45%
Springer Verlag	4	2.45%
University of Chicago Press	4	2.45%
Otras	53	32.5%

Khun (1962:19) asocia la aparición de revistas especializadas en una disciplina a momentos de cambio de paradigma. En la sociología las publicaciones pioneras son *American Journal of Sociology*, creada en 1895 por Albion W. Small en la Universidad de Chicago y *Sociological Review* fundada en Reino Unido en 1908 por Leonard T. Hobhouse. Tras estas publicaciones aparecen otras tres americanas, *Social Forces* en 1922, *American Sociological Review* en 1936 y el *Journal of Marriage and Family* en 1939. La figura 2 muestra las revistas por década de aparición. En la década de los 60 hay un aumento importante de aparición de revistas, este aumento alcanza su máximo en la década de los años 70 con más de 40 revistas. En esta década encontramos *Annual Review of Sociology*, *Sociological Methods and Research* y *Social Networks*. Un 40% de las revistas aparecen en las tres últimas décadas, aunque en la década más reciente apenas empiezan su andadura 7 revistas.

Figura 2. Distribución de revistas por década de aparición



Las publicaciones quedan clasificadas de acuerdo con su influencia en la disciplina a través de métricas de impacto. La métrica más sencilla es el número total de citas, JCR lo calcula teniendo en cuenta las citas entre las revistas disponibles en su base de datos. Sin embargo, la métrica por excelencia para medir el impacto es el *Journal Impact Factor (JIF)* que se calcula anualmente al dividir las citas de la revista ese año por el número de artículos en los dos años anteriores. En la figura 3 representamos las revistas con su impacto en el eje horizontal y las citas en el eje vertical. Aparece para cada revista el promedio de los JIF anuales y el total de citas más reciente. Solo han quedado representadas aquellas revistas con más de 100 citas con un eje vertical expresado en escala logarítmica en base 10 para ajustar las importantes diferencias en el total de citas entre las revistas.

Figura 3. Revistas organizadas por factor de impacto y número de citas

En la parte superior derecha del gráfico aparecen las revistas pioneras y generalistas. *American Sociological Review* y *American Journal of Sociology* ocupan los más lugares destacados. Junto a estas dos aparece *Annual Review of Sociology*, uno de los denominados *review journals*, que por este motivo obtienen un número elevado de citas (Moed, 2005). Tras estas publicaciones aparecen revistas temáticas tales como *Journal of Marriage and Family* y *Journal of Health and Social Behaviour* que incorporan investigación de interés actual sobre el género, la familia la salud y la medicina. Más centrada en el gráfico aparece *Social Forces*, una de las revistas prominentes en el campo y junto a *American Sociological Review* y *American Journal of Sociology* perteneciente a la denominada “Triple Corona”. Su diferencia de citas e impacto con las otras dos posiblemente sea debido a la capacidad de las otras dos de distanciarse de las demás (Jacobs, 2016). Alrededor de *Social Forces* se encuentra *Information Communication & Society*, revista que comienza en 2001 y centrada en temáticas de creciente interés como son los estudios sobre la sociedad de la comunicación y el impacto de las nuevas tecnologías. Las otras tres revistas de este grupo, todas de los años 70, presentan una aproximación empírica y metodológica. Son *Sociological Methods and Research*, *Social Networks* y *Social Indicators Research*. Un poco más abajo aparece *Sociological Review*, otra de las pioneras pero que no llega a tener los indicadores de influencia de ASR y AJS. Las publicaciones en otros idiomas como el alemán, francés o castellano aparecen en la parte inferior izquierda.

Explorar la procedencia y el impacto de las revistas es un buen punto de partida. Sin embargo, es necesario incluir en el análisis a los autores e instituciones de las revistas. Es pues preciso bajar un nivel de profundidad y estudiar los artículos. Para ello se emplean los datos de MAG filtrando los artículos de las 163 revistas JCR de sociología de 1997 a 2018. Hay cinco revistas para las

que no hay artículos en MAG en estos años: *Ethology and Sociobiology*, *Innovation*, *Australian and New Zealand Journal of Sociology* y *Studies in Symbolic Interaction*. En total se cuenta con 164.036 artículos que han sido revisados para eliminar las recensiones de libros eliminando aquellos títulos que contienen las palabras “book review”. Al final de este proceso se obtuvieron 139.452 artículos. Finalmente, tras descartar los artículos repetidos se consiguió la cifra final de 137.178 artículos de estas 158 revistas entre 1997 y 2018.

En las tablas de abajo (tabla 2) se presentan los listados de las diez revistas con más artículos y citas por artículo. La revista *Contemporary Sociology* cuenta con 6.648 representando casi el 5% del total. *American Sociological Review* es la revista con la media más alta de citas por artículo con 119, además sus citas representan un 4.7% del total de citas en esta base de datos.

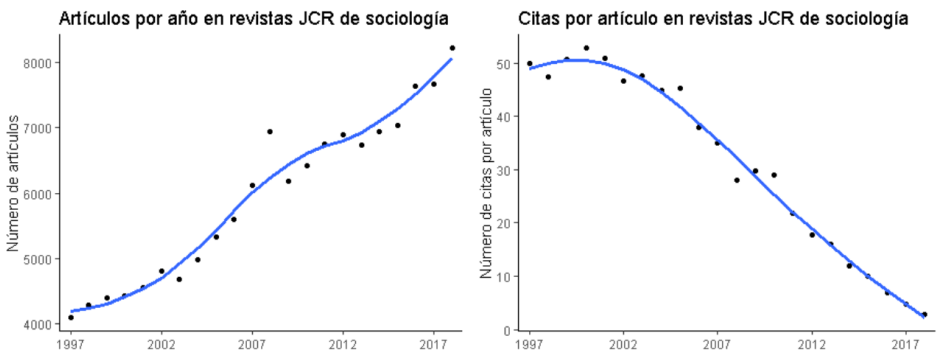
Tabla 2. Arriba las diez revistas con más artículos en los datos y abajo la diez revistas con la media de citas por artículos más alta en los datos

Revistas	Número de artículos en MAG	Proporción con respecto al total
Contemporary Sociology	6.648	4.85%
Social Indicators Research	3.904	2.85%
American Journal of Sociology	3.482	2.54%
Ethnic and Racial Studies	2.901	2.11%
Journal of Marriage and Family	2.372	1.73%
Society	2.028	1.48%
Human Ecology	1.956	1.43%
Social Forces	1.906	1.39%
Social Science Quarterly	1.782	1.3%
The Sociological Review	1.779	1.3%

Revistas	Número de citas en MAG	Media de citas por artículo	Proporción de citas con respecto al total
<i>American Sociological Review</i>	128.115	119	4.69%
Journal of Health and Social Behavior	60.218	91.5	2.20%
Sociology Of Education	31.323	68.1	1.15%
Journal of Marriage and Family	134.822	56.8	4.93%
Gender & Society	46.993	53.8	1.73%
Sociological Methodology	16.599	53.0	0.61%
Social Networks	44.956	50.9	1.65%
Social Problems	34.936	49.9	1.28%
Evolution and Human Behavior	54.333	46.6	2.00%
Population and Development Review	42.699	45.8	1.57%

De aquí en adelante, y con el fin de intuir mejor las tendencias temporales, representaremos las series temporales con un método de curva suavizada de ajuste de regresión polinómica. En los gráficos siguientes (figura 4) se representa la evolución anual del número de artículos y citas de las 158 revistas de sociología. En 20 años el número de artículos se duplica y las medias más altas de citas por artículo se acumulan en los primeros años.

Figura 4. Evolución anual del número de artículos y de citas. Las curvas utilizan el método de suavizado de ajuste de regresión polinómica.



Al fijarnos en los artículos con más citas (tabla 3), resalta que hay 3 de ellos del *American Journal of Sociology*, que la mitad son de 1997 o 1998 y que presentan una diversidad de enfoques y temáticas desde temas de salud, a elementos metodológicos y teóricos.

Tabla 3. Los diez artículos con más citas en MAG

Citas acumuladas	Referencias
5.901	Idler, Ellen L., and Yael Benyamini. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." <i>Journal of Health and Social Behavior</i> 38, no. 1 (1997): 21-37.
4.539	Burnham, K. P., & Anderson, D. R. (2004). "Multimodel Inference: Understanding AIC and BIC in Model Selection". <i>Sociological Methods & Research</i> , 33(2), 261–304.
3.700	Connell, R. W., & Messerschmidt, J. W. (2005). "Hegemonic Masculinity: Rethinking the Concept". <i>Gender & Society</i> , 19(6), 829–859.
3.325	Woolcock, Michael. "Social Capital and Economic Development: Toward a Theoretical Synthesis and Policy Framework." <i>Theory and Society</i> 27, no. 2 (1998): 151-208.
2.878	Inglehart, Ronald, and Wayne E. Baker. "Modernization, Cultural Change, and the Persistence of Traditional Values." <i>American Sociological Review</i> 65, no. 1 (2000): 19-51

Citas acumuladas	Referencias
2.730	Burt, Ronald S. "Structural holes and good ideas." <i>American Journal of Sociology</i> 110, no. 2 (2004)
2.647	Heckathorn, Douglas D. "Respondent-driven Sampling: a New Approach to the Study of Hidden Populations." <i>Social Problems</i> 44, no. 2 (1997): 174-199.
2.478	Meyer, John W., John Boli, George M. Thomas, and Francisco O. Ramirez. "World society and the nation-state." <i>American Journal of sociology</i> 103, no. 1 (1997): 144-181.
2.463	Reckwitz, Andreas. "Toward a theory of social practices: A development in culturalist theorizing." <i>European Journal of Social Theory</i> 5, no. 2 (2002): 243-263.
2.359	Emirbayer, Mustafá, and Ann Mische. "What is Agency?". (1998) <i>American Journal of Sociology</i> . Vol 13, no. 4, Jan 1998. 962-1023.

Los datos de MAG proporcionan información de los 126.744 autores presentes en la selección de artículos. Además, hay 4.981 instituciones relacionadas con estos autores. Esta información aparece en un 37% de las relaciones entre autor y artículo. Las tablas 4 y 5 presentan los autores e instituciones con más citas. Paul R. Amato, de la Universidad de Pennsylvania, es el autor con más citas, seguido de Robert J. Sampson de la Universidad de Harvard, Alejandro Portes de la Universidad de Princeton, Ed Diener de la Universidad de Illinois, Tom A. B. Snijders de la Universidad de Oxford y Ellen L. Idler de la Universidad de Emory. Entre las instituciones con más citas encontramos un predominio de universidades americanas de gran prestigio como Michigan, Pennsylvania, Cornell, Ohio, Texas, Harvard, California (Berkeley y UCLA) o University of Wisconsin-Madison.

Tabla 4. Autores con más citas

Autor	Número de artículos	Número de citas	Revistas distintas	Media de citas por artículo	Artículo con más citas de cada autor
Paul R. Amato	48	8.163	7	170.0	Amato, Paul R. "The consequences of divorce for adults and children." <i>Journal of Marriage and Family</i> 62, no. 4 (2000): 1269-1287.
Robert J. Sampson	23	7.306	11	317.6	Sampson, Robert J., and Stephen W. Raudenbush. "Systematic social observation of public spaces: A new look at disorder in urban neighborhoods." <i>American Journal of Sociology</i> 105, no. 3 (1999)

<i>Autor</i>	<i>Número de artículos</i>	<i>Número de citas</i>	<i>Revistas distintas</i>	<i>Media de citas por artículo</i>	<i>Artículo con más citas de cada autor</i>
<i>Alejandro Portes</i>	43	6.322	16	147.0	Portes, Alejandro, Luis E. Guarnizo, and Patricia Landolt. "The study of transnationalism: pitfalls and promise of an emergent research field." <i>Ethnic and Racial Studies</i> 22, no. 2 (1999): 217-237.
<i>Ed Diener</i>	23	6.314	2	274.5	Diener, Ed, and Eunkook Suh. "Measuring quality of life: Economic, social, and subjective indicators." <i>Social Indicators Research</i> 40, (1997)
<i>Tom A. B. Snijders</i>	43	6.239	6	145.0	Snijders, Tom AB, Gerhard G. Van de Bunt, and Christian EG Steglich. "Introduction to stochastic actor-based models for network dynamics." <i>Social networks</i> 32, no. 1 (2010): 44-60.
<i>Ellen L. Idler</i>	7	6.078	3	868.2	Idler, Ellen L., and Yael Benyamini. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." <i>Journal of Health and Social Behavior</i> 38, no. 1 (1997): 21-37.
<i>Stephen P. Borgatti</i>	22	5.994	3	272.4	Borgatti, Stephen P. "Centrality and network flow." <i>Social Networks</i> 27, no. 1 (2005): 55-71.
<i>Douglas D. Heckathorn</i>	10	5.723	5	572.3	Heckathorn, Douglas D. "Respondent-driven sampling: a new approach to the study of hidden populations." <i>Social Problems</i> 44, no. 2 (1997)
<i>Douglas S. Massey</i>	73	5.506	19	75.4	Massey, Douglas S., and Kristin E. Espinosa. "What's driving Mexico-US migration? A theoretical, empirical, and policy analysis." <i>American Journal of Sociology</i> 102, no. 4 (1997): 939-999.
<i>Michel Callon</i>	17	5.169	5	304.0	Callon, Michel. "Introduction: the embeddedness of economic markets in economics." <i>The Sociological Review</i> 46, no. 1_suppl (1998): 1-57.

Tabla 5. Instituciones con más citas

<i>Institución</i>	<i>Número de artículos</i>	<i>Número de citas</i>	<i>Número de autores</i>	<i>Media de citas por artículo</i>
<i>University of Michigan</i>	1346	72.135	722	53.59
Pennsylvania State University	1537	65.473	645	42.60
Cornell University	1148	52.972	488	46.14
Ohio State University	1034	52.611	440	50.88
University of Texas at Austin	1138	47.811	502	42.01
University of Wisconsin-Madison	1050	43.810	572	41.72
University of California, Los Angeles (UCLA)	971	43.753	520	45.06
Harvard University	924	43.558	534	47.14
University of North Carolina at Chapel Hill	857	42.833	424	49.98
University of California, Berkeley	836	38.743	497	46.34

7.1. Análisis de citas, coautoría y proporción de género por tipos de revistas

En los anteriores análisis descriptivos hemos visto que hay diferentes tipos de revistas: las revistas más antiguas, las temáticas, las metodológicas o las de lengua no inglesa. ¿Tienen estos tipos de revistas distintos patrones de citas, coautoría o proporción de género? Para averiguarlo organizamos algunas de las revistas en los cuatro tipos que indicamos en la tabla abajo (tabla 6). Incluimos las tres revistas de la triple corona, todas las revistas de habla no inglesa y todas las revistas puramente metodológicas. Para las revistas temáticas seleccionamos las trece primeras en cuanto a su factor de impacto y número de citas en JCR.

Tabla 6. Agrupación de algunas revistas en cuatro grupos

Revistas de la triple corona	Revistas temáticas
American Sociological Review	Ethnic and Racial Studies
American Journal of Sociology	Gender & Society
Social Forces	Information Communication & Society
Revistas de lengua no inglesa	International Journal of Intercultural Relations
Archives Européennes de Sociologie	International Political Sociology
Berliner Journal für Soziologie	Journal for The Scientific Study of Religion
Convergencia-Revista de Ciencias Sociales	Journal of Consumer Culture
Deviance Et Societe	Journal of Health and Social Behavior
Drustvena Istrazivanja	Journal of Marriage and Family
Filosofija-Sociologija	Society & Natural Resources

Kolner Zeitschrift fur Soziologie Und Sozialpsychologie	Socio-Economic Review
Revista Española de Investigaciones Sociológicas	Sociology of Education
Revista Internacional de Sociología	Sociology of Health & Illness
Revue Française de Sociologie	Revistas de métodos
Sociologicky Casopis	Journal of Mathematical Sociology
Sociologie Du Travail	Qualitative Research
Sociologija I Prostor	Qualitative Sociology
Sociologisk Forskning	Social Indicators Research
Sotsiologicheskie Issledovaniya	Social Networks
Soziale Welt-Zeitschrift fur Sozialwissenschaftliche Forschung Und Praxis	Sociological Methodology
Zeitschrift fur Soziologie	Sociological Methods & Research
	Journal of Mathematical Sociology

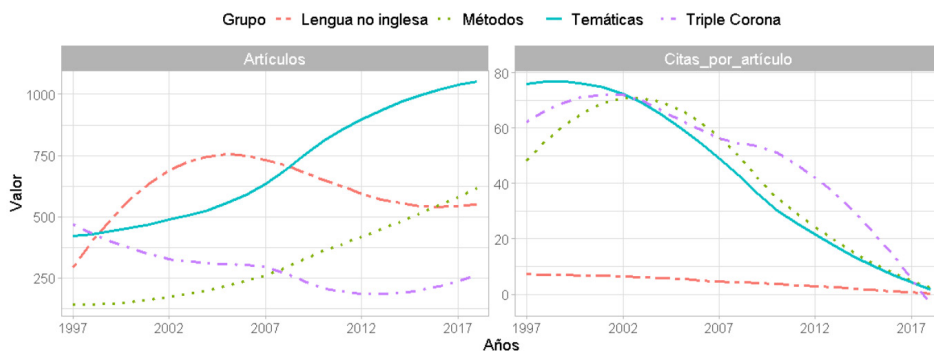
Para cada uno de estos grupos se obtuvieron el número de artículos y la media de citas por artículo (tabla 7). Las revistas de la triple corona son las que poseen la media más alta de citas por artículo, casi 55, seguidas de las temáticas y las de métodos. Las revistas de lengua no inglesa tienen la media de citas por artículo más baja con tan solo 4,5.

Tabla 7. Artículos, media de citas por artículo y desviación estándar por grupo de revistas

Grupo	Artículos	Media de citas por artículo	Desviación estándar
Triple corona	5.984	54,6	142,0
Temáticas	14.531	37,3	98,9
Métodos	6.294	34,2	108,0
Lengua no inglesa	12.719	4,5	16,4

Al representar anualmente los artículos y la media de citas por artículo (figura 6) puede apreciarse que el número de artículos de las revistas de la triple corona desciende con los años. Este efecto se debe principalmente al descenso de artículos publicados en la *American Journal of Sociology*. Por otro lado, el número de artículos de las revistas temáticas y de métodos aumenta debido al aumento de artículos en revistas como *Social Indicators Research* o *Ethnic and Racial Studies*. Basándose en las citas por artículo, las revistas de los tres grupos tienen tendencias similares a primera vista. Por su lado, las revistas de lengua no inglesa alcanzan su máximo de artículos en torno al año 2005, pero como ya se ha comentado están lejos de las revistas americanas o británicas en citas por artículo.

Figura 6. Evolución del número de artículos y citas por artículo por grupos de revistas. Las curvas utilizan el método de suavizado de ajuste de regresión polinómica.



Para comprobar si las diferencias en la media de citas por artículo entre los cuatro grupos de revistas son significativas se empleó el análisis de varianza (ANOVA) para probar la hipótesis nula de que no hay diferencias entre las medias de los diferentes grupos. Los resultados de ANOVA entre grupos verifican que la diferencia de medias es estadísticamente significativa ($p < .05$) con una $F_{3,39,524}$ de 494,4. La prueba post hoc de Tukey (tabla 8) señala que todos los grupos difieren significativamente ($p < .05$), excepto la media de citas por artículo de las revistas temáticas y las de métodos. Las revistas de la triple corona son las que obtienen la media de citas por artículos más alta y con mayor diferencia que las otras revistas en los otros grupos.

Tabla 8. Resultados de la prueba de Tukey para las medias de citas

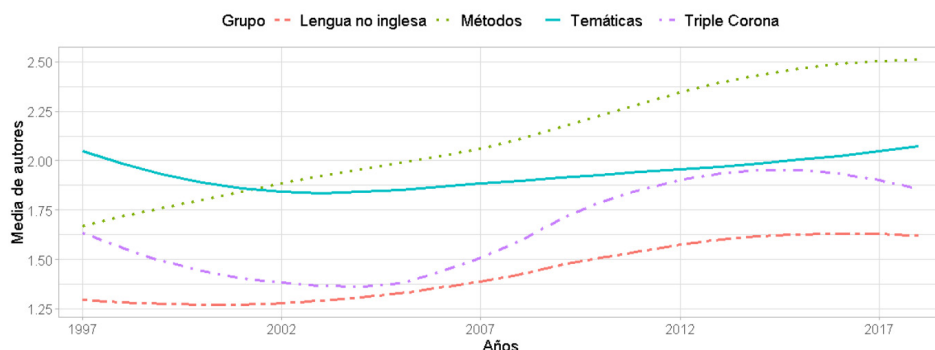
<i>Resultados dos a dos</i>	<i>Diferencia</i>	<i>p-valor</i>
Métodos - Lengua no inglesa	29,73	0
Temáticas - Lengua no inglesa	32,84	0
Triple Corona - Lengua no inglesa	50,05	0
Temáticas – Métodos	3,1	0,11
Triple Corona – Métodos	20,13	0
Triple Corona - Temáticas	17,21	0

El número de coautores por artículo se ha ido incrementando en todas las disciplinas. En sociología es algo cada vez más común motivado por razones de competitividad y progresión académica, el aumento de las oportunidades de colaboración o la propia naturaleza de la investigación interdisciplinar (Taylor & Francis, 2017). Al calcular la media de autores por artículo y por grupo de revistas (tabla 9), pueden observarse diferencias entre los distintos grupos. Las revistas de métodos con más de 2 autores por artículo superan a las temáticas con 1,9 y a las de la triple corona con 1,6.

Tabla 9. Media de coautores por artículo y desviación estándar por grupo de revistas

Grupo	Media de autores por artículo	Desviación estándar
<i>Triple Corona</i>	1,59	0,99
Temáticas	1,94	1,52
Métodos	2,22	1,52
Lengua no inglesa	1,42	0,81

El gráfico de evolución anual (figura 7) muestra un crecimiento en la media de coautores para los todos los grupos a excepción de las revistas de métodos que parten de una media de dos autores y la mantienen en el tiempo. Las revistas temáticas son las que experimentan mayor aumento pasando de 1,6 autores por artículo a casi 2,5.

Figura 7. Media anual de coautores por grupos de revistas

De nuevo, se comprueba que existen diferencias significativas en la media de autores por artículo entre los cuatro grupos de revistas utilizando el análisis de varianza (ANOVA) para probar la hipótesis nula de que no hay diferencias entre las medias de autores de los diferentes grupos. Los resultados de ANOVA entre grupos verifican la diferencia estadísticamente significativa ($p < .05$) de estas medias ($F_{3,39,400} = 737$, $p < 2e-16$). La prueba post hoc de Tukey (tabla 10) señala que todos los grupos difieren significativamente ($p < .05$). Las revistas de métodos son las que obtienen la media de autores por artículos más alta superando en 0.6 coautores por artículo a las de la triple corona, 0.23 a las temáticas y en 0.74 a las de lengua no inglesa.

Tabla 10. Resultados de la prueba de Tukey para las medias de autores

<i>Resultados dos a dos</i>	<i>Diferencia</i>	<i>P adj</i>
<i>Métodos - Lengua no inglesa</i>	0,79	0
Temáticas - Lengua no inglesa	0,51	0
Triple Corona - Lengua no inglesa	0,17	0
Temáticas – Métodos	-0,28	0
Triple Corona – Métodos	-0,62	0
Triple Corona - Temáticas	-0,34	0

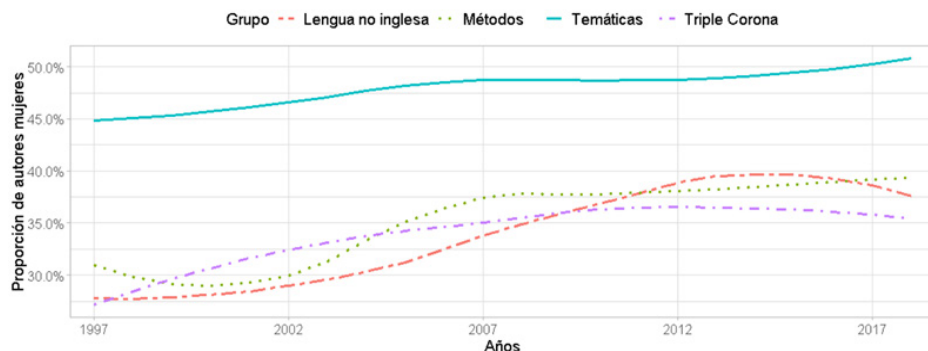
Otra de las cuestiones que puede abordarse es la referente a la existencia de sesgo de género en las publicaciones de estas revistas. A través de los nombres de los autores se puede asignar un género con el fin de estudiar la proporción de mujeres en las publicaciones. Del total de 19.985 nombres únicos de autores se ha dispuesto del género en 18.341 casos, un 92%. Un 44% de los autores de los artículos son mujeres. La proporción de género es distinta en los artículos de los cuatro grupos de revistas (tabla 11) con las revistas temáticas con la proporción de mujeres más alta.

Tabla 11. Proporción media de autores de género femenino por artículo y desviación estándar por grupo de revistas

<i>Grupo</i>	<i>Proporción media de autoras por artículo</i>	<i>Desviación estándar</i>
<i>Triple Corona</i>	0,329	0,425
Temáticas	0,477	0,436
Métodos	0,356	0,401
Lengua no inglesa	0,321	0,431

En la figura 8 puede observarse el aumento en la proporción de autores de género femenino desde 1997 para los cuatro grupos y la importante diferencia de las revistas temáticas con las de los otros tres grupos. Revistas como *Gender & Society* cuentan con una proporción del 83% de mujeres autores o *Journal of Marriage and Family* con un 60%. Llama también la atención encontrar dentro de las revistas de métodos dos con una proporción de mujeres por encima de las demás en el grupo, se trata de *Qualitative Sociology* y *Qualitative Research*.

Figura 8. Evolución anual de la proporción de autores de género femenino para los cuatro grupos de revistas



El análisis de varianza ANOVA nos permite ver si la diferencia entre las medias de citas y de autores es significativa, pero no nos vale para la diferencia entre proporciones de género. Para este caso usamos una prueba de proporciones con la que comprobamos si la diferencia entre las proporciones de autores de género femenino de los distintos grupos es estadísticamente significativa. Utilizamos la hipótesis nula de que las proporciones son iguales entre los grupos. El resultado de la prueba confirma las diferencias significativas ($p < .05$) entre las proporciones de los cuatro grupos. Las revistas temáticas son las que obtienen la proporción más alta de mujeres por artículo superando en un 15% a las de la triple corona y las de lengua no inglesa y en un 12% a las temáticas.

8. LA RED DE AFILIACIÓN ENTRE AUTORES Y REVISTAS

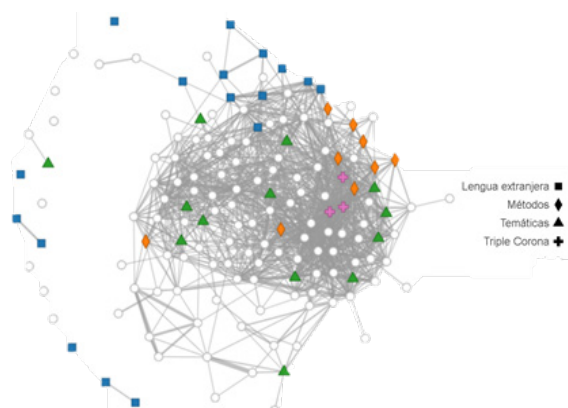
Con el fin de entender mejor la relación entre los autores y las revistas JCR de sociología, se representarán sus relaciones a través de una red de afiliación donde los actores son los autores de las publicaciones y los grupos las revistas donde publican. A través de esta red podrán detectarse aquellos grupos de revistas con comunidades comunes de autores.

Para los cálculos del análisis de coincidencias eliminamos los artículos de *Contemporary Sociology* para evitar el efecto que pueda tener una revista con una proporción de artículos de revisión elevada, pero con un impacto bajo. Así partimos de la matriz de incidencias X que tiene 127.368 filas que representan a los autores y 157 columnas, una por revista. Los valores x_{ij} de esta matriz toman el valor 1 si el autor i -ésimo ha publicado alguna vez en la revista j -ésima y cero en caso contrario. A partir de esta matriz binaria de incidencias X se aplica el análisis de coincidencias y se calcula la matriz de adyacencias A . La red de afiliación pasa de ser bipartita a modo uno formada por nodos que representan las 158 revistas cuyas relaciones, establecidas mediante el residuo de Haberman,

representan la fuerza de la relación entre las revistas. Cuantos más autores hayan publicado en dos revistas, más fuerte será su relación. Para simplificar la red solo se tienen en cuenta aquellas relaciones con Haberman > 3 asegurando que las relaciones son significativas con un nivel de confianza mayor del 99,8% en pruebas de una sola cola.

La red resultante se muestra en la figura 9. Los nodos representan las revistas JCR y las aristas representan los enlaces que establecen la relación entre las revistas. Utilizamos el algoritmo basado en fuerzas de atracción repulsión de redes de Fruchterman-Reingold (Fruchterman y Reingold, 1991). Esta red consta de un componente central con 139 nodos conectados y otros 18 nodos separados. Las formas de los nodos representan los grupos de la sección anterior. Todas las revistas de la triple corona y todas las de métodos forman parte del componente central mientras que una de las temáticas y varias de las revistas de lengua no inglesa están fuera de él. Las tres revistas más prestigiosas se encuentran próximas a la derecha del componente central. Las revistas de métodos se agrupan arriba a la derecha del componente principal, aunque dos de ellas (las cualitativas) aparecen en otras zonas de este mismo componente. Las revistas temáticas están distribuidas por varias zonas sin una agrupación clara. Finalmente, las revistas de lengua no inglesa están en la parte superior de la red y siete de ellas desconectadas del componente central.

Figura 9. Red de afiliación de revistas utilizando la forma del nodo para representar los grupos de revistas



Se calcularon dos medidas de centralidad de redes: el grado con pesos y la centralidad de intermediación. En la red de la figura 10 los colores de los nodos representan el grado ponderado en una escala de color. Las revistas de la triple corona tienen todos valores altos, siendo *Social Forces* la revista con mayor número de enlaces con otras revistas. Al fijarnos en la tabla 12 con las quince revistas con mayor grado con peso identificamos revistas americanas y británicas

Figura 10. Redes de afiliación con escalas de color para representar el grado con pesos de los distintos nodos. La red de la derecha es la misma que la de la izquierda, pero con centrada en los nodos con mayor grado con pesos junto de sus nombres.

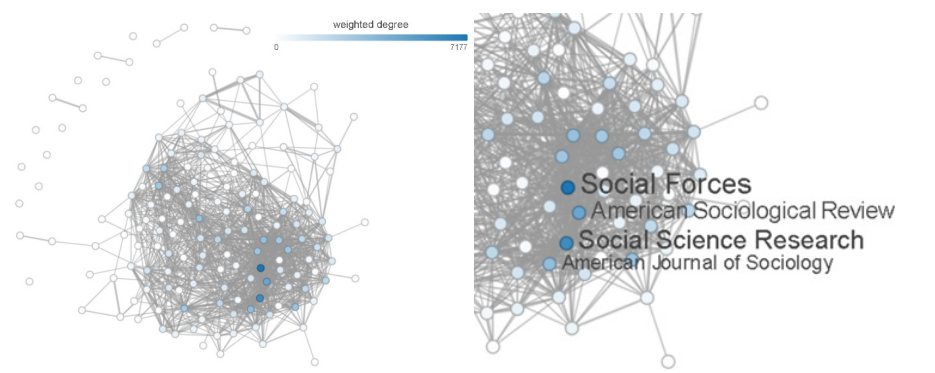


Tabla 12. Las quince revistas con mayor grado con pesos de la red de afiliación

Revista	Grado con pesos
Social Forces	7.177
Social Science Research	6.079
American Sociological Review	4.694
American Journal of Sociology	3.447
Sociological Forum	3.284
British Journal of Sociology	3.142
Journal of Marriage and Family	3.132
Sociological Quarterly	3.014
Social Problems	2.960
Sociology	2.856
European Sociological Review	2.850
Sociological Perspectives	2.813
The Sociological Review	2.546
Sociology Compass	2.157
Ethnic and Racial Studies	2.066

En la figura 11 los colores de los nodos denotan el grado de intermediación y la tabla 13 presenta los datos de las primeras 20 revistas. Los nodos con mayor

grado de intermediación son *British Journal of Sociology*, *European Societies* y *Sociological Theory*. Estos nodos tienen un papel importante en la red ya que suelen actuar como controladores de flujo de información al ser puentes entre grupos.

Figura 11. Redes de afiliación con escalas de color para representar la intermediación de los distintos nodos. La red de la derecha es la misma que la de la izquierda, pero centrada en los nodos con mayor grado de intermediación junto con sus nombres.

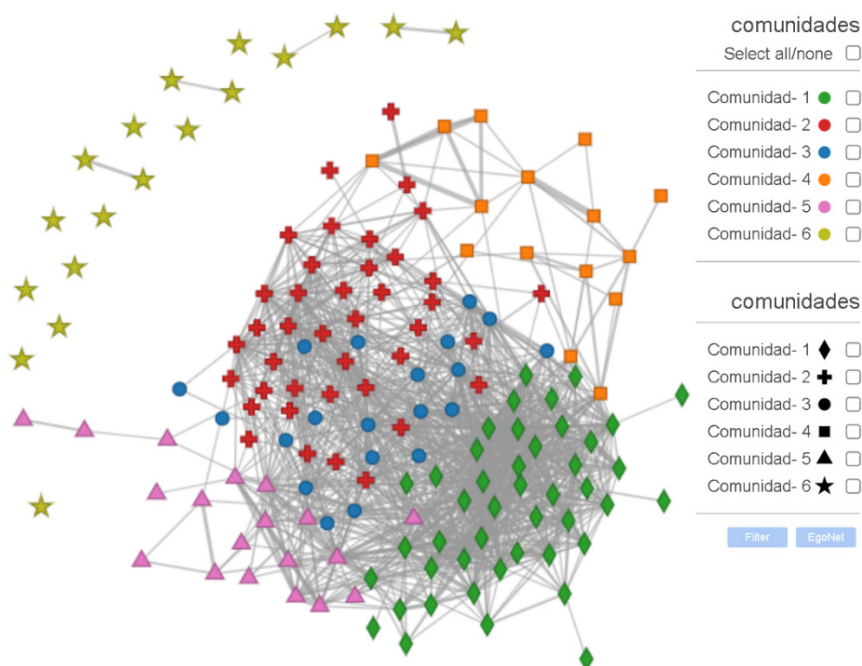


Tabla 13. Las quince revistas con mayor grado de intermediación de la red de afiliación

Revista	Centralidad de intermediación
British Journal of Sociology	611,02
European Societies	478,88
Sociological Theory	474,16
Social Forces	440,56
Human Ecology Review	420,08
International Sociology	370,58
Social Science Research	367,06
Sociological Perspectives	354,68
Rural Sociology	346,39
American Sociological Review	325,63
European Sociological Review	320,33
Sociological Quarterly	306,29
Research in Social Stratification and Mobility	305,68
European Journal of Social Theory	289,02
Sociological Forum	281,95

Tras estos análisis, se aplicó el algoritmo de Louvain para el cálculo de comunidades (Blondel et al., 2008) y se obtuvieron 19 comunidades en la red de afiliación. Como 5 de las comunidades acumulan el 90% de los nodos las restantes se agruparon en una única comunidad (Comunidad-6). La figura 12 presenta la disposición de estas seis comunidades.

Figura 12. La red de afiliación con la forma y color de los nodos identificando las comunidades obtenidas por el algoritmo de Louvain.



A continuación, se describirán cada una de las comunidades. Para ello se representarán en la figura 13 las subredes mediante el grado como tamaño de los nodos y un degradado de color para el factor de impacto de los nodos. Los enlaces también utilizan un degradado de color que representa el número de autores coincidentes entre las revistas.

La comunidad 1, situada a la derecha de la red, es la más extensa con 45 nodos. Son los nodos con mayor factor de impacto, con predominancia de revistas americanas, especialmente las más tradicionales. Las revistas de la triple corona ocupan un lugar central en esta comunidad que incluso contiene la mayoría de las revistas de métodos cuantitativos. Además, contiene revistas temáticas enfocadas a la religión, la sociología militar, el género, la familia, la juventud y la salud. La comunidad 2, es la siguiente en tamaño con 40 nodos y aparece a

la izquierda del componente principal de la red. Esta comunidad está dominada por cuatro revistas del Reino Unido: *British Journal of Sociology*, *Sociology*, *Sociological Review* y *Sociological Research Online*. Cuenta con una revista de métodos cualitativa *Qualitative Research* e incluye dos de las revistas temáticas de mayor impacto *Information, Communication & Society* y *Journal of Consumer Culture*. Además, contiene revistas que cubren temáticas como la sociología del lenguaje, salud, media y cultura, inmigración, estudios raciales, nacionalismo y estudios internacionales.

La comunidad situada entre las dos anteriores es la comunidad 3 y consta de 20 nodos. Contiene revistas teóricas como *Sociological Theory* y *Theory and Society*, además de revistas de corte cualitativo como *Qualitative Sociology*, *Ethnography* y *Journal of Contemporary Ethnography*. En la parte superior de la red está la comunidad 4, esta comunidad está compuesta por revistas temáticas que abordan materias como la sociología rural, la agricultura, los recursos naturales, el ocio o el deporte. La comunidad que aparece en la parte inferior de la red es la comunidad 5 e incluye gran cantidad de revistas de lengua no inglesa enlazadas con el componente principal a través de revistas de temática internacional como *European Sociological Review*, *International Journal of Comparative Sociology* y *European Societies*. Finalmente, la comunidad 6 agrega 18 revistas desconectadas del componente principal donde encontramos enlaces entre revistas de los mismos países como las dos revistas españolas *Revista Española de Investigaciones* y la *Revista Internacional de Sociología* o las croatas *Sociologija I Prostor* y *Drustvena Istrazivanja*. También hay enlaces entre revistas de temática común como *Society & Animals* y *Anthrozoos*.

Figura 13. Las seis comunidades representadas como subgrafos usando el grado para el tamaño de los nodos y el factor de impacto para el color



9. DISCUSIÓN DE RESULTADOS

Las revistas JCR de sociología y sus artículos conforman una rica fuente para conocer el conjunto de temáticas, sistema de comunicación compuesto por investigadores, instituciones y editoriales que complementa sostienen el conocimiento conjunto de la sociología como disciplina. Como expresó operativamente Inkeles (1964, 8): la sociología puede ser abordada como “lo que hacen los

sociólogos” y en la actualidad la mayor parte de los sociólogos académicos se dedican a publicar artículos en revistas de impacto.

Los análisis realizados muestran una sociología dominada por grupos editoriales internacionales y en menor medida por unas pocas editoriales universitarias. Como es de común conocimiento, el idioma predominante es el inglés y dos de cada tres revistas tienen su origen en Estados Unidos y Reino Unido. Las universidades de ambos países también aportan los autores con más artículos y citas.

Destacan en la disciplina las revistas de la llamada triple corona: *American Journal of Sociology*, *American Sociological Review* y *Social Forces*. En el listado de las incluidas en el listado JCR, son las primeras en aparecer y ejercen un dominio absoluto en citas e impacto. Por otro lado, estas revistas poseen un grado de coautoría bajo y la mayor desproporción de género entre sus autores. Tras ellas, en los años setenta aparecen gran cantidad de revistas temáticas y metodológicas. Las revistas metodológicas tienen la media más alta de coautores con más de dos autores por artículo aportando similares resultados a los obtenidos por Moody (2004). Por su parte, Grant y Ward (1991) también señalaron un mayor equilibrio de género entre los autores que publican en las revistas temáticas y, en especial, en aquellas con enfoques de género, familia o salud, que obtienen proporciones más altas de autoría de género femenino (Grant y Ward, 1991).

Las revistas se agrupan en comunidades a través de las redes de afiliación que forman con los autores. La red resultante corresponde claramente con el primer tipo de red de colaboración que propone Moody (2004). Se trata de una red compuesta por múltiples especialidades desconectadas y con subredes altamente agrupadas reflejando la falta de una teoría unificada de la disciplina. Las tres revistas americanas más importantes lideran la comunidad con mayor número de miembros como ya anticipaban (Moody y Light, 2006). Esta subred está también compuesta por un grupo de revistas metodológicas más recientes y otras sobre temas de religión, sociología militar, juventud y salud. La siguiente comunidad en tamaño la dominan las revistas del Reino Unido donde hay una sola revista metodológica de corte cualitativo y otras temáticas de inmigración, estudios raciales o nacionalismo. Esta división entre Estados Unidos y Reino Unido coincide con el análisis de Zougiris (2018). Además, nos encontramos con otros dos tipos de comunidades, las temáticas y las de lengua no inglesa. Las temáticas abordan temas como la sociología rural, la agricultura, los recursos naturales, las relaciones humanos-animales, el ocio y el deporte. Las de lengua no inglesa, responden a las comunidades nacionales de Vanderstraeten (2010) e incluyen revistas alemanas, francesas, españolas o del este europeo con una conexión fuerte con las revistas de enfoque europeo.

10. CONCLUSIONES

En este artículo se ha caracterizado el conjunto de la disciplina sociológica a partir de las publicaciones en revistas especializadas de impacto en los últimos años. La novedad del trabajo consiste en haber utilizado una fuente de datos de las denominadas *big scholarly data*. En concreto se ha utilizado el Microsoft Academic Graph que cuenta con más de 300 millones de publicaciones. El artículo aporta varias estrategias que permiten seleccionar los datos de interés entre los millones de publicaciones y reducir su dimensionalidad para representarlos en forma de red. La selección inicial se realiza a través de las revistas conforme al ranking de revistas JCR de sociología. Tras esto se vuelve a realizar una selección de revistas y se reúnen en tres grupos para los que se comparan sus citas, coautoría y sesgo de género. Finalmente, se utiliza el análisis de coincidencias para establecer una red de afiliación entre autores y revistas que se proyecta a una red de modo uno formada por solo revistas. Esta red permite localizar agrupaciones de revistas enlazadas por autores comunes.

Los resultados corroboran análisis de otros estudios presentando la sociología como una disciplina dominada por revistas generalistas anglosajonas. Es una disciplina que abarca una amplia variedad de temáticas y con enfoques metodológicos diversos que dependen del ámbito geográfico en el que se desarrollan. Unas pocas revistas dominan las citas, mientras que los artículos de revistas metodológicas tienen un grado mayor de coautoría y las revistas temáticas un menor sesgo de género.

Estos resultados pueden verse condicionados por el sesgo anglosajón en la cobertura de las revistas JCR y por la utilización únicamente de artículos de revistas y no de libros o comunicaciones en congresos.

Una posible investigación para estudios futuros consistiría en clasificar temática y metodológicamente los artículos de las revistas seleccionadas y poder evaluar cuestiones tales como la relación entre el número de citas y las distintas temáticas o metodologías, la evolución de las tendencias en los últimos años en relación con la distribución geográfica de las revistas o la correspondencia entre el género, las materias y los métodos que se utilizan.

11. BIBLIOGRAFÍA

- AMATURO, E. y PUNZIANO, G. (2017): “Blurry Boundaries: Internet, Big-New Data, and Mixed-Method Approach”, en *Data Science and Social Research*, Cham, Springer International Publishing, pp. 35-55
- BLOK, A., CARLSEN, H. B., JØRGENSEN, T. B., MADSEN, M. M., RALUND, S. y PEDERSEN, M. A. (2017): “Stitching Together the Heterogeneous Party: A Complementary Social Data Science Experiment”, *Big Data & Society*, 4(2), DOI: 10.1177/2053951717736337

- BLONDEL, V. D., GUILLAUME, J.-L., LAMBIOTTE, R. y LEFEBVRE, E. (2008): "Fast Unfolding of Communities in Large Networks", *Journal of Statistical Mechanics: Theory and Experiment*, 2008(10), DOI: 10.1088/1742-5468/2008/10/p10008
- BOURDIEU, P. (2004): *Science of Science and Reflexivity*, Cambridge, Polity Press
- BOYD, D. y CRAWFORD, K. (2012): "Critical Questions for Big Data", *Information, communication & society*, 15(5), pp. 662-79. DOI: 10.1080/1369118X.2012.678878
- BURROWS, R. y SAVAGE, M. (2014): "After the Crisis? Big Data and the Methodological Challenges of Empirical Sociology", *Big Data & Society*, 1(1), DOI: 10.1177/2053951714540280
- CARROLL, S. (2009): "Defining the Scientific Method", *Nature Methods*, 6(4), pp. 237-37. DOI: 10.1038/nmeth0409-237
- CASTELLS, M. (1998): *The Rise of the Network Society*, Oxford, Blackwell Publishers
- CLARIVATE ANALYTICS (2020): *Journal Citation Reports*, disponible en <https://jcr.clarivate.com/> [consulta: Agosto 2019]
- CLEMENS, E. S., POWELL, W. W., MCILWAINE, K. y OKAMOTO, D. (1995): "Careers in Print: Books, Journals, and Scholarly Reputations", *American Journal of Sociology*, 101(2), pp. 433-94. DOI: 10.1086/230730
- CLOGG, C. C. (1992): "The Impact of Sociological Methodology on Statistical Methodology", *Statistical Science*, 7(2), pp. 183-96.
- DE HAAN, J. (1997): "Authorship Patterns in Dutch Sociology", *Scientometrics*, 39(2), pp. 197-208. DOI: 10.1007/BF02457448
- DURKHEIM, E. (1982): "Sociology and the Social Sciences (1903)", en *The Rules of Sociological Method: And Selected Texts on Sociology and Its Method*, London, Macmillan Education UK, pp. 175-208
- ESCOBAR, M. (2015): "Studying Coincidences with Network Analysis and Other Multivariate Tools", *Stata Journal*, 15(4), pp. 1118-56.
- ESCOBAR, M. y TEJERO, C. (2018): "El Análisis Reticular De Coincidencias", *Empiria. Revista de metodología de ciencias sociales*, 39(2018) DOI:10.5944/empiria.39.2018.20879
- ESCOBAR, M. y MARTINEZ-URIBE, L. (2020): "Network Coincidence Analysis: The Netcoin R Package", *Journal of Statistical Software*, 93(11), pp. 1-32, DOI: 10.18637/jss.v093.i11
- ESPELAND, W. N. y STEVENS, M. L. (2008): "A Sociology of Quantification", *European Journal of Sociology*, 49(3), pp. 401-36. DOI: 10.1017/S0003975609000150
- EVANS, J. A. y ACEVES, P. (2016): "Machine Translation: Mining Text for Social Theory", *Annual Review of Sociology*, 42(1), pp. 21-50. DOI: 10.1146/annurev-soc-081715-074206
- FRANK, M. R., WANG, D., CEBRIAN, M. y RAHWAN, I. (2019): "The Evolution of Citation Graphs in Artificial Intelligence Research", *Nature Machine Intelligence*, 1(2), pp. 79-85. DOI: 10.1038/s42256-019-0024-5
- FRUCHTERMAN, T. M. J. y REINGOLD, E. M. (1991): "Graph Drawing by Force-Directed Placement", *Software: Practice and Experience*, 21(11), pp. 1129-64. DOI: 10.1002/spe.4380211102
- GANTMAN, E. R. y DABÓS, M. P. (2018): "Research Output and Impact of the Fields of Management, Economics, and Sociology in Spain and France: An Analysis Using Google Scholar and Scopus", *Journal of the Association for Information Science and Technology*, 69(8), pp. 1054-66. DOI: 10.1002/asi.24020

- GIARDULLO, P. (2016): "Does 'Bigger' Mean 'Better'? Pitfalls and Shortcuts Associated with Big Data for Social Research", *Quality & Quantity*, 50(2), pp. 529-47. DOI: 10.1007/s11135-015-0162-8
- GRANT, L. y WARD, K. B. (1991): "Gender and Publishing in Sociology", *Gender and Society*, 5(2), pp. 207-223
- GRIMMER, J. (2015): "We Are All Social Scientists Now: How Big Data, Machine Learning, and Causal Inference Work Together", *Political Science & Politics*, 48(1), pp. 80-83. DOI: 10.1017/S1049096514001784
- GUPTA, B. y BHATTACHARYA, S. (2004): "Bibliometric Approach Towards Mapping the Dynamics of Science and Technology", *DESIDOC Journal of Library & Information Technology*, 24(1)
- HABERMAN, S.J. (1973): "The Analysis of Residuals in Cross-Classified Tables", *Biometrics*, 29(1), pp. 1-25. DOI: 10.18637/jss.v014.i15
- HALFORD, S. y SAVAGE, M. (2017): "Speaking Sociologically with Big Data: Symphonic Social Science and the Future for Big Data Research", *Sociology*, 51(6), pp. 1132-48. DOI: 10.1177/0038038517698639
- HEALY, K. y MOODY, J. (2014): "Data Visualization in Sociology", *Annual Review of Sociology*, 40(1), pp. 105-28. DOI: 10.1146/annurev-soc-071312-145551
- JACOBS, J. A. (2016): "Journal Rankings in Sociology: Using the H Index with Google Scholar", *The American Sociologist*, 47(2), pp. 192-224. DOI: 10.1007/s12108-015-9292-7
- KHUN, T. S. (1962): *The Structure of Scientific Revolutions*, United States, The University of Chicago Press
- KOROM, P. (2019): "A Bibliometric Visualization of the Economics and Sociology of Wealth Inequality: A World Apart?", *Scientometrics*, 118(3), pp. 849-68. DOI: 10.1007/s11192-018-03000-z
- KOROM, P. (2020): "The Prestige Elite in Sociology: Toward a Collective Biography of the Most Cited Scholars (1970-2010)", *The Sociological Quarterly*, 61(1), pp. 128-63. DOI: 10.1080/00380253.2019.1581037
- LATOUR, B. y WOOLGAR, S. (1987): *Laboratory life*, New Jersey, Princeton University Press
- LAZER, D., PENTLAND, A., ADAMIC, L., ARAL, S., BARABÁSI, A.-L., BREWER, D., CHRISTAKIS, N., CONTRACTOR, N., FOWLER, J., GUTMANN, M., JEBARA, T., KING, G., MACY, M., ROY, D. y VAN ALSTYNE, M. (2009): "Computational Social Science", *Science*, 323(5915), pp. 721. DOI: 10.1126/science.1167742
- MANOVICH, L. (2015): "Data Science and Digital Art History", *International Journal for Digital Art History*, 0(1), DOI: 10.11588/dah.2015.1.21631
- MARTINHO, D. T. (2018): "Researching Culture through Big Data: Computational Engineering and the Human and Social Sciences", *Social Sciences*, 7(12), DOI: 10.3390/socsci7120264
- McFARLAND, D. A. y McFARLAND, H. R. (2015): "Big Data and the Danger of Being Precisely Inaccurate", *Big Data & Society*, 2(2), DOI: 10.1177/2053951715602495
- McFARLAND, D. A., LEWIS, K. y GOLDBERG, A. (2016): "Sociology in the Era of Big Data: The Ascent of Forensic Social Science", *The American Sociologist*, 47(1), pp. 12-35. DOI: 10.1007/s12108-015-9291-8
- MILLS, C. W. (1959): *The Sociological Imagination*, New York, Oxford University Press

- MOED, H. F. (2005): "Citation Analysis of Scientific Journals and Journal Impact Measures", *Current Science*, 89(12), pp. 1990-96.
- MOODY, J. (2004): "The Structure of Social Science Collaboration Network: Disciplinary Cohesion from 1963 to 1999", *American Sociological Review*, 69(2), pp. 213-238. DOI: 10.1177/000312240406900204
- MOODY, J. y LIGHT, R. (2006): "A view from Above: The Evolving Sociological Landscape", *The American Sociologist*, 37(2), pp. 67-86
- MOKSONY, F., HEGEDŰS, R. y CSÁSZÁR, M. (2014): "Rankings, Research Styles, and Publication Cultures: A Study of American Sociology Departments", *Scientometrics*, 101(3), pp. 1715-29. DOI: 10.1007/s11192-013-1218-y
- MORETTI, F. (2000): "Conjectures on World Literature", *New left review*, pp. 54-68.
- NERESINI, F. (2017): "On Data, Big Data and Social Research. Is It a Real Revolution?" en *On Data, Big Data and Social Research. Is It a Real Revolution?*, Cham, Springer International Publishing, pp. 9-16
- O'REILLY, K. (2009): "For Interdisciplinarity and a Disciplined, Professional Sociology", *Innovation: The European Journal of Social Science Research*, 22(2), pp. 219-32. DOI: 10.1080/13511610903075761
- OROMANER, M. (1981): "Cognitive Consensus in Recent Mainstream American Sociology: An Empirical Analysis", *Scientometrics*, 3(2), pp. 73-84. DOI: 10.1007/BF02025631
- PERITZ, B. C. (1983): "Are Methodological Papers More Cited Than Theoretical or Empirical Ones? The Case of Sociology", *Scientometrics*, 5(4), pp. 211-18. DOI: 10.1007/BF02019738
- PHELAN, T. J. (2000): "Bibliometrics and the Evaluation of Australian Sociology", *Journal of Sociology*, 36(3), pp. 345-63. DOI: 10.1177/144078330003600305
- PINCH, T. J. y BIJKER, W. E. (1984): "The Social Construction of Facts and Artefacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other", *Social Studies of Science*, 14(3), pp. 399-441. DOI: 10.1177/030631284014003004
- PRESS, G. (2014): "12 Big Data Definitions: What's Yours?" *Forbes*. Disponible en la página web: <https://www.forbes.com/sites/gilpress/2014/09/03/12-big-data-definitions-whats-yours/> [consulta: Septiembre 2021]
- RAFTERY, A. E. (2000): "Statistics in Sociology, 1950-2000", *Journal of the American Statistical Association*, 95(450), pp. 654-61. DOI: 10.1080/01621459.2000.10474245
- RIVIERA, E. (2015): "Testing the Strength of the Normative Approach in Citation Theory through Relational Bibliometrics: The Case of Italian Sociology", *Journal of the Association for Information Science and Technology*, 66(6), pp. 1178-88. DOI: 10.1002/asi.23248
- RODRIGUEZ-YUNTA, L. (2009): "Revistas españolas en WoS", *Anuario ThinkEPI*, 2010(4), pp. 250-253 web: <https://recyt.fecyt.es/index.php/ThinkEPI/article/view/31268> [consulta: Septiembre 2021]
- SANTAMARÍA, L. y MIHALJEVIĆ, H. (2018): "Comparison and Benchmark of Name-to-Gender Inference Services", *PeerJ Computer Science*, 4, pp. e156.
- SAVAGE, M. y BURROWS, R. (2007): "The Coming Crisis of Empirical Sociology", *Sociology*, 41(5), pp. 885-99. DOI: 10.1177/0038038507080443
- SCHWEMMER, C. y WIECZOREK, O. (2019): "The Methodological Divide of Sociology: Evidence from Two Decades of Journal Publications", *Sociology*, 54(1), pp. 3-21. DOI: 10.1177/0038038519853146

- SEALE, C. (2008): "Mapping the Field of Medical Sociology: A Comparative Analysis of Journals", *Sociology of health & illness*, 30(5), pp. 677-95. DOI: 10.1111/j.1467-9566.2008.01090.x
- SINHA, A., SHEN, Z., SONG, Y., MA, H., EIDE, D., HSU, B.-J. y WANG, K. (2015): "An Overview of Microsoft Academic Service (Mas) and Applications." en *An Overview of Microsoft Academic Service (Mas) and Applications*, pp. 243-46
- SMALL, A. W. (1906): "The Relation between Sociology and Other Sciences", *American Journal of Sociology*, 12(1), pp. 11-31.
- SMELSER, N. J. (2014): "The Optimum Scope of Sociology (1969)." en *Getting Sociology Right*, University of California Press, pp. 15-34
- STICHWEH, R. (2008): "The Sociology of Scientific Disciplines: On the Genesis and Stability of the Disciplinary Structure of Modern Science", *Science in Context*, 5(1), pp. 3-15. DOI: 10.1017/S0269889700001071
- STINCHCOMBE, A. L. (1984): "The Origins of Sociology as a Discipline", *Acta Sociologica*, 27(1), pp. 51-61. DOI: 10.1177/000169938402700104
- SU, H.-N. y LEE, P.-C. (2010): "Mapping Knowledge Structure by Keyword Co-Occurrence: A First Look at Journal Papers in Technology Foresight", *Scientometrics*, 85(1), pp. 65-79. DOI: 10.1007/s11192-010-0259-8
- TAYLOR & FRANCIS (2017): Co-Authorship in the Humanities and Social Sciences, disponible en página web <https://authorservices.taylorandfrancis.com/wp-content/uploads/2017/09/Coauthorship-white-paper.pdf> [consulta: Agosto 2019]
- TINATI, R., HALFORD, S., CARR, L. y POPE, C. (2014): "Big Data: Methodological Challenges and Approaches for Sociological Analysis", *Sociology*, 48(4), pp. 663-81. DOI: 10.1177/0038038513511561
- TUBARO, P. (2014): "Sociology and Social Networks." en *Sociology and Social Networks*, SAGE Publications Sage UK: London, England, pp.
- VANDERSTRAETEN, R. (2010): "Scientific Communication: Sociology Journals and Publication Practices", *Sociology*, 44(3), pp. 559-76. DOI: 10.1177/0038038510362477
- WANG, K., SHEN, Z., HUANG, C., WU, C.-H., EIDE, D., DONG, Y., QIAN, J., KANAKIA, A., CHEN, A. y ROGAHN, R. (2019): "A Review of Microsoft Academic Services for Science of Science Studies", *Frontiers in Big Data*, 2, pp. 45.
- XIA, F., WANG, W., BEKELE, T. M. y LIU, H. (2017): "Big Scholarly Data: A Survey", *IEEE Transactions on Big Data*, 3(1), pp. 18-35. DOI: 10.1109/TBDA-TA.2016.2641460
- ZOUGIRIS, K. (2018): "Detecting Topical Divides and Topical Bridges Across National Sociologies". *The American Sociologist*, 50, pp. 63-84. DOI: 10.1007/s12108-018-9392-2