

Búsqueda bibliográfica

Cómo repensar las formas de buscar, recopilar y analizar la producción científica escrita

Viviana Martinovich



SERIE SALUD COLECTIVA

El médico y la medicina: autonomía y vínculos de confianza en la práctica profesional del siglo XX
Lilía Blima Schraiber, 2019

Gobernantes y gestores: las capacidades de gobierno a través de narrativas, puntos de vista y representaciones
Hugo Spinelli, Jorge Arakaki, Leonardo Federico, 2019

Morir de alcohol: saber y hegemonía médica
Eduardo L. Menéndez, 2020

Violencia obstétrica en América Latina: conceptualización, experiencias, medición y estrategias
Patrizia Quattrocchi, Natalia Magnone (compiladoras), 2020

Pensar en salud
Matio Testa, 2020

Adiós, señor presidente
Carlos Matus, 2020

Método Paideia: análisis y cogestión de colectivos
Gastão Wagner de Sousa Campos, 2021

Gestión en salud: en defensa de la vida
Gastão Wagner de Sousa Campos, 2021

Desafíos para la salud colectiva en el siglo XXI
Jairmilson Silva Paim, 2021

Estado sin ciudadanos: seguridad social en América Latina
Sonia Fleury, 2021

Teoría del juego social
Carlos Matus, 2021

La salud persecutoria: los límites de la responsabilidad
Luis David Castiel, Carlos Álvarez-Dardet, 2021

Salud: cartografía del trabajo vivo
Emerson Elias Merhy, 2021

Sentirjugarhacerpensar: la acción en el campo de la salud
Hugo Spinelli, 2022

Saber en salud: La construcción del conocimiento
Mario Testa, 2022

El líder sin estado mayor: la oficina del gobernante
Carlos Matus, 2022

La historia de la salud y la enfermedad interpelada: Latinoamérica y España (siglos XIX-XXI)
Gustavo Vallejo, Marisa Miranda, Adriana Álvarez, Adrián Carbonetti, María Silvia Di Liscia, 2022

Precariedades del exceso: Información y comunicación en salud colectiva
Luis David Castiel, Paulo Roberto Vasconcellos-Silva, 2022

SERIE CLÁSICOS

Política sanitaria argentina
Ramón Carrillo, 2018

Medicina del trabajo al servicio de los trabajadores
Instituto de Medicina del Trabajo, 2019

Geopolítica del hambre: Ensayo sobre los problemas de la alimentación y la población del mundo
Josué de Castro, 2019

La salud mental en China
Gregorio Bermann, 2020

SERIE TRAYECTORIAS

Vida de sanitarista
Mario Hamilton, 2021

SERIE DIDÁCTICA

Teorías dominantes y alternativas en epidemiología
Marcelo Luis Urquía, 2019

Método Altadir de planificación popular
Carlos Matus, 2021

SERIE INFORMES TÉCNICOS

Salud en cárceles: Informe de auditoría de la situación sanitaria en el Servicio Penitenciario Bonaerense, 2013-2014
Instituto de Salud Colectiva, 2020

Búsqueda bibliográfica

Cómo repensar las formas de buscar, recopilar
y analizar la producción científica escrita

Viviana Martinovich



EDUNLA
COOPERATIVA

Secretaría de Investigación y Posgrado

Martinovich, Viviana

Búsqueda bibliográfica : cómo repensar las formas de buscar, recopilar y analizar la producción científica escrita / Viviana Martinovich. - 1a ed. - Remedios de Escalada : De la UNLa - Universidad Nacional de Lanús, 2022.

Libro digital, PDF - (Cuadernos del ISCo / Hugo Spinelli ; Didáctica ; 27)

Archivo Digital: descarga y online

ISBN 978-987-8926-16-2

1. Búsqueda Documental. 2. Búsqueda y Recuperación de Información. 3. Técnicas de Búsqueda de Información. I. Título.

CDD 025.5240218

Colección *Cuadernos del ISCo*

Serie *Didáctica*

Dirección científica: *Hugo Spinelli*

Dirección editorial: *Viviana Martinovich*

Edición ejecutiva: *Jorge Arakaki, Ignacio Yannone*

Coordinación editorial de esta obra: *Ignacio Yannone*

Fotografía de tapa e interiores: *Carlos Castilla*

© 2022, Viviana Martinovich

© 2022, EDUNLa Cooperativa

ISBN 978-987-8926-16-2

DOI 10.18294/9789878926162

EDUNLa Cooperativa

Edificio "José Hernández"

29 de Septiembre 3901, B1826GLC Remedios de Escalada, Buenos Aires, Argentina

Teléfono: (54-11) 5533-5600 int. 5727. edunla@unla.edu.ar

Instituto de Salud Colectiva

Edificio "Leonardo Wertheim"

29 de Septiembre 3901, B1826GLC Remedios de Escalada, Buenos Aires, Argentina

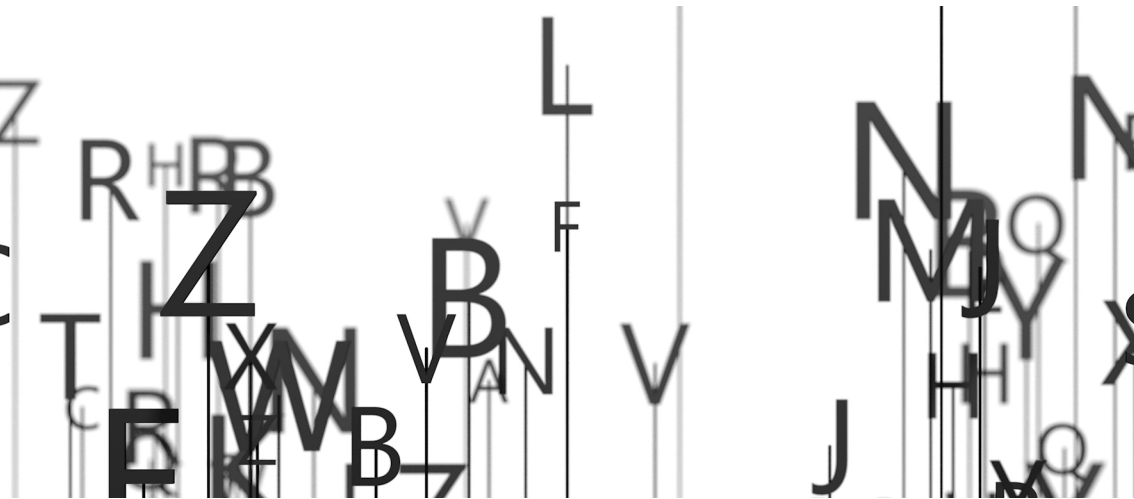
Teléfono: (54-11) 5533-5600 int. 5958. <http://cuadernosdelisco.unla.edu.ar>



Esta obra está bajo licencia internacional Creative Commons

Reconocimiento-NoComercial-CompartirIgual 4.0.

Las y los autores conservan sus derechos autorales y les permiten a otras personas copiar y distribuir su obra siempre y cuando reconozcan la correspondiente autoría y no se utilice la obra con fines comerciales.



Mientras un texto sea mudo no ha iniciado aún su comprensión.

Pero un texto puede empezar a hablar [...] Entonces no se limita a decir su palabra, siempre la misma, en una rigidez inerte, sino que da nuevas respuestas al que le pregunta y formula nuevas preguntas al que le contesta. Comprender textos es entenderse en una especie de conversación.

Hans-Georg Gadamer (2006, p. 130)

Viviana Martinovich

Se graduó como Editora en la Facultad de Filosofía y Letras, Universidad de Buenos Aires. Es profesora adjunta en los posgrados del ISCo-UNLa. Desde 2005, es editora ejecutiva de la revista científica *Salud Colectiva*, y directora editorial de la colección de libros Cuadernos del ISCo, ambos editados por el ISCo-UNLa.

Desde la historia de las revistas científicas, su área de estudio, se centra en los complejos procesos de mediación y la trama social de intereses que posibilitan que un texto científico acceda a la esfera pública y se inscriba en una conversación. Desde esta perspectiva publicó diversos editoriales en la revista científica *Salud Colectiva*, entre ellos “Diez años de Salud Colectiva: una aproximación a las reglas de juego del campo editorial científico” (2014); “La dimensión política del acceso abierto: ¿el conocimiento como bien público o como mercancía?” (2015); “Salud Colectiva: una apuesta al acceso abierto” (2016); como así también una serie de entrevistas a editoras y editores, entre ellas “Práctica editorial contextualizada: Carlos Augusto Monteiro y la *Revista de Saúde Pública*” (2016) y “Práctica editorial contextualizada: Maria Cecília de Souza Minayo y la revista *Ciência & Saúde Coletiva*” (2016).

Entre sus artículos de investigación, se destacan “Indicadores de citación y relevancia científica: Genealogía de una representación” (*Dados*, 2020); y “Ciencia abierta en América Latina: repensar la interdependencia dinámica entre las ciencias y la sociedad” (*Espacios*, 2021).

Recuperando la dimensión política de la edición científica, publicó algunas notas de divulgación como “Los dueños de la ciencia” (*Anfibia*, 2017); “El negocio de las revistas científicas” (*Página 12*, 2018); “COVID-19 y dependencia científica” (*El Cohete a la Luna*, 2020); “Barajar y dar de nuevo: ¿No será hora de apostar a crear una industria editorial científica nacional?” (*El Cohete a la Luna*, 2020).

Como docente invitada, ha dictado cursos en posgrados de universidades de Argentina, Chile y Colombia; y desde su labor editorial, ha sido invitada como expositora a diversos encuentros en Argentina, Brasil, Perú, Cuba y Colombia.



Índice

Introducción	1
Capítulo 1 Fase exploratoria de la investigación	7
Capítulo 2 Investigaciones basadas en fuentes escritas	19
Capítulo 3 ¿Qué es una revisión bibliográfica?	29
Capítulo 4 Elección de las bases de datos de bibliografía científica	35
Capítulo 5 ¿Cómo buscar en una base de datos?	49
Capítulo 6 ¿Cómo crear un mapa de términos de búsqueda?	63
Capítulo 7 Diseño de un instrumento de búsqueda bibliográfica	69
A modo de cierre	73
Bibliografía	75

C
L
F
N
V
J
G
A
W
S
K
Z
G
O
N

Introducción

Este libro es resultado de un proceso que comenzó a gestarse en el año 2011, con las primeras guías de ejercicios de búsqueda bibliográfica desarrolladas desde el Instituto de Salud Colectiva de la Universidad Nacional de Lanús, en las que se presentaban algunas técnicas y algunos conceptos que año tras año se fueron ampliando y enriqueciendo sobre la base de las nuevas preguntas que surgían en las clases. Las discusiones con estudiantes y docentes pusieron de manifiesto, por un lado, que las formas de buscar y recuperar bibliografía están estrechamente vinculadas a las posibilidades posteriores de analizar ese material y, por otro, que repensar qué es un texto y cuántos planos y dimensiones de análisis propone, nos abre un universo muy vasto de formas de concebir la búsqueda y recuperación de las fuentes escritas. Y, en este sentido, este texto es una invitación a transitar ese proceso.

Desde diversas áreas de investigación, la palabra “bibliografía” suele ser utilizada para designar el listado de referencias consultadas o el listado de textos propuestos para un curso, pero lo cierto es que la bibliografía es un campo de estudios en sí mismo, una disciplina con tradiciones, rupturas y discusiones internas, de la cual los distintos tipos de trabajos de “revisión bibliográfica” tomaron históricamente prestado muchas de sus características esenciales. Sin embargo, mientras que, en la década de 1980, la bibliografía como disciplina rompió con sus propios moldes descriptivos y enumerativos, los distintos tipos de estudios que toman la bibliografía como insumo siguen atrapados en su impronta de contabilizar, enumerar y describir existencias.

Entonces ¿por qué no intentar recuperar las nuevas perspectivas de la bibliografía, de la historia y de la sociología de los textos para enriquecer las formas en que las diversas áreas del conocimiento analizan las producciones textuales?; ¿por qué no abordar la propia narrativa científica con la riqueza con que ciertas áreas de las ciencias analizan otras producciones discursivas? ¿Es posible hacer este proceso sin repensar las formas de recuperar e interpretar esa bibliografía?

Para iniciar el recorrido, nuestra propia etapa exploratoria nos permitió recuperar y analizar un cuerpo amplio de trabajos de revisión que aplicaban distintas técnicas desde perspectivas disciplinarias y teóricas diversas. En gran parte de esos trabajos, si bien en las páginas introductorias se situaba y justificaba socialmente el problema de investigación, identificando incluso a distintos actores sociales involucrados en la construcción de ese problema, a

la hora de conformar el corpus, optar por una técnica de revisión, seleccionar los términos de búsqueda, las bases de datos, e incluso analizar los objetos textuales, la técnica se tornaba irreflexiva, descontextualizada, y el problema social pasaba a ser analizado como un dato, cuya naturaleza se tornaba irrelevante. La aplicación de una técnica de revisión en modo “autómata” parecía generar la ilusión de poder sistematizar y sintetizar la complejidad de un problema a través de un recorte de artículos científicos y, al no problematizar las formas de crear el *corpus* ni los fundamentos teóricos desde los cuales se analizan los textos recuperados, se desdibujaban las limitaciones y sesgos, generando un espejismo respecto del grado de representatividad conceptual, geográfica y temporal, y de las conclusiones a las que se arribaba. Pero enriquecer las formas de interpretar los textos ¿no demanda repensar las formas en que buscamos e interpretamos ese recorte de fuentes escritas que componemos?

El efecto Selecciones del Reader's Digest

En aquellos trabajos de revisión bibliográfica, cuyas preguntas centrales suelen plantear “qué se sabe respecto de tal tema”, es posible hacer una analogía con el efecto que producían los resúmenes de novelas publicados por *The Reader's Digest*, revista estadounidense que desde la década de 1940 se convirtió en uno de los fenómenos editoriales más importantes del mundo occidental y que el público hispanohablante conoció como *Selecciones del Reader's Digest*. Esta revista publicaba una selección de resúmenes de grandes obras literarias, que tenían un doble efecto: producían la ilusión no solo de haber leído las obras originales, sino de haber aprehendido la producción literaria internacional de la segunda mitad del siglo XX. Pero la realidad es que quienes leían los resúmenes publicados por *Selecciones del Reader's Digest* desconocían tanto los textos originales como la riqueza existente más allá de las fronteras del canon de lecturas propuesto por quienes dirigían la revista.

También nos encontramos con otros trabajos de revisión bibliográfica en los que la técnica de búsqueda y recuperación de artículos científicos primaba por sobre la formulación de un problema de investigación, poniendo el énfasis en la conformación y descripción del corpus, sin plantear un problema de investigación que superara el plano estrictamente informativo-descriptivo.

Más allá de dónde se ubicara el foco, gran parte de las propuestas analizadas seguían manteniendo la impronta de “recopilación” presente en los trabajos de revisión desde la década de 1940, o planteaban una “síntesis” de esa recopilación sin proponer un referencial teórico desde el cual interpelar las diversas perspectivas de la comunidad científico-académica sobre

un determinado tema y sin situar y problematizar el contexto político, económico, ideológico y cultural de esa producción escrita.

Como las técnicas de revisión tienen su origen en la era del papel, momento en que la accesibilidad a la bibliografía científica era mucho más limitada que en la actualidad, hasta fines del siglo XX, realizar estudios que simplemente recopilaran y describieran un número de trabajos recuperados de un catálogo compuesto por varios volúmenes impresos, era un aporte relevante al conocimiento sobre el tema. Sin embargo, el acceso remoto a grandes volúmenes de registros electrónicos sobre la producción científica de diversas regiones del mundo, es un proceso relativamente reciente que torna un tanto irrelevante la tarea de recopilación y síntesis y obliga a ir un poco más allá, generando nuevas técnicas y formas de entender las investigaciones basadas en fuentes escritas.

Hulme: “De la ausencia de un tipo de bibliografía no se puede sacar una conclusión válida”

Para poder ampliar y enriquecer las perspectivas teóricas desde las cuales abordar de forma crítica, reflexiva y contextualizada la producción científica, deberíamos reflexionar sobre las formas de buscar y recuperar el material para nuestro análisis. Si bien este libro no se centra en dimensiones analíticas sino en las distintas etapas de búsqueda y recuperación de fuentes escritas, entendemos que reflexionar sobre las formas de recuperación de las fuentes puede abrir las puertas para enriquecer el tipo de análisis que se realizará de esos textos. Es decir, al igual que en otros tipos de investigaciones, la etapa exploratoria ya debería estar orientada por un problema o una pregunta de investigación que se sustente en una dimensión analítica determinada, de manera de utilizar la etapa exploratoria para identificar posibles planos de análisis y a partir de allí formular un problema o una pregunta que delimite la técnica de búsqueda y recuperación de artículos.

En este sentido, proponemos ir más allá de aquella bibliografía que cataloga y estudia los textos como formas registradas, desde un plano meramente descriptivo. Desde que un tema se torna un problema de investigación, se publica e ingresa a la esfera pública, hasta que llega a manos de alguien que lo interpreta, se produce una gran cantidad de procesos de mediación que pueden potenciar un diálogo o limitarlo, ampliar su difusión o reducirla. Por lo tanto, no es posible disociar los textos que llegan hasta nuestras manos de la matriz de pensamiento y de las acciones del entorno que los puso en

circulación porque son dimensiones que se retroalimentan. En esos procesos de mediación convergen aspectos macroestructurales vinculados a intereses políticos, económicos y simbólicos –como, por ejemplo, el uso de las revistas científicas como estrategias de legitimación de los productos comercializados por los complejos industriales (Angell, 2006; Kearns *et al.*, 2015; Parker *et al.*, 2018)–, pero también una serie de prácticas mediadoras que operan a nivel microsociedad, es decir, saberes editoriales, prácticas de investigación, capitales sociales, culturas y políticas científicas, que son los que van a ampliar o limitar la circulación de determinadas líneas de investigación. Los textos se abren así a una variedad de dimensiones analíticas que deberíamos empezar a explorar.

Al comenzar a transitar este camino, nos encontramos con ciertos presupuestos instalados, prácticas recurrentes y limitaciones propias de los trabajos basados en fuentes escritas que creíamos necesario desandar. Como generalmente se suelen ignorar y, por lo tanto, anular analíticamente los procesos de mediación por los que atraviesa toda investigación —desde que se formula el problema hasta que el texto ya evaluado, reelaborado, editado y publicado llega a la esfera pública— se suele pensar, por un lado, que aquellos textos a los que se logra acceder es el universo de lo publicado sobre el tema (primera falacia) y que lo publicado es el universo de lo investigado (segunda falacia). Ya en las primeras décadas del siglo XX, más precisamente en 1923, en un estudio titulado “*The growth of modern civilization*” (El crecimiento de la civilización moderna), en uno de los primeros estudios de *bibliografía estadística*, Edward Wyndham Hulme reflexiona sobre las formas de acceso al contenido y al tipo de conclusiones a las que se puede llegar:

Cuando decimos que la bibliografía constituye el mejor reflejo de la mente humana, debe admitirse que la metáfora requiere alguna aclaración. En el espectro bibliográfico es cierto que encontramos líneas aquí y allá de cuya presencia pueden extraerse conclusiones confiables. Pero lo contrario no es cierto. De la ausencia de un tipo de bibliografía no se puede sacar una conclusión válida en cuanto al estado del conocimiento con respecto a esa clase. La práctica de un arte puede haber precedido a su exposición sistemática en la bibliografía, a menudo por muchos siglos. Se sabe que el proceso de expansión del círculo de la bibliografía es lento y gradual. Una vez más, una clasificación de temas no es una clasificación del conocimiento:

sus unidades no siempre coinciden con las divisiones temáticas¹.
(Hulme, 1923, p. 9)

¿Cuántas de las investigaciones recientes tienen en cuenta que los artículos recuperados a través de una base de datos no se relacionan necesariamente con el estado del conocimiento ni con la variedad de investigaciones realizadas sobre el tema que se propone abordar? ¿Cuántas de esas investigaciones reconocen la decisión de ciertas bases de datos de negar el ingreso a un gran número de revistas editadas en Europa del Este, Asia, América Latina o África? (Martinovich, 2020). Es necesario recuperar el espíritu presente en las palabras de Hulme para reconocer las limitaciones propias de la búsqueda y recuperación de bibliografía científica.

Una de esas limitaciones es que, al analizar la producción científica a través de los registros disponibles en los canales de distribución de esa producción —es decir, las bases de datos de bibliografía científica—, se recopilan registros al final de una cadena productiva muy extensa, intermediada por múltiples prácticas e intereses. Para ejemplificar esta idea podríamos decir que sería como analizar cuándo, dónde y quiénes producen tomates, a través de los registros de existencias de una cadena de hipermercados. Indudablemente que toda la producción de tomate no se distribuye envasada en latas a través de cadenas de supermercados y no todas las marcas de tomates envasados están disponibles en todas las cadenas. Si en vez de latas de tomates, quisiéramos analizar cómo se produce el tomate envasado de producción orgánica sin agroquímicos, cuya elaboración es menos industrializada y quizá existan marcas que ni siquiera ingresen a una cadena de hipermercados, el resultado que se obtendría de los registros de ese producto sería inexistente. ¿Y eso implica que ese tipo de producción es inexistente? Por absurdo que parezca, esta es la lógica que suele aplicarse en los trabajos de revisión bibliográfica. Esto nos muestra que los resultados encontrados en una base de datos estarían altamente condicionados por las formas de producción y distribución específicas de cada entorno social y por las reglas de juego que impongan las instancias de mediación. Tal como señala Juan Samaja:

¹Traducción del original: *When we speak of literature constituting the best reflection of the human mind it must be admitted that the metaphor requires some qualification. In the literary spectrum it is true we find here and there lines from the presence of which trustworthy conclusions may be drawn. But the converse does not hold true. From the absence of a literary class no valid conclusion can be drawn as to the state of knowledge in respect of that class. The practice of an art may have preceded its systematic exposition in literature—often by many centuries. The process of the extension of the circle of literature is known to have been a slow and gradual one. [...] Again, a book classification is not a classification of knowledge: its units do not always coincide with those of subject-matter.* (Hulme, 1923, p. 9)

El tema de las fuentes de datos suele ser concebido muy pobremente. Con el término “fuente de datos” se asocia la idea de un “cierto lugar” en donde “ocasionalmente” están los datos. Yo creo que esa representación encubre un concepto mucho más profundo y complejo: el concepto de *la praxis que la sociedad ha logrado generar en el campo del objeto de estudio, y por medio del cual él llega a ser objeto de conocimiento*. [...] En efecto, la palabra “fuente de datos” resume en una representación cosificada lo que una cierta sociedad obra en una cierta realidad, transformándola así en una realidad “intersubjetiva” y, consecuentemente, en “objeto” de conocimiento. (Samaja, 2007, p. 256) (cursivas del original)

En este sentido, el mundo virtual en el que se inscriben las bases de datos suele crear cierta ilusión de universo paralelo desvinculado de la participación de actores sociales y exento de intereses económicos y políticos, como una especie de oráculo cibernético que brinda respuestas inequívocas. Sin embargo, el análisis y discusión de los datos obtenidos a través de búsquedas bibliográficas no debería omitir, por ejemplo, que los modos de producción de la ciencia, las políticas científicas, el mercado editorial, o incluso la ausencia de un mercado, condicionan los resultados obtenidos a través de los canales de distribución.

Desde esta perspectiva, analizar qué expresan los registros obtenidos en una búsqueda bibliográfica requiere de un esfuerzo adicional. Las técnicas de búsqueda y recopilación de bibliografía científica reconstruyen un escenario posible en una temática en particular y, al hacerlo a través de los canales de distribución, el resultado queda necesariamente atravesado y condicionado no solo por las asimetrías en cuanto a los recursos económicos públicos y/o privados destinados a la investigación de cada temática o a la disparidad entre países respecto de la inversión en investigación y desarrollo, sino también por las propias prácticas de publicación, por el interés o el desinterés que le otorguen quienes investigan y quienes editan revistas a la distribución de sus trabajos a través de bases de datos. El corpus es una construcción, un recorte intermediado, cuyo análisis debería contemplar los intereses de quienes participan de los procesos de mediación para abordar el recorte de la realidad que intenta representar.

Este texto se pensó como material de apoyo para estudiantes, pero también como guía para docentes, tutores y directores en la tarea de acompañar la búsqueda y recuperación de fuentes escritas para ser utilizadas en todo el proceso de investigación.

Capítulo 1

Fase exploratoria de la investigación

En todo trabajo de investigación, uno de los primeros pasos suele ser atravesar una fase de búsqueda exploratoria de fuentes escritas, que permita delimitar un problema factible de ser abordado, reconstruir cuál es el cuerpo de conocimientos preexistente sobre ese problema, cuáles son sus referentes, con los que podremos concordar o no, pero con certeza no podremos desconocer ni omitir. ¿Cómo acceder a esta gran variedad de fuentes? La forma de acceso varía según el tipo de fuente, y en algunos casos tendrá caminos más sistematizados y, en otros, mucho más intuitivos e, incluso, azarosos.

Esta fase exploratoria es de vital importancia dado que, a la vez que amplía nuestro universo referencial sobre el tema, permite redefinir y recortar tanto el problema, la pregunta, como el objeto de estudio a partir de cotejar qué es lo que otras personas se han preguntado sobre el tema, desde qué marcos referenciales lo han abordado, con qué recursos metodológicos han tratado de responder a la pregunta planteada, qué dimensiones se suelen abordar con mayor frecuencia y cuáles no, entre muchas otras cuestiones.

Nos hemos encontrado algunos trabajos en los que la elección de la metodología y la técnica de análisis parecía preceder al planteamiento de un problema de investigación. Suelen ser investigaciones en las que prima el nivel descriptivo, y que obligan a una adecuación del problema para que pueda ser abordado por medio de una técnica específica. Si entendemos el *diseño de la investigación* como “el momento de la adopción de una estrategia metodológica para la resolución del problema” (Samaja, 2007, p. 203), la fase exploratoria debería centrarse en la identificación de problemas cuyo análisis se enriquezca con la perspectiva metodológica adoptada y no a la inversa. Tal como lo expresa María Cecilia de Souza Minayo:

...la investigación es un trabajo artesanal que no prescinde de la creatividad, se realiza fundamentalmente con un lenguaje basado en conceptos, proposiciones, hipótesis, métodos y técnicas; lenguaje que se construye con un ritmo propio y particular. Este ritmo lo

denominamos *ciclo de la investigación*, o sea, un peculiar proceso de trabajo en espiral que comienza con una pregunta y termina con una respuesta o producto que, a su vez, da origen a nuevos interrogantes. (Minayo, 2012, p. 30) (cursivas del original)

En los diferentes capítulos, incorporamos ejemplos de artículos en los que la identificación del problema estuviera estrechamente vinculada a la elección metodológica. Y cuando esto sucede, la apuesta creativa planteada en preguntas poco convencionales se suele trasladar a la reformulación de ciertas técnicas.

Por ejemplo, el artículo “*Criticism of the NOVA classification: who are the protagonists?*” (Mialon *et al.*, 2018), se origina en el marco del debate y la disputa de intereses en torno a la clasificación de alimentos NOVA propuesta por Carlos Monteiro desde la Universidade de São Paulo, que ganó un gran respaldo internacional al exponer los efectos nocivos de los productos ultraprocesados, pero también recibió muchas críticas. En ese contexto, Melissa Mialon, Paulo Sêrodiob y Fernanda Scagliusia se plantearon como interrogante si existía algún vínculo entre la industria alimentaria y las personas y las organizaciones que criticaban la clasificación NOVA, que pudiera ser entendido como conflicto de intereses. Luego de hacer una primera búsqueda exploratoria en PubMed, Scopus y Web of Science, en la que recuperaron todos aquellos documentos que hicieran mención al término “ultraprocesado”, refinaron las búsquedas y seleccionaron todos los trabajos que cuestionaban la clasificación NOVA. Cruzaron los nombres de autoras y autores de las publicaciones seleccionadas con declaraciones de conflicto de intereses, agradecimientos, financiamientos, patrocinios y colaboraciones vinculadas con la cadena productiva de los alimentos ultraprocesados, incluyendo a las industrias, las grandes empresas de distribución y asociaciones empresariales y de promoción comercial. La técnica, en este caso, se puso al servicio de intentar abordar un problema complejo como es la manipulación de ciertas “evidencias”, adecuándolas a un tipo de retórica científica que encubre otros intereses. Este es uno de los tantos usos de las revistas científicas, como dispositivos de legitimación de productos o intereses de las industrias.

En este ejemplo, pudieron realizar el cruce de información y establecer redes de colaboración, porque rastrearon una gran cantidad de fuentes escritas e incluso audiovisuales muy diversas, ¿pero estas fuentes son la totalidad de las consultadas en la investigación? Por supuesto que no, porque para justificar el problema de investigación y discutir sus resultados necesitaron de otras fuentes escritas que fueron recuperadas por otras vías menos sistematizadas.

De esto es lo que trata este capítulo: de la gran variedad de fuentes escritas y de los diversos usos que tienen dentro de la estructura narrativa de una investigación.

Tipos de fuentes escritas

En el proceso de exploración inicial es importante distinguir los distintos tipos de fuentes con que trabajaremos en una investigación y los distintos usos que les daremos. Según Carmen Medina-López *et al.*:

Es el investigador el que debe decidir la necesidad de acudir a las diferentes opciones posibles [...] sin duda, la variedad de fuentes consultadas enriquecerá el trabajo realizado. Serán la naturaleza y los objetivos de la investigación desarrollada los que definen la combinación de fuentes que es necesario utilizar. (2010, p. 16)

Fuentes primarias y secundarias

Para algunas áreas disciplinarias, las fuentes primarias son las que aún no existen y que se construyen en la propia investigación, por ejemplo, las transcripciones de entrevistas, de grupos focales o de cuestionarios; la creación de datos estadísticos a partir de encuestas poblacionales, etc.; y las fuentes secundarias son documentos preexistentes, elaborados por otras personas, es decir, fuentes publicadas, que incluyen todas aquellas obras y documentos que abordan nuestro tema de estudio.

Sin embargo, tal como expresa Umberto Eco, para otros marcos disciplinares, un texto publicado también puede ser considerado una fuente primaria, según el uso dado dentro de la investigación:

Una tesis estudia un *objeto* valiéndose de determinados *instrumentos*. Muchas veces el objeto es un libro y los instrumentos son otros libros. Tal es el caso de una tesis, supongamos, sobre *El pensamiento económico de Adam Smith*, en la cual el objeto está constituido por los libros de Adam Smith mientras que los instrumentos son otros libros de Adam Smith. En tal caso diremos que los escritos de Adam Smith constituyen las *fuentes primarias* y los libros sobre Adam Smith

constituyen las *fuentes secundarias* o la *literatura crítica*. (Eco, 1992, p. 69) (cursivas del original)

Desde esta perspectiva, las fuentes secundarias serían todos aquellos textos que abordan nuestro tema, pero no forman parte de nuestro corpus de investigación. Es decir, un mismo texto puede tornarse una fuente primaria o secundaria según el uso dado en la investigación. Retomando el ejemplo de Umberto Eco, si en vez de trabajar sobre *el pensamiento económico de Adam Smith*, nuestro problema de investigación se centrara en *la presencia del pensamiento económico liberal en la prensa escrita: la configuración del rol del Estado frente a las desigualdades*, nuestras fuentes primarias serían artículos periodísticos de un recorte geográfico y temporal definido, y podríamos incorporar ciertos conceptos de los textos escritos por Adam Smith para que nutran nuestro marco teórico. En este caso, esos mismos libros que en el ejemplo de Umberto Eco eran fuentes primarias, en nuestra investigación pasarían a ser fuentes secundarias.

Esta distinción es de suma relevancia para comprender las diversas formas de interpretar un artículo, un libro o cualquier otro tipo de texto publicado. Cuando la lectura articula el discurso del texto con nuestros propios interrogantes, el texto cobra una nueva vida que se reactualiza en cada nueva lectura y en cada nuevo presente, abriendo la posibilidad de restituir, entre muchas otras dimensiones, las múltiples relaciones con el entorno social que permitió su existencia pública. En términos de Ricoeur:

Como lectores, podemos permanecer en la suspensión del texto, tratarlo como texto sin mundo y sin autor [...]. O bien podemos levantar la suspensión del texto [...] y restituirlo a la comunicación viva, con lo cual lo interpretamos. (Ricoeur, 1986, p. 135)

Desde esta perspectiva, nuestras propias nociones respecto de qué es un texto van a dar lugar a un tipo de lectura y a un tipo de apropiación. Según la perspectiva gadameriana, un texto es mudo mientras no se haya iniciado aún su comprensión, y comprender un texto implica mucho más que su apropiación instrumental. Según Gadamer (2006), la persona que lee tiene ante sí la posibilidad de que un texto comience a hablar, y cuando habilita esa posibilidad el texto “no se limita a decir su palabra, siempre la misma, en una rigidez inerte, sino que da nuevas respuestas al que le pregunta y formula nuevas preguntas al que le contesta” (Gadamer, 2006, p. 130). Esta forma dialógica de comprender los textos, aplicada a las producciones discursivas propias de una comunidad como la científica, tiene la potencialidad de transformar la

lectura de textos mudos en una lectura interpretativa de textos vivos. Según Gadamer, “comprender textos es entenderse en una especie de conversación” (2006, p. 130).

Fuentes de primera mano y segunda mano

Otra distinción que realiza Umberto Eco (1992) sobre las fuentes, se refiere al tipo de acceso. Las *fuentes de primera mano* son aquellos textos a los que accedemos en su versión original, no a través de la cita realizada por otros. Si leemos en un artículo que “según, el Banco Mundial, la tasa de incidencia de la pobreza a nivel mundial descendió del 38,2% en 1985 al 10,1% en 2015”, y ese dato nos interesa, deberíamos recurrir al sitio web del Banco Mundial para buscar, cotejar y citar la información.

A ese artículo que retoma los datos del Banco Mundial, Umberto Eco lo denomina *fuentes de segunda mano*, es decir, cuando tomamos un concepto, un dato o la descripción de un evento a través de una cita encontrada en otros textos. En estos casos deberíamos intentar recuperar el texto original, de lo contrario se corren serios riesgos de tergiversaciones. En teoría, la cita de una cita solo debería utilizarse en el caso de que sea confirmada en varias publicaciones, pero aún así se corren serios riesgos de reproducir errores.

Un ejemplo interesante de este tipo de tergiversaciones es la cita que figura en la versión en español del libro *Cómo escribir y publicar trabajos científicos* de Robert Day (2005). En el texto, al hablar de los estilos de referencias bibliográficas, Day menciona:

Las revistas varían mucho en la forma de tratar las referencias. Una autora que examinó 52 revistas científicas encontró 33 estilos diferentes de enumerar las referencias (M. O'Connor, Br. Med. J. 1(6104): 31, 1978). (Day, 2005, p. 49–50)

Según Robert Day, una “autora” había examinado 52 revistas y encontrado los 33 estilos diferentes de referencias. Si buscamos esa expresión en Internet, encontraremos una gran cantidad de textos que transcribieron literalmente la frase mencionada por Day. Incluso, en algunos casos, la recrearon del siguiente modo:

Las revistas científicas varían mucho en la forma de tratar las referencias. La autora M. O'Connor examinó 52 revistas científicas y

encontró 33 estilos diferentes de enumerar las referencias. (Castillo Blasco, 2002, p. 1)

En este caso, “la autora” —que mencionaba Robert Day— pasó a ser “M. O’Connor” y según Lourdes Castillo Blasco, había sido la propia O’Connor quien examinó las 52 revistas. A estas alturas, la cita había comenzado a despertar nuestro interés, por lo que procuramos el texto original del que Robert Day extrajo la mención, y para nuestra sorpresa encontramos lo siguiente:

At the BMJ-ELSE (European Life Science Editors) meeting in Winchester one speaker said that he had found 33 different styles of references in 52 journals. [En la reunión del BMJ-ELSE (European Life Science Editors) en Winchester, un orador dijo que había encontrado 33 estilos diferentes de referencias en 52 revistas]. (O’Connor, 1978, p. 31)

No era la “autora”, mencionada por Robert Day, ni mucho menos Maeve O’Connor quien había examinado las revistas. La mención había sido realizada por un orador y escuchada por Maeve O’Connor. Si bien nuestro ejemplo es bastante inofensivo, esta práctica puede producir tergiversaciones mucho más complejas, por lo que debería evitarse. Según Eco, “lo único que no puedo hacer es citar [...] a través de las citas hechas por otro. En teoría, un trabajo científico serio no tendría que citar *nunca* a partir de otras citas” (1992, p. 76) (cursivas del original).

Distintos usos de las fuentes dentro de una investigación

Como hemos visto, la selección del tipo de fuente varía según sea el uso que le demos dentro de la investigación. Por ejemplo, para la justificación del problema, muchas veces se utilizan datos de encuestas nacionales, censos, informes de organismos internacionales mientras que, para delimitar el marco teórico-conceptual, en general, se utilizan libros de autores de referencia.

Si bien cada tipo de investigación sugiere un abanico de fuentes posibles, cada una de las partes que integran el proyecto de investigación o el propio problema de investigación o su forma de abordarlo, delimitan ciertos usos y ciertas formas de recuperación, como veremos a continuación.

Justificación del problema

La “justificación del problema” suele ser parte del apartado Introducción. El uso dado a las fuentes escritas es justificar por qué el problema de investigación es relevante, para quién es relevante, qué actores sociales están involucrados, etc., por lo tanto, las fuentes suelen aportar datos o descripciones puntuales que validen esa relevancia social y académica, según a qué actores se invoque como interlocutores claves del problema (organizaciones de la sociedad civil, organismos públicos, asociaciones profesionales, agrupaciones políticas, movimientos sociales, medios de prensa, organizaciones no gubernamentales, etc.).

En el caso de ser una investigación teórica, las fuentes utilizadas para la justificación del problema central a tratar suelen vincularse a la necesidad de reflexionar, cuestionar o repensar un concepto, una teoría, o prácticas específicas. Por ejemplo, en el artículo “Desigualdades en salud: nuevas perspectivas teóricas”, Naomar de Almeida Filho (2020) se propone articular “la problemática teórica de las desigualdades económicas y sociales con la determinación social de la salud” (p. 1), partiendo de la hipótesis de que “está en curso un proceso de transformación de las desigualdades económicas con nuevas formas de inequidades sociales e iniquidades en salud” (p. 1). En la justificación de este problema teórico, retoma los aportes de aquellos referentes del pensamiento epidemiológico que décadas atrás discutieron la necesidad de entender la epidemiología social, tanto como una cuestión teórica, como problemática metodológica y como desafío político, es decir, justifica el problema exponiendo y apoyándose en las discusiones teóricas precedentes.

En el caso del artículo “Guías bilingües: una estrategia para disminuir las barreras culturales en el acceso y la atención en salud de las comunidades wayuu de Maicao, Colombia” (Patiño-Londoño *et al.*, 2016), el cuerpo de investigación está conformado por la transcripción de entrevistas semiestructuradas, y para la justificación del problema de investigación, se centraron en los indicadores de salud de las comunidades wayuu, para lo cual utilizaron datos publicados por la Organización Mundial de la Salud, datos estadísticos del Ministerio de Salud y Protección Social de Colombia, del Censo poblacional, de leyes nacionales, etc.

El acceso a estas fuentes escritas no suele realizarse de forma sistematizada. Es posible que busquemos específicamente algunos datos en ciertos organismos que se dediquen al estudio específico del tema o recuperemos discusiones teóricas, utilizando la técnica de bola de nieve que describimos en el apartado sobre “estado del arte”.

Marco teórico-conceptual

En este caso, la función de las fuentes escritas es explicitar la tradición teórica y los conceptos dentro de esa tradición desde los cuales se va a abordar el problema planteado. Según Minayo, una teoría es “un conjunto coherente de proposiciones, que interrelaciona principios, definiciones, tesis e hipótesis y sirve para organizar lógicamente la interpretación de la realidad empírica” (Minayo, 2013, p. 145).

En este sentido, elegir y adoptar una tradición teórica desde la cual interpretar un recorte de la realidad empírica implica recuperar los trabajos escritos por quienes crearon o redefinieron los conceptos centrales de esa teoría. Por ejemplo, si dentro de la tradición teórica de la Escuela de Fráncfort, adoptamos la noción de “razón técnica” que Herbert Marcuse discute en *Industrialización y capitalismo en Max Weber* (1969), no deberíamos citar el concepto mencionado por otros autores o por otras autoras sino, sino que deberíamos recuperar la obra de Marcuse. Incluso, para reconstruir las discusiones que recuperan posteriormente esa noción, podríamos recurrir a la relectura de la obra de Marcuse que realiza el filósofo y sociólogo alemán Jürgen Habermas en *Ciencia y técnica como ideología* (2007).

Por la propia extensión que requieren los desarrollos teóricos, es más probable que debamos consultar libros o capítulos de libros, a los que podemos acceder a través de lecturas previas, seminarios específicos sobre determinadas corrientes teóricas, o como propuesta de quien dirige nuestro trabajo de investigación. Otra forma de acercarnos a ciertos desarrollos teóricos que puedan ser de nuestro interés es recorrer los artículos recuperados en las búsquedas exploratorias iniciales, para explorar la autoría de los conceptos mencionados.

Estado del arte

Uno de los puntos que señala Umberto Eco al definir qué es la científicidad en una investigación es que “la investigación tiene que decir sobre este objeto cosas que todavía no han sido dichas o bien revisar con óptica diferente las cosas que ya han sido dichas” (1992, p. 49). ¿Y cómo sabemos lo que ya ha sido dicho sobre el problema de investigación que nos proponemos abordar? Parte de este camino ya lo deberíamos haber transitado en las búsquedas de textos iniciales realizadas en la etapa exploratoria. Pero para redactar el estado del arte, esas búsquedas exploratorias no nos alcanzan. La escritura del estado del arte implica la explicitación narrativa de ese conocimiento precedente,

y eso demanda una selección, una delimitación y una organización textual específica de lo que vamos a considerar como el núcleo del estado del arte de nuestro problema de investigación.

En el caso de que nuestro *tema* ya haya sido muy abordado en la bibliografía científica y estemos formulando un *problema* desde una perspectiva teórica que articula una mirada distinta e innovadora sobre ese tema que ya tiene una larga tradición, quizás pueda ser interesante centrar el estado del arte, no sobre el tema, sino sobre esas distintas miradas que ha tenido el problema.

Supongamos que nos interesara trabajar el tema “consumo de fármacos psiquiátricos”, y en las búsquedas y lecturas exploratorias detectamos que ya ha sido bastante tratado en la bibliografía, sobre todo desde la perspectiva psiquiátrica, como un problema a ser abordado desde el ámbito profesional e institucional, sin integrar la perspectiva de las personas que reciben el tratamiento. Pero supongamos que en esas mismas búsquedas iniciales detectamos otros tipos de trabajos como, por ejemplo, el artículo “Subjetividad y autonomía: significados y narrativas sobre la discontinuación de fármacos psiquiátricos” (2018), en el que Tatiana Parada aborda el tema desde otro marco referencial, y parte de las personas que deciden abandonar la medicación como una decisión autónoma. Al reconstruir los antecedentes, recupera los estudios que “han abordado los testimonios de personas que han vivido la experiencia de la psiquiatrización” y aquellas investigaciones más recientes que “se han enfocado en abordar la dimensión subjetiva del consumo de psicofármacos, enfatizando la percepción de efectos adversos, así como las estrategias que posibilitan la suspensión de la medicación psiquiátrica” (Parada, 2018, p. 514). Es decir, la reconstrucción del estado del arte no se centró en la concepción más generalizada del tema, sino que buscó e incorporó aquellas investigaciones que recuperaban otra perspectiva.

Esta es la diferencia entre la reconstrucción más general sobre el tema que se suele hacer en las búsquedas exploratorias iniciales y las decisiones narrativas que tendremos que adoptar a la hora de escribir el estado del arte. Esta aclaración es importante porque es muy probable que esta redefinición requiera realizar nuevas búsquedas de bibliografía. Por su grado de especificidad, estas nuevas búsquedas pueden realizarse también a través de la técnica de “bola de nieve”: partimos de unos pocos textos muy pertinentes para abordar nuestro problema y, luego de leerlos, recuperamos del listado de referencias bibliográficas aquellos trabajos citados que aborden el problema desde la perspectiva que estamos intentando reconstruir. Esta misma operación la reiteramos con los nuevos textos que vayamos incorporando

hasta conformar una de las tantas tramas de conversaciones previas sobre nuestro problema de investigación.

Metodología

Las fuentes citadas en la metodología suelen fundamentar las elecciones adoptadas respecto del tipo de diseño de estudio; la conformación de la muestra o el corpus; los canales en los que se obtuvieron los datos cuantitativos o las fuentes escritas, orales o gráficas; los instrumentos de recolección de los datos; las categorías y las técnicas de análisis, etc. Y estas elecciones suelen estar estrechamente relacionadas con el marco teórico adoptado en la investigación.

Por ejemplo, en el trabajo “Geografías para un buen nacer: una aproximación geo(corpo)gráfica del parto en domicilio en Chile”, en las primeras líneas del apartado metodológico, en las que se suele justificar la elección del diseño de estudio, Pía Rodríguez-Garrido menciona:

Considerando la importancia de poner en el centro las narrativas y experiencias de vida de las mujeres, es que este estudio se encuadra en una metodología cualitativa de investigación^(39,40) con elementos del paradigma fenomenológico feminista, ya que otorga “una comprensión que surge de una interrelación ontológica que no solo ofrece la explicación de la experiencia incorporada, sino también, como una percepción encarnada que subyace a la producción de conocimiento y fundamentos políticos”⁽⁴¹⁾. (Rodríguez-Garrido, 2022, p. 4)

Las fuentes citadas por la autora en ese primer párrafo de la metodología son libros sobre métodos cualitativos de investigación, pero no sobre cualquier escuela o corriente de pensamiento, sino centrados en la perspectiva fenomenológica y, más específicamente, en el paradigma fenomenológico feminista. En este caso, la autora puso en diálogo la teoría y la técnica.

El acceso a estas fuentes suele darse a través de lecturas previas, seminarios específicos sobre metodologías vinculadas a determinadas corrientes teóricas, o como propuesta de quien dirige nuestro trabajo de investigación. Otra forma de acercarnos a ciertas técnicas propuestas por líneas teóricas específicas es recorriendo el apartado metodológico de los artículos recuperados en las búsquedas exploratorias iniciales. Este suele ser un buen punto de partida para explorar nuevos recursos teórico-metodológicos.

Resultados

La centralidad de las fuentes citadas en este apartado se corresponde con ese recorte de datos cuantitativos o de fuentes escritas, orales o gráficas (publicadas o no) desde el cual proponemos abordar el problema de investigación y sobre el cual vamos a llevar a cabo el análisis.

Desde ciertas áreas suelen analizar los fragmentos extraídos de las fuentes del corpus desde la perspectiva teórico-conceptual e incluso con el estado del arte del problema, para lo cual recuperan las fuentes citadas en los apartados anteriores. Por ejemplo, en el artículo “Sufrimiento de mujeres en una ocupación urbana de São Paulo, Brasil: un desafío para el trabajo en atención primaria de la salud”, en el apartado de Resultados, Aline Messias Mota y Rubens Bedrikow analizan las transcripciones de las personas entrevistadas, retomando la perspectiva teórica que presentaron en la Introducción del artículo:

Esos fragmentos explicitan que, aunque la presencia del Estado garantice algunos servicios básicos, son transitorios, insuficientes y no eliminan los riesgos a la salud. Según Rol-nik⁽³²⁾ el acto de ceder terrenos y servicios no es nuevo o exclusivo de Brasil, y tiene como fondo el mantenimiento de una relación de dependencia por parte de los vulnerables. (Mota & Bedrikow, 2021, p. 8)

En este caso, esa referencia a Raquel Rolnik fue utilizada al presentar los conceptos o categorías centrales de la investigación al comienzo del artículo, y es retomada en los resultados para analizar en términos conceptuales las transcripciones de los relatos que surgieron en las entrevistas en profundidad.

Discusión

En este apartado se retoman los resultados obtenidos en la investigación para ponerlos en discusión las diversas perspectivas presentadas en el estado del arte, es decir, los resultados se enriquecen al ponerlos en discusión con las conversaciones precedentes sobre el problema. Se suelen invocar textos que fueron citados en los apartados anteriores, sobre todo, aquellos textos que plasman las conversaciones previas sobre nuestro problema de investigación.

Si bien gran parte de los trabajos que participan en la discusión deberían haber sido recuperados en las etapas iniciales, cuando la investigación se reduce para ser publicada en formato de artículo, la estructura narrativa suele tener ciertas diferencias según las distintas culturas escritas dentro

de la comunidad científica. En algunos casos daría la sensación de que hay referencias que aparecen por primera vez en la discusión, aunque esto no sea realmente así. Al recortar la investigación y presentarla en formato de artículo, desde algunas áreas disciplinarias hacen más hincapié en el estado del arte, por lo que incluyen un apartado de antecedentes al inicio del artículo y suelen omitir la discusión. Sin embargo, desde otras áreas disciplinarias, utilizan la introducción para la justificación social del problema e incorporan los “antecedentes” en el apartado de discusión, luego de los resultados, en el cual discuten los hallazgos con la bibliografía preexistente sobre el tema. En estos casos, si bien parecería que esa bibliografía fue recuperada con posterioridad a obtener los resultados, es una ilusión creada por la trama narrativa: los textos que se retoman para poner en discusión los resultados deberían haber sido recuperados en las etapas iniciales de la investigación, dado que son los que permiten configurar y recortar el problema de estudio como tal.

Capítulo 2

Investigaciones basadas en fuentes escritas

Si decidimos adoptar las fuentes documentales como fuentes primarias de nuestra investigación se abre un universo muy vasto de posibilidades en cuanto al tipo de fuentes a analizar. Se pueden conformar cuerpos de investigación con documentos legislativos, artículos periodísticos, fallos judiciales, documentos técnicos de organismos nacionales e internacionales, registros de interacciones en redes sociales, pautas publicitarias gráficas, fotografías, artículos de investigación científica, entre tantos otros.

En todos estos casos, no hay un contacto directo con quienes produjeron esas fuentes ni con las instituciones, los eventos, los procesos o las personas analizadas, sino que el contacto es con el objeto textual, audiovisual, gráfico o fotográfico, que pertenece a ciertas comunidades de diálogo, las cuales crean arquitecturas discursivas propias que, en términos de Foucault, están conformadas por “un conjunto de reglas que son inmanentes a una práctica” (2002, p. 76). De allí que, por ejemplo, sobre un tema como la “locura”, el ámbito periodístico construye una estructura narrativa y una discursividad muy distinta a la de un fallo judicial, un artículo de investigación o un proyecto de ley. Los diversos actores sociales (poder legislativo, medios periodísticos, comunidad científica, poder judicial, fuerzas policiales) abordan problemas desde miradas y perspectivas diferentes, y generan producciones discursivas con propósitos y características propias. Cada uno de estos “géneros discursivos”, en términos de Bajtin (2011), o “formaciones discursivas”, en términos de Foucault (2002), al tener formatos y reglas propias de circulación pública, conforman dispositivos de comunicación y comunidades discursivas factibles de ser analizadas.

Los procesos de recuperación de las fuentes para la conformación de un cuerpo de investigación varían según el tipo de fuente, la pregunta de investigación y las categorías analíticas que van a delimitar también la pertinencia del recorte temporal de cada investigación. Daremos algunos ejemplos.

Uno de los objetivos de Paula Guitelman (2006) fue indagar “el tipo de subjetividad infantil que la revista *Billiken* conformó en el plano discursivo durante la última dictadura”¹. En este caso, el cuerpo de investigación va a estar conformado por las revistas *Billiken* publicadas en Argentina entre los años 1976 y 1978, y justifica el recorte del siguiente modo: “dicho trienio fue el momento en el que se produjeron la mayor parte de las desapariciones de personas, y también cuando el régimen dictatorial tuvo menor oposición”, es decir, el recorte está estrechamente vinculado al problema y a las categorías de análisis de la investigación. La forma de recopilar y conformar este cuerpo bien definido de revistas impresas, que pueden no estar digitalizadas, es necesariamente distinta a la forma de reconstruir otros cuerpos documentales. Si no contara con la colección, podría consultar los ejemplares en el acervo de alguna biblioteca pública.

María Julieta Oddone (2013) se propuso analizar:

...los mensajes emitidos por los libros de lectura para la escuela primaria editados en Argentina, en el período comprendido entre los años 1880 y 2012 para profundizar sobre la imagen de la vejez y el envejecimiento que la sociedad tiene y transmite a las nuevas generaciones y el papel que se le asigna a este grupo generacional. (2013, p. 27).

En este caso, el cuerpo de investigación estuvo conformado por los libros de lectura escolar, editados en Argentina, utilizados “en segundo grado de escolaridad primaria a nivel nacional y, de estos, se seleccionaron aquellos más usados en las escuelas dependientes del Estado” (2013, p. 28). El período temporal es muy amplio dado que corresponde a una investigación realizada en distintas etapas y la organización interna está definida de la siguiente manera: “Los períodos históricos que dan contexto a los datos fueron definidos a partir de los ritmos marcados por la permanencia o los puntos de cambio (*turning point*) de los valores sociales transmitidos en las lecturas” (2013, p. 28). En este caso, la propia investigadora es quien va rastreando los libros que año a año llegan a las aulas, seleccionados por las comisiones federales de evaluación del gobierno nacional.

Otro ejemplo, es el trabajo de Anahí Farji Neer (2015) basado en las transcripciones taquigráficas de los debates parlamentarios sobre las leyes de

¹Se hace alusión a la última dictadura cívico-militar de Argentina que tuvo lugar entre los años 1976-1983. *Billiken* es una revista infantil semanal, creada por Constancio Vigil, cuyo primer ejemplar se publicó 1919.

identidad de género y fertilización asistida, que tuvieron lugar en el parlamento argentino durante los años 2011 y 2013, cuyo recorte temporal está definido por las propias fechas en que se realizaron las sesiones parlamentarias. Las transcripciones de las sesiones son documentos públicos, digitalizados, y de libre acceso a través de página web, por lo que acceder a estas fuentes no suele presentar grandes dificultades.

En los tres casos mencionados se utilizaron cuerpos documentales que no se encontraban conformados, sino que fueron creados como tales a partir de la apuesta creativa planteada por las propias investigadoras. Esos objetos textuales no se analizan como unidades aisladas, sino como un cuerpo textual producido por una comunidad narrativa particular, desde un género discursivo particular, que se inserta en momentos históricos, sociales y políticos determinados, es decir, el texto es entendido como un “dispositivo” a ser interpretado desde múltiples dimensiones, que porta mucha más información y sentidos que un “dato” publicado a ser abordado de forma descriptiva.

Pero estos cuerpos documentales ¿son las únicas fuentes que se utilizan en una investigación? Claramente no, la distinción entre las fuentes que conforman el cuerpo de investigación y las utilizadas en los restantes apartados suele suscitar dudas, por lo que es importante utilizar la etapa exploratoria de búsqueda bibliográfica inicial para delinear una primera definición del tipo de fuentes secundarias que se utilizarán en los distintos apartados que, a su vez, podrán retroalimentar y redefinir la propia propuesta de investigación.

En este gran universo de posibilidades que ofrecen las investigaciones basadas en fuentes documentales, es importante tener en cuenta que los diversos actores sociales abordan los temas desde miradas y perspectivas diferentes, y generan producciones discursivas con características muy diversas. A continuación, incluimos algunos ejemplos, según el género discursivo desde el cual fueron creados los documentos que conforman el corpus de investigación.

Ejemplos de investigaciones basadas en fuentes documentales

Género científico

Investigaciones basadas en cuerpos documentales poco tradicionales del género discursivo científico (registros de ensayos clínicos, historias clínicas, manuales, etc.):

- “El uso de placebo en ensayos clínicos de fase III en Brasil” (Rubenich *et al.*, 2015): cuerpo documental conformado por registros de ensayos clínicos, que verifica el grado de cumplimiento de ciertas normativas éticas.
- “El saber ver y el saber hacer en el hospital-asilo de Oliva (1914-1950): el caso Adela” (Vanadia & Rodríguez, 2019): cuerpo documental conformado por historias clínicas y publicaciones institucionales que analiza cómo los comportamientos de los pacientes y sus allegados, los valores y usos sociales definían tipos de ingresos, diagnósticos y posibilidades de tratamiento.
- “Factores de riesgo y bases genéticas: el caso del trastorno por déficit de atención e hiperactividad” (Martinhago *et al.*, 2019) : retoma como fuente escrita la caracterización del TDAH en el manual DSM-5 para discutir los numerosos intentos de describir las bases biológicas del trastorno para usarlos con fines de diagnóstico y como marcadores de riesgo.

Género legislativo

Investigaciones basadas en cuerpos documentales del género discursivo legislativo (versiones taquigráficas de discusiones parlamentarias, leyes, reglamentaciones, etc.):

- “¿Libertad de enseñanza? ¿Atentado contra la libertad de conciencia? Discusiones parlamentarias en torno al uso de los manuales del Hermano Damasceno en las aulas laicas del Uruguay de primera mitad del siglo XX” (García, 2019): retoma las transcripciones taquigráficas de las discusiones parlamentarias ocurridas en 1932, para analizar los argumentos relacionados con la separación jurídica de la Iglesia y el Estado.
- “¿Articulación imposible? Debates parlamentarios sobre el voto joven y el voto extranjero en Argentina” (Baidera, 2018): recupera las versiones taquigráficas de la Cámara de Senadores de la Nación, para analizar cómo se articularon las demandas emergentes para la ampliación del voto, de los jóvenes (argentinos o nacionalizados) de 16 a 17 años de edad y de los extranjeros con residencia permanente en Argentina.
- “Cuerpo, derechos y salud integral: Análisis de los debates parlamentarios de las leyes de Identidad de Género y Fertilización Asistida (Ar-

gentina, 2011-2013)” (Farji Neer, 2015): reúne las versiones taquigráficas de los debates parlamentarios para analizar cómo fue caracterizado el problema público al cual ambas leyes darían respuesta, cómo fue conceptualizada la misión de ambas leyes y en qué medida dicha definición implicó poner en entredicho las nociones de salud y enfermedad con relación a la cobertura de tratamientos de optimización o modificación corporal por parte del sistema público de salud.

Género periodístico

Investigaciones basadas en cuerpos documentales del género discursivo periodístico (editoriales, notas firmadas, etc.):

- “(De)construir la otredad: las mujeres inmigrantes en la prensa escrita española” (Creighton, 2013): a través de las noticias sobre inmigración publicadas en los periódicos españoles *El País* y *ABC*, durante los primeros seis meses de 2011, se analizan las imágenes creadas sobre las mujeres inmigrantes.
- “La construcción de la malnutrición infantil en la prensa escrita argentina durante la crisis de 2001” (Demonte, 2011): analiza la construcción de la malnutrición infantil en las notas publicadas en *La Nación*, *Clarín* y *Crónica* entre diciembre de 2001 y diciembre 2002, e identifica la persistencia de ciertas representaciones ligadas al problema de la malnutrición infantil, aun en contra de los datos epidemiológicos producidos durante el mismo período.
- “La representación social negativa de los procesos de salud/enfermedad/atención en la prensa escrita” (Menéndez & Di Pardo, 2008): analiza la información publicada por los diez principales periódicos mexicanos de circulación nacional y observa la constante tendencia a presentarla en términos negativos y sensacionalistas.

Género jurídico

Investigaciones basadas en cuerpos documentales del género discursivo jurídico (fallos de juzgados, cortes supremas nacionales, tribunales internacionales, etc.):

- “Derechos en el final de la vida: el rechazo de tratamientos médicos en los tribunales argentinos” (Alonso, 2016) se recuperan fallos de tribunales argentinos para analizar la judicialización de decisiones médicas en el final de la vida, en el marco del avance de la justicia en la regulación de la práctica médica y de un mayor reconocimiento de la autonomía de los pacientes.
- “Vida digna y derecho a la salud en los fallos de la Corte Interamericana de Derechos Humanos” (Iosa, 2021): se recuperan los casos contenciosos tramitados ante la Corte IDH en los que se constataron violaciones al derecho a la salud, para analizar los diferentes usos del concepto “vida digna” y reflexionar sobre las implicancias de estos usos en materia sanitaria.
- “La problemática del arsénico en el servicio de agua en la provincia de Buenos Aires, Argentina: Análisis de casos jurisprudenciales” (Minaverry *et al.*, 2016): se recuperan una serie de casos jurisprudenciales vinculados con la calidad del agua potable que se consume en el ámbito de la provincia de Buenos Aires para analizar la creciente contaminación del recurso hídrico y la judicialización de distintos conflictos por el acceso y calidad del agua.

Género publicitario

Investigaciones basadas en cuerpos documentales del género discursivo publicitario (piezas gráficas institucionales o comerciales, publicidades audiovisuales, etc.):

- “Modelos de envejecimiento en la publicidad gráfica: un análisis de género” (Pochintesta & Mansinho, 2014): conforma un corpus documental de 52 imágenes seleccionadas a partir del relevamiento de 144 números de *Revista Viva*, *La Nación* y *Gente* del año 2000, con el propósito de identificar estereotipos de género en la vejez.
- “El humor en la publicidad gráfica de la Dirección General de Tráfico en España (1960-2009)” (Ojeda *et al.*, 2015): conforma un corpus documental de 246 anuncios, integrado por toda la publicidad gráfica de la institución desde el año 1959 al 2009, para analizar el humor como modalidad semántica en las campañas publicitarias e identificar las emociones expresadas en los anuncios que utilizan este recurso.

- “Pequeños grandes clientes: La publicidad de sucedáneos de la leche materna en dos revistas pediátricas de Argentina entre 1977 y 2006” (Vallone, 2009): se conformó un cuerpo de publicidades de sucedáneos de leche materna aparecidas en dos colecciones de revistas pediátricas argentinas publicadas entre 1977 y 2006 con el propósito de analizar el rigor científico en la información brindada en las publicidades y las estrategias de comercialización que desalientan las prácticas de amamantamiento.

Géneros digitales

Investigaciones basadas en cuerpos documentales de géneros discursivos digitales (posteos o interacciones en las redes sociales, piezas gráficas, piezas audiovisuales, etc.):

- “*Mapping social dynamics on Facebook: The Brexit debate*” (Mapeo de la dinámica social en Facebook: el debate sobre el Brexit) (Del Vicario *et al.*, 2017): se recuperan las interacciones de 1 millón de usuarios de Facebook, para analizar la polarización y la evolución de las narrativas centrales en los debates en línea relacionadas al Brexit.
- “Política y redes sociales en Argentina. El caso de los candidatos presidenciales de 2011 en Twitter” (Slimovich, 2016): corpus formado por las cuentas de Twitter de los candidatos a la presidencia de las elecciones Argentina 2011 para analizar sus estrategias mediáticas.
- “Violencia en la red social: una indagación de expresiones online en adolescentes de sectores populares marginalizados del Área Metropolitana de Buenos Aires” (Linne & Angilletta, 2016): se construyó un corpus de 3.000 observaciones virtuales de perfiles de la red social Facebook, para analizar las expresiones online de violencia ejercida o padecida por las y los adolescentes de sectores populares marginalizados del Área Metropolitana de Buenos Aires.

Estos son tan solo algunos pocos ejemplos de la gran variedad de posibilidades de corpus de fuentes documentales. En ninguno de estos ejemplos hay un contacto directo con los sujetos, las instituciones, los eventos o los procesos a ser analizados. Ese contacto está intermediado por diversas circunstancias, entre ellas, los intereses y la perspectiva de quienes se formularon la pregunta de investigación, se posicionaron en determinado marco conceptual,

seleccionaron una metodología determinada para encontrar una posible respuesta a su pregunta, y plasmaron un recorte de todo ese proceso en un formato textual determinado. Tal como menciona Moscovici:

La comunicación nunca se reduce a transmitir los mensajes originales o a transportar informaciones inmutables, sino que diferencia, traduce, interpreta, combina, así como los grupos inventan, diferencian o interpretan los objetos sociales o las representaciones de los otros grupos. (1979, p. 18)

Es necesario tener en cuenta, además, que el corpus es una construcción, un recorte intermediado, cuyo análisis debería tener en cuenta a los sujetos y procesos de esa mediación para abordar la realidad que intenta representar. En términos de Caravedo:

El resultado no es una construcción mimética de la realidad, por naturaleza multiforme y desordenada [...] Paradójicamente, para servir a los fines científicos, el corpus termina siendo una construcción artificial, que presenta mutilada, parcializada o innaturalmente condensada, la realidad compleja que pretende captar. (Caravedo, 1999, p. 65)

Características del corpus documental de investigación

El corpus de la investigación es el conjunto de materiales discursivos (orales o escritos, publicados o sin publicar) que conforman el objeto de estudio de una investigación. Tal como menciona Adriana Bolívar “no cualquier colección de materiales es un corpus” (Bolívar, 2013). Una cosa es reunir materiales de interés y otra es compilar textos, fotografías o mapas a partir de alguna pregunta de investigación para explicar un fenómeno discursivo o un problema social.

Las decisiones respecto de cómo delimitar o constituir un corpus son complejas porque detrás de todo corpus hay un tipo de abordaje disciplinar, una posición teórica y un problema de investigación que puede ser de naturaleza muy variada: “tanto sus formas como sus materiales (si se nos permite esta riesgosa distinción analítica) deberán ser el resultado (siempre provisorio) de un proceso de indagación y de análisis” (Aguilar *et al.*, 2014).

Para lograr que ese *corpus* tenga un cierto grado de homogeneidad que permita ser analizado a través de categorías y dimensiones extraídas del

propio material o propuestas previamente, se requiere que las fuentes admitan ser interrogados por ciertos ejes comunes; es decir, ese corpus, para constituirse como tal y poder ser analizado, debería responder a ciertos principios de representatividad, homogeneidad y consistencia.

Representatividad

Se refiere al alcance del corpus, a la capacidad de dar cuenta en términos temporales, geográficos y conceptuales de la pregunta y del tipo de análisis propuesto en la investigación. En algunos casos, el corpus puede ser totalmente representativo del recorte propuesto. Por ejemplo, si la investigación se centra en las revistas *Billiken* publicadas entre 1976 y 1978 y compilamos los ejemplares publicados entre esos años, el cuerpo documental creado será representativo con relación a los objetivos propuestos; pero, si en cambio, lográramos recuperar solo las revistas *Billiken* del año 1976, deberíamos reformular el proyecto de investigación, porque el corpus ya no sería representativo del trienio “en el que se produjeron la mayor parte de las desapariciones de personas, y también cuando el régimen dictatorial tuvo menor oposición” (Guitelman, 2006).

Hay otros casos en los que no es posible mensurar el número total de ejemplares por lo que necesariamente cambia la noción de representatividad. Por ejemplo, la cantidad de artículos científicos publicados sobre un determinado tema es un número desconocido e imposible de determinar, por lo que la noción de representatividad se relaciona necesariamente con diversas dimensiones vinculadas a la propia investigación y se ciñe a los trabajos que se han podido recuperar dentro del universo de *documentos indizados en las bases de datos seleccionadas*, a través de ciertos instrumentos de búsqueda diseñados *ad hoc*, y es allí donde las formas de recuperar esos artículos comienza a cobrar cierta relevancia.

Homogeneidad

Los textos seleccionados deberían responder a un mismo *género discursivo* (periodístico, jurídico, legislativo, literario o científico), es decir, que estén escritos por una misma comunidad interpretativa, tengan un tipo de estructura narrativa que admita ser interrogada por ciertos ejes comunes, a través de categorías o conceptos extraídos de las propias fuentes o de nuestro marco conceptual. Por ejemplo, un cuerpo de notas periodísticas, un conjunto de

fallos judiciales sobre un tema, una serie de transcripciones de discusiones parlamentarias, cada conjunto en su interior, mantienen, además de una estructura narrativa similar, ciertos usos léxicos y retóricos propios de cada una de las comunidades interpretativas.

Consistencia

Dentro de esa homogeneidad, la consistencia se refiere al tipo de objeto seleccionado dentro de un mismo género discursivo. Por ejemplo, dentro de la comunidad científica existe un abanico muy amplio de comunidades interpretativas que generan objetos textuales muy diversos (abordajes jurídicos; ensayos clínicos; etnografías; estudios históricos, epidemiológicos o sociológicos; etc.) y cada uno de estos tipos de estudios y de abordajes condicionan el tipo de análisis a realizar. El género periodístico también se expresa a través de distintos lenguajes, como el fotográfico, el audiovisual, el escrito, etc., y los recursos metodológicos para analizar un cuerpo de fotografías son distintos que los utilizados para abordar un video o una nota escrita.

Capítulo 3

¿Qué es una revisión bibliográfica?

Como vimos en el capítulo anterior, existe una amplia gama de investigaciones basadas en fuentes escritas que no se autodenominan como “revisiones bibliográficas”, sino que son investigaciones que reconstruyen cuerpos textuales que son abordados desde distintas perspectivas teórico-metodológicas. Sin embargo, desde ciertas áreas disciplinarias, el análisis de un cuerpo de artículos científicos parecería quedar indefectiblemente asociado a la noción de *revisión bibliográfica* o incluso a técnicas específicas como PRISMA (*Preferred Reporting Items for Systematic reviews and Meta-Analyses*), diseñadas desde marcos referenciales que buscan “evaluar la certeza de la evidencia” (Page *et al.*, 2021). Como la racionalidad presente en este tipo de técnicas se orienta exclusivamente a la correcta elección de procedimientos técnicos, tal como lo expresa Habermas (2007), lo que en realidad se produce es un ocultamiento de la trama social de intereses y una anulación de la dimensión políticas del problema analizado, para centrar la racionalidad solo en situaciones de empleo posible de la técnica. Por lo tanto, entendemos que resulta contradictorio extrapolar estas técnicas de forma acrítica hacia investigaciones que intentan abordar problemas sociales desde marcos referenciales que reivindicarían dimensiones históricas, políticas y culturales.

Según Elizabeth Smythe y Deborah Spence (2012), en el mundo académico parece darse por sentado que hay una sola forma de hacer una revisión bibliográfica. Sin embargo, esta denominación engloba una gran diversidad de tipos de estudios, y cada una de esas técnicas se posiciona desde un referencial teórico desde el cual propone observar, enumerar, clasificar o cuantificar cierta información; como así también delimitar, excluir o crear parámetros para conformar un determinado recorte, y no siempre se expresa ese referencial de manera explícita.

A diferencia de las investigaciones que construyen un cuerpo de fuentes primarias (escritas, orales o gráficas) para ser abordadas desde ciertas categorías analíticas, las revisiones bibliográficas tienen la impronta de sistematizar y sintetizar lo publicado sobre un tema en particular.

Maria Grant y Andrew Booth (2009) enumeran una serie de técnicas como, por ejemplo, revisión crítica (*critical review*); revisión bibliográfica (*literature review*); metaanálisis (*meta-analysis*); revisión sistemática cualitativa (*qualitative systematic review*); revisión del estado del arte (*state-of-the-art review*); entre muchas otras. Del mismo modo, Jong-Myon Bae en su artículo titulado “*Narrative reviews*” (2014) recupera una serie de técnicas cualitativas, las cuales nuclea en cuatro grupos:

- 1) Énfasis en síntesis cualitativa:
 - Resumen narrativo (*narrative summary*)
 - Síntesis temática (*thematic synthesis*)
 - Síntesis narrativa textual (*textual narrative synthesis*)
 - Síntesis interpretativa crítica (*critical interpretative synthesis*)
 - Síntesis del marco teórico/conceptual (*framework synthesis*)
 - Síntesis realista (*realist synthesis*)

- 2) En contraste con la revisión sistemática cuantitativa
 - Revisión sistemática cualitativa (*qualitative systematic review*)
 - Síntesis de la investigación cualitativa (*synthesis of qualitative research*)
 - Revisión sistemática narrativa (*narrative systematic review*)
 - Análisis comparativo cualitativo (*qualitative comparative analysis*)
 - Meta-síntesis cualitativa (*qualitative meta-synthesis*)
 - Meta-resumen cualitativo (*qualitative meta-summary*)

- 3) En contraste con el metaanálisis
 - Metasíntesis (*meta-synthesis*)
 - Revisión metanarrativa (*meta-narrative review*)
 - Metaetnografía (*meta-ethnography*)
 - Metaestudio (*meta-study*)
 - Metainterpretación (*meta-interpretation*)

- 4) Incluyen investigaciones cuantitativas y cualitativas
 - Métodos mixtos (*mixed methods research*)
 - Triangulación ecológica (*ecological triangulation*)

Técnicas de revisión bibliográfica

Todas estas formas de revisión tienen en común una impronta más de síntesis descriptiva, que interpretativa. El deseo de abarcar el crecimiento exponencial de la producción científica dio lugar a la proliferación de una serie de técnicas de revisión, que se incrementaron considerablemente con el acceso remoto a un gran número de registros electrónicos. Cuando la técnica pierde su historia y su referencial teórico, suelen generarse nuevos nombres para prácticas e instrumentos preexistentes. De hecho, hay algunas denominaciones como, por ejemplo, revisión bibliográfica, análisis bibliográfico o análisis documental, que están presentes en la bibliografía, en algunos casos, desde el siglo XVIII.

La Tabla 1 muestra una selección de técnicas de revisión, según los tipos de estudios que admiten en el corpus y sus principales referentes. Algunas de las técnicas señaladas tienen su origen en determinados marcos conceptuales como, por ejemplo, la revisión sistemática, que se sustenta en la medicina basada en la evidencia, o las técnicas cuantitativas que se centran en la medición de la innovación científica como motor de la economía. Si bien muchas de estas investigaciones se sustentan en una matriz de corte positivista, ninguna de ellas suele explicitar un plano teórico-conceptual desde el cual analizar los resultados.

A diferencia de las técnicas de revisión bibliográfica, que se abocan al análisis de contenido de un cuerpo de textos, la bibliometría no suele analizar el texto, sino que se centra en los metadatos (los datos que conforman un registro bibliográfico), los cuales proveen información acerca de dónde, cuándo y quiénes abordaron un tema en particular, en qué marcos institucionales, quién otorgó el financiamiento, etc. Sin embargo, las técnicas de “minería de texto” basadas en la lingüística han abierto un abanico muy vasto de posibilidades al incorporar el procesamiento de lenguaje natural (NLP, por sus siglas en inglés), un tipo de análisis asistido por sistemas de lenguaje humano. Estos sistemas pueden “extraer conceptos de forma inteligente (incluidas frases compuestas). Además, el conocimiento del lenguaje subyacente permite la clasificación de conceptos en grupos relacionados (como, por ejemplo, productos, organizaciones o personas) utilizando el significado y el contexto” (IBM, 2021).

En este sentido, las técnicas de revisión aportan tan solo los instrumentos para reconstruir un escenario posible de una temática particular y, al hacerlo a través de los canales de distribución, el resultado queda necesariamente atravesado y condicionado no solo por las asimetrías en cuanto a los recursos económicos públicos y/o privados destinados a la investigación

Tabla 1. Principales técnicas de revisión de bibliografía científica, según tipos de artículos analizados, descripción y principales referentes.

Nombre de la técnica	Tipo de artículos del corpus	Descripción	Principales referentes
Revisión bibliográfica	Todo tipo de estudios	Recupera la bibliografía publicada recientemente sobre un tema como forma de sintetizar y actualizar el estado del arte.	Está presente en la bibliografía desde el siglo XVIII
Análisis bibliométrico	Todo tipo de estudios	Aplica métodos estadísticos para analizar la producción científica y generar indicadores (por ejemplo, de producción, de circulación, de dispersión, de uso, de visibilidad o impacto, de colaboración, de obsolescencia, etc.).	Lotka (1926) Bradford (1934) Garfield (1955) De Solla Price (1963)
Cienciometría	Todo tipo de estudios	Aplica métodos estadísticos para analizar el impacto económico de la innovación científico-tecnológica	Manual de Frascati (OECD, 1963) Manual de Oslo (OECD, 1992) Manual de Camberra (OECD, 1995) (OECD, 1992)
Metaanálisis	Estudios cuantitativos	Integra los hallazgos de estudios individuales a través de la aplicación de métodos estadísticos.	Glass (1976)
Revisión sistemática	Ensayos clínicos aleatorizados, prácticas clínicas	Reúne evidencias que se correspondan con criterios de elegibilidad establecidos previamente, evalúa la validez de los resultados de los estudios incluidos, sintetiza las características y resultados de los estudios incluidos.	The Cochrane Collaboration (1993) <i>Preferred Reporting Items for Systematic Reviews and Meta-Analyses</i> (PRISMA) (2009)
Metaetnografía	Estudios cualitativos	Compara y analiza textos creando nuevas interpretaciones en el proceso.	Noblit & Hare (1988)
<i>Scoping review</i>	Todo tipo de estudios	Realiza un mapeo rápido de los conceptos claves que sustentan un área de investigación, las principales fuentes y evidencias disponibles	Arksey & O'Malley (2005)
Síntesis interpretativa crítica	Estudios cualitativos	Integra conceptos y teorías presentes en los artículos analizados. El análisis interpretativo es conceptual, y el producto principal no es la agregación de datos, sino de teorías.	Dixon-Woods <i>et al.</i> (2006)

Fuente: Elaboración propia.

de cada temática o a la disparidad entre países respecto de la inversión en investigación y desarrollo, sino también por las propias prácticas de publicación, por el interés o el desinterés de quienes investigan y quienes editan revistas en distribuir sus trabajos a través de bases de datos y, sobre todo, por nuestros propios conocimientos acerca de cómo usar las bases de datos y cómo reconstruir un recorte que enriquezca el problema a ser abordado. Si bien el cuerpo de textos recuperados será siempre parcial respecto de lo que se investiga sobre el tema, en ocasiones, además de parcial puede ser sesgado, si no se hace un esfuerzo adicional de compensar la escasa participación de determinadas regiones del mundo en ciertas bases de datos, o de recuperar el abanico de términos que reconstruya las distintas formas de nominar el fenómeno estudiado.

C
L
F
B
V
AN
V
Z
PO
W
W
S
V
L
G
E
O

Capítulo 4

Elección de las bases de datos de bibliografía científica

Si bien las prácticas para la obtención de datos primarios se centraban, en términos históricos, en la experimentación y la observación, tal como expresa Samaja (2007), ese “campo de observación” varía según la tecnología social imperante en cada sociedad, que habilita dispositivos observacionales muy diferentes. Desde esta perspectiva, podemos concebir una base de datos de bibliografía científica como un dispositivo observacional de la producción científica escrita:

...como toda tecnología no flota en un espacio social vacío de valores, podemos sostener que toda “fuente de datos” constituye una cierta configuración práctica *instituida* socialmente. Las vías de acceso a los hechos mismos plantean problemas de variadísima gama de asuntos y de distintos niveles de profundidad, la mayor parte de los cuales debieron ponerse en la mesa de discusión durante el proceso de elaboración del marco teórico [...] (Samaja, 2007, p. 257) (cursivas del original)

En este sentido, una buena pregunta de investigación puede obtener respuestas sesgadas si no se tienen en cuenta los complejos procesos de mediación que operan desde que un tema se torna un problema de investigación, se publica e ingresa a la esfera pública, hasta que llega a manos de quien lo interpreta. Esa trama social de intereses que actúan en las instancias mediadoras pueden potenciar un diálogo o silenciarlo, ampliar su difusión o limitarla.

Si bien existe una gran cantidad de bases de datos de bibliografía científica, no existe la base de datos que nos brinde acceso a todo el material publicado en el mundo. Cada base de datos nos ofrece un recorte parcial que deberíamos conocer para poder comprender los resultados que obtenemos en las búsquedas.

Existen bases multidisciplinarias que contienen solo las revistas científicas editadas por una editorial, como ScienceDirect que incluye solo el contenido de las revistas editadas por Elsevier, o las bases de las editoriales Sage, Springer o Nature Journals. Estas bases de datos no deberían utilizarse para componer el estado del arte sobre un tema o crear el corpus de artículos de una investigación, dado que contienen un recorte muy limitado. Solo tiene sentido que sean incorporadas si su inclusión estuviera justificada, como en el artículo de Grieneisen & Zhang (2012) que analizaremos más adelante.

Para la búsqueda y recuperación de bibliografía científica suelen utilizarse otro tipo de bases de datos denominadas *bases referenciales*, *índices de resúmenes*, o *agregadores* como, por ejemplo, Web of Science (WoS), Scopus, PubMed, Biblioteca Virtual en Salud (BVS), Scientific Electronic Library Online (SciELO), Red de Revistas Científicas de América Latina y el Caribe, España y Portugal (Redalyc), Directory of Open Access Journals, etc., que compilan registros bibliográficos del material publicado por diversas editoriales.

Estas bases de datos responden a distintos intereses económicos, políticos y sociales, por lo que proporcionan un recorte temporal, espacial, temático y conceptual que debería tenerse en cuenta tanto al momento de seleccionar qué tipo de bases se van a utilizar, como también al momento de analizar los resultados obtenidos en las búsquedas. Es bastante común encontrar investigaciones que utilizan solo una base de datos que cuenta con una muy baja cobertura de determinadas regiones del mundo, en las que se llega a conclusiones del tipo “EE.UU. sigue siendo el país líder en cuanto a bibliografía publicada”. Tal como menciona Cecília Minayo, “el endiosamiento de las técnicas produce o un formalismo árido o respuestas estereotipadas” (Minayo, 2012, p. 19).

Por ejemplo, Scopus tiene buena cobertura de Europa o América del Norte, pero la participación de otras regiones es muy baja (Figura 1). Sin embargo, esto no está necesariamente vinculado a las asimetrías en cuanto a los porcentajes del PBI destinados a investigación y desarrollo: para 2018, Corea del Sur invertía el 4,53% y Japón el 3,28% del PBI en investigación y desarrollo (Banco Mundial, 2022), valores que en muchos casos duplican los porcentajes invertidos en diversos países de Europa, y eso no se ve reflejado en el número de revistas incluidas o en el número de publicaciones. Esto muestra que, al analizar la producción a través de los canales de distribución, los resultados quedan atravesados por problemáticas propias de la distribución.

En el caso de la disparidad entre cantidad de revistas y cantidad de publicaciones que se visualiza en la Figura 1, los datos de la región de Asia muestran que los grupos de investigación están publicando en revistas que no son de su región, por lo que las políticas en ciencia y tecnología, muy

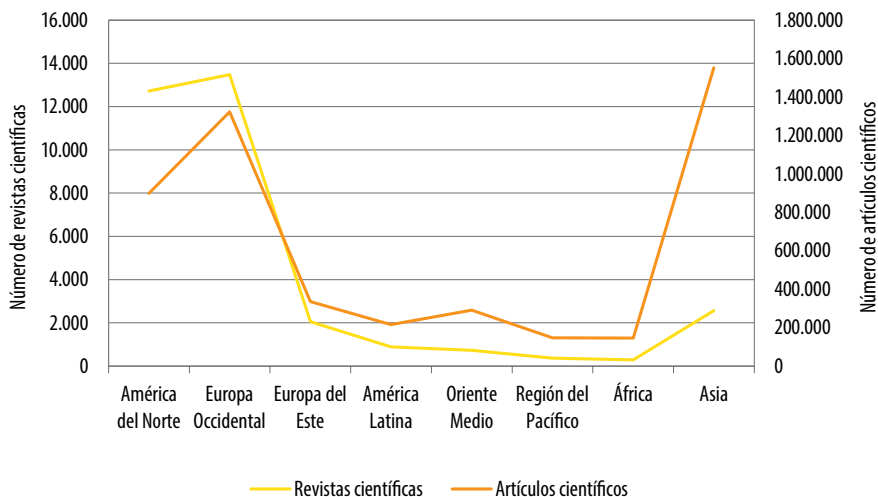


Figura 1. Número de revistas y de publicaciones científicas, según regiones geográficas. Scopus, 2020.

Fuente: Scimago Journal and Country Rank.

Nota: El término “artículos” incluye artículos de investigación, ensayos, editoriales, reseñas, revisiones.

probablemente no estén orientadas a fortalecer a sus propias revistas, sino dirigidas a fomentar la publicación en revistas de la gran industria editorial.

En este sentido, es interesante ver algunas investigaciones basadas en otro tipo de elecciones respecto de las bases de datos. En el trabajo titulado “*A comprehensive survey of retracted articles from the scholarly literatura*”, con el objetivo de analizar el fenómeno de las retractaciones¹, los autores decidieron consultar 42 bases de datos y las dividieron en tres categorías: a) bases de datos multieditoriales y agregadores, b) motores de búsqueda, y c) bases de datos de editoriales y de revistas científicas individuales (Grieneisen & Zhang, 2012). Y justifican la decisión de consultar un amplio espectro de bases de datos en las prácticas dispares de cada una de ellas:

¹La retractación es un mecanismo por el cual se retira un artículo ya publicado y se sustituye por una nota en la que se explicitan las razones por las cuales el artículo fue retractado. Según el Committee on Publication Ethics, las retractaciones pueden usarse para alertar a los lectores en casos de publicación redundante, plagio, manipulación del proceso de revisión por pares, reproducción del material o de los datos sin autorización, infracción de los derechos de autor u otros problemas legales, investigaciones no éticas y/o no revelar un interés competitivo importante que hubiera influido indebidamente en las interpretaciones y en las recomendaciones.

...aunque la base de datos X cubra la revista Y, es posible que la base de datos no identifique los artículos retractados por esa revista. Un ejemplo es la base de datos Inspec, que cubre una amplia gama de revistas de ingeniería, pero no indiza las retractaciones². (Grieneisen & Zhang, 2012)

Para poder detectar ese tipo de prácticas editoriales que pueden incidir en los resultados es necesario explorar, observar e identificar previamente los distintos actores que participan de esas múltiples instancias de mediación y analizar qué acciones pueden estar modificando o condicionando aquello que estamos intentando abordar. Las bases de datos no surgen por generación espontánea, sino que son construcciones sociales que responden a diversos intereses, valores y prácticas sociales. Descifrar qué prácticas pueden condicionar el resultado, es un interesante desafío que permitiría encontrar formas de minimizar posibles sesgos y enriquecer la discusión.

Otro ejemplo de adopción de bases de datos alternativas es el artículo “El uso de placebo en ensayos clínicos de fase III en Brasil” de Rubenich, Heck, Hellmann, & Schlemper Junior (2015), en el que a partir de los cambios normativos del año 2008, en Brasil, respecto de la restricción del uso de placebo en “las situaciones en las que no existieran métodos probados de profilaxis, diagnóstico o tratamiento”, el estudio se propone “conocer la frecuencia real del uso de placebo en los ensayos clínicos de fase III realizados en Brasil, tomando como punto de corte el año 2008”. Y para ello utilizan la base de datos ClinicalTrials.gov, desarrollada por la National Library of Medicine, que cuenta con alrededor de 250.000 registros provenientes de 201 países.

En el caso del estudio “*Reviewers’ ratings and bibliometric indicators: Hand in hand when assessing over research proposals?*” (Cabezas-Clavijo et al., 2013), los autores analizan críticamente la relación entre la evaluación de los postulantes a becas de investigación y los indicadores bibliométricos obtenidos por los investigadores, con el propósito de cotejar si esos indicadores afectan las posibilidades de los solicitantes de ser financiados, para lo cual cruzan los datos de los proyectos presentados ante la Agencia Nacional de Evaluación y Prospectiva de España y los indicadores bibliométricos calculados por el Journal Citation Reports.

En los tres ejemplos anteriores, la elección metodológica se plantea como instrumento a través del cual se pueden encontrar algunas respuestas a

²Traducción del original: *...even if database X covers journal Y, identifying the retracted articles from that journal in the database may not be possible. One example is the Inspec database, which covers a wide range of engineering journals but does not index retraction notices.* (Grieneisen & Zhang, 2012)

ciertos problemas formulados previamente, por lo que la técnica se adapta al problema y no a la inversa.

Para caracterizar las bases de datos más utilizadas, podríamos agruparlas por distintos ejes: el recorte temático, la cobertura temporal de lo publicado, si son públicas o pertenecen a iniciativas privadas, etc. En este caso, del gran universo de bases de datos, proponemos una selección de aquellas que permiten realizar búsquedas avanzadas, con posibilidad de incluir más de dos campos de búsquedas, que cuentan con filtros y formas de exportar los registros³.

Biblioteca Virtual en Salud (BVS)

- *Año de inicio:* 1979 (Index Medicus Latinoamericano)
- *Cobertura temática:* ciencias biomédicas, ciencias de la salud.
- *Cobertura geográfica:* América Latina y El Caribe
- *Cobertura temporal retrospectiva:* hasta 1982.
- *Acceso al contenido:* abierto.
- *Idioma de indización:* portugués, español, inglés y francés.
- *Idioma de búsqueda:* portugués, español, inglés y francés.
- *Tipo de contenido:* *Índices de resúmenes* (no muestra el texto completo del artículo). En los casos de artículos de acceso abierto, puede contener el enlace al texto completo publicado por la revista.
- *Enlace:* <https://pesquisa.bvsalud.org/portal/advanced/?lang=es>

A diferencia de las bases de datos que cuentan con una administración centralizada, la BVS es un portal integrador de recursos, conformado por una red de bibliotecas cooperantes pertenecientes a 30 países, coordinada por el Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud (BIREME), centro especializado de la Organización Panamericana de la Salud (OPS), con sede en San Pablo, Brasil. Su creación se remonta a fines de la década de 1970, cuando Latinoamérica comienza a generar redes cooperativas de circulación de la producción científica de la región, como el Index Medicus Latinoamericano (IMLA):

³Existen muchas otras bases de datos como la [Red de Revistas Científicas de América Latina y el Caribe, España y Portugal](#) (Redalyc), o el [Directory of Open Access Journals](#) (DOAJ), de gran valor bibliográfico, pero que no cuentan con interfaces de búsqueda avanzada.

El IMLA reunirá la producción científica de unas 250 publicaciones, presentadas en un formato único, actualizando la literatura médica de América Latina y poniéndola a disposición de los interesados de esta y otras regiones. Se contará así con un documento orgánico que permitirá conocer en forma global las publicaciones latinoamericanas del sector salud, hasta ahora dispersas. (Acuña, 1979)

A fines de la década de 1980, Bireme desarrolla el índice de Descriptores en Ciencias de la Salud (DeCS), que inicialmente contenía la traducción de los *Medical Subject Headings* (MeSH) al español y al portugués. Con el paso de los años, Bireme decidió ampliar el índice e incorporar nuevos descriptores, por lo que ambos índices dejaron de ser equivalentes. Esto es un punto muy relevante que puede generar ciertas confusiones, dado que existen DeCS que no puedan ser utilizados como MeSH. Por ejemplo, si nuestro tema se centrara en la provisión industrial o institucional de alimentos, en el índice DeCS encontraremos el descriptor “*collective feeding*”. Sin embargo, si realizamos en Pubmed una búsqueda con el término “*collective feeding*” dentro del campo *Mesh Term* el resultado sería “0” registros. Y esto no quiere decir que en PubMed no existan artículos que contengan la expresión “*collective feeding*”, sino que al no ser un *Mesh Term*, no existe ningún documento rotulado con ese término⁴.

Tal como puede verse en el [Apéndice](#), el portal de la BVS nuclea diversas bases de datos como Bibliografía Brasileña de Odontología (BBO); Banco de Datos de Enfermería (BDENF); Bibliografía Nacional en Ciencias de la Salud (BINACIS); Bibliografía Cubana (CUMED); Índice Bibliográfico Español en Ciencias de la Salud (IBECS); Literatura Latinoamericana y del Caribe en Ciencias de la Salud (LILACS); Literatura Peruana en Ciencias de la Salud (LIPECS), etc. Los registros de todas estas bases pueden ser recuperados a través de la interfaz de búsqueda del portal de la BVS (Figura 2).

Con relación a la cobertura temporal retrospectiva de la BVS, al nuclear diversas colecciones resulta difícil de determinar. En el caso de LILACS, una de las bases incluidas, contiene documentos publicados a partir de 1982, y se da “preferencia al procesamiento del material más reciente” (Bireme *et al.*, 2008).

⁴Este punto se desarrolla con mayor profundidad en el capítulo “¿Cómo buscar en una base de datos?”.



Figura 2. Interfaz de búsqueda avanzada de la Biblioteca Virtual en Salud, 2022.

Fuente: National Library of Medicine.

PubMed

- *Año de inicio:* 1879 (Index Medicus).
- *Cobertura temática:* ciencias biomédicas, ciencias de la salud.
- *Cobertura geográfica:* internacional (con muy baja participación de ciertas regiones del mundo).
- *Cobertura temporal retrospectiva:* hasta 1970.
- *Acceso al contenido:* abierto.
- *Idioma de indexación:* inglés.
- *Idioma de búsqueda:* inglés.
- *Tipo de contenido:* Índices de resúmenes (no muestra el texto completo del artículo). En los casos de artículos de acceso abierto, puede contener el enlace al texto completo publicado por la revista.
- *Enlace:* <https://pubmed.ncbi.nlm.nih.gov/advanced>

El origen de PubMed se remonta a 1879, cuando se publicó el primer volumen del *Index Medicus*, con 20.000 registros de publicaciones del área de ciencias de la salud (U.S. National Library of Medicine, 2011; 2016a). Desde un inicio, su estructura relacional se basó en “entradas temáticas”. Esa forma

de clasificación inicial se fue completando y complejizando hasta que, en 1954, se publica el primer listado oficial de “*subject headings*” bajo el título *Subject Heading Authority List* y luego, en 1960, se publica una versión revisada y actualizada bajo el título *Medical Subject Heading* (MeSH). Mientras en 1963, el listado contenía 5.700 descriptores, en 2004 pasó a tener 22.000 descriptores (U.S. National Library of Medicine, 2017a).

Si bien a partir de 1964, los registros impresos comenzaron a incluirse en sistemas computarizados que permitían la búsqueda y recuperación de bibliografía, recién en 1996 los registros comenzaron a convertirse, de forma paulatina, de copia impresa a un formato legible por máquinas (U.S. National Library of Medicine, 2017b). Hasta el momento, la cobertura retrospectiva llega a 1946. Los registros son públicos por lo que pueden ser consultados sin restricciones, pero como el propósito de la base de datos no es brindar acceso al texto completo, sino dar a conocer la existencia de las investigaciones incluidas en la base de datos, no todos los registros incluyen vínculos al artículo, sino solo aquellos registros alojados en PubMed Central, SciELO, o en los sitios web de editoriales que adhieran a políticas de acceso abierto. La Figura 2 muestra la interfaz a través de la cual podemos acceder a los registros.



Scientific Electronic Library Online (SciELO)

- *Año de inicio:* 1997
- *Cobertura temática:* multidisciplinaria.
- *Cobertura geográfica:* colecciones nacionales de 15 países y la colección temática Salud Pública.
- *Cobertura temporal retrospectiva:* depende de cada revista.
- *Acceso al contenido:* abierto.
- *Idioma de indización:* portugués, español e inglés (todos los contenidos, más allá del idioma de publicación, deben contar con título y resumen en inglés)
- *Idioma de búsqueda:* portugués, español e inglés.
- *Tipo de contenido:* artículos de acceso abierto a texto completo.
- *Enlace:* <https://search.scielo.org>

Para fines de la década de 1990, sobre la base de 20 años del desarrollo de proyectos colaborativos y en red coordinados por Bireme, América Latina contaba con canales de distribución de la producción vinculada a ciencias de la salud, pero no existían canales de distribución para el resto de las áreas de conocimiento. Por aquel entonces, las discusiones se centraban en torno a la falta de “visibilidad” de la producción latinoamericana. La única gran base de datos multidisciplinaria a nivel internacional era Web of Science, en la que la cobertura de la producción latinoamericana era muy baja.

En este contexto, surge en Brasil el proyecto *Scientific Electronic Library Online* (SciELO), en 1997, fruto de la colaboración entre la Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), el Centro Latinoamericano y del Caribe de Información en Ciencias de la Salud (BIREME) de la Organización Panamericana de la Salud (OPS) y una serie de editores científicos brasileños (Packer et al., 2006).

Se trata de un proyecto que planteaba la apertura del contenido a texto completo, para lo cual desarrolla una metodología de procesamiento descentralizada de los archivos. El modelo nace con la colección SciELO Brasil, que fue presentada públicamente en el *Seminário sobre Avaliação da Produção Científica*, en marzo de 1998 (Packer et al., 2006). Esta nueva modalidad de publicación aporta *visibilidad* a la producción latinoamericana y plantea un cambio en cuanto a dos de los principios básicos del acceso abierto: la *apertura*, entendida como la puesta a disposición inmediata de los contenidos

a texto completo en Internet, y la *gratuidad*, la accesibilidad en línea de la información científica sin costo para los lectores.

Desde el punto de vista estructural, es una biblioteca electrónica conformada por una red de colecciones de revistas científicas, organizadas por país. Actualmente existen 15 colecciones de países, más la colección SciELO Salud Pública. Si bien cada colección tiene su propia interfaz de búsqueda, a través de SciELO.org se puede acceder a los registros de todas las colecciones (Figura 3).

Su finalidad es proveer el contenido a texto completo, con acceso abierto, libre y gratuito, por lo que en ninguna de las colecciones hay registros que pertenezcan a revistas editadas bajo modelos cerrados de publicación.

Respecto a la cobertura temporal retrospectiva, como la responsabilidad del procesamiento del material para ser incorporado en la base de datos recae en los editores, la disponibilidad de números anteriores dependerá de la capacidad de cada revista de procesar su propio contenido.



Figura 3. Interfaz de búsqueda de SciELO.org, 2022.

Fuente: SciELO.org.

Scopus

- *Año de inicio:* 2004
- *Cobertura temática:* multidisciplinaria.
- *Cobertura geográfica:* internacional (con muy baja participación de ciertas regiones del mundo)
- *Cobertura temporal retrospectiva:* hasta 1970.

- *Acceso al contenido*: restringido. En América Latina, el acceso suele estar centralizado y administrado por el Estado, quien paga el canon y otorga permisos de acceso a las instituciones de investigación vinculadas al sistema nacional de ciencia y tecnología de cada país.
- *Idioma de indización*: inglés
- *Idioma de búsqueda*: inglés
- *Tipo de contenido*: *Índices de resúmenes* (no muestra el texto completo del artículo). En los casos de artículos de acceso abierto, puede contener el enlace al texto completo publicado por la revista.
- *Enlace*: <https://www.scopus.com>

Creada en el año 2004 por la editorial Elsevier, es la segunda base de datos multidisciplinaria más grande (en cantidad de registros) luego de Web of Science. En algo más de una década llegó a sumar más de 60 millones de registros. Según Susan Fingerman:

Cuando Elsevier introdujo Scopus en 2004, atrajo mucha atención. Esto se debió, por un lado, a un gran esfuerzo de marketing de su parte y, por otro, debido al hecho de que Scopus fue el primer gran desafío al predominio de Thomson ISI Web of Science (WoS)⁵. (Fingerman, 2006)

Al igual que WoS, el acceso a Scopus es pago y, en el caso de Argentina, forma parte de los recursos que la Biblioteca Electrónica en Ciencia y Tecnología pone a disposición de las universidades e institutos de investigación. Un 40% de su contenido corresponde al material publicado por la industria editorial científica. La Figura 4 muestra la interfaz a través de la cual podemos realizar las búsquedas.

Respecto a la cobertura retrospectiva de registros, a fines de 2016 se incorporaron cerca de 240 millones de referencias provenientes de 12 millones de artículos, lo que aumentó el alcance temporal hasta 1970 (Elsevier, 2016). Como se visualiza en la Figura 1, la cobertura de las distintas regiones del mundo es muy dispar: cuenta con una alta cantidad de registros de Europa y América del Norte, y una muy baja participación de regiones como Asia y América Latina.

⁵Traducción del original: *When Elsevier introduced Scopus in 2004, it attracted lots of attention. This was partly due to an enormous marketing effort on their part, and partly due to the fact that Scopus was the first major challenge to the supremacy of the Thomson ISI Web of Science (WoS).* (Fingerman, 2006)

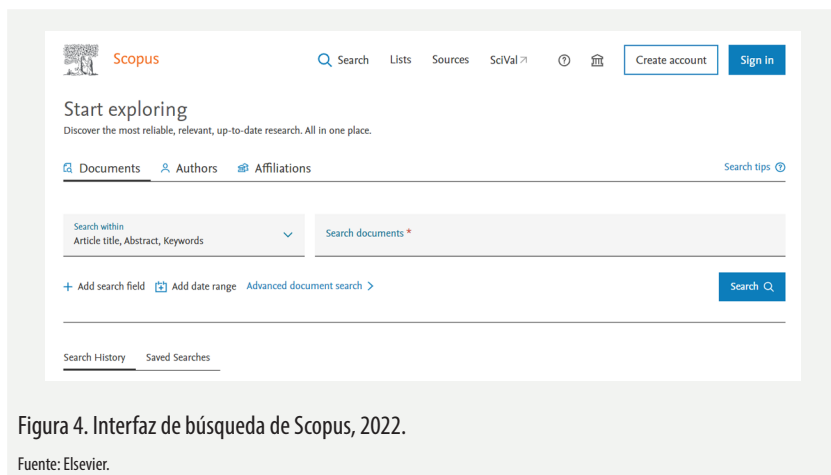


Figura 4. Interfaz de búsqueda de Scopus, 2022.

Fuente: Elsevier.

Web of Science

- *Año de inicio*: 1963
- *Cobertura temática*: multidisciplinaria.
- *Cobertura geográfica*: internacional (recientemente incorporó a Chinese Science Citation Database; Russian Science Citation Index; SciELO Citation Index, entre otros recursos)
- *Cobertura temporal retrospectiva*: hasta 1900 (varía según la cobertura temporal retrospectiva de cada una de las bases de datos incluidas, ver [Apéndice](#))
- *Acceso al contenido*: restringido. En América Latina, el acceso suele estar centralizado y administrado por el Estado, quien paga el canon y otorga permisos de acceso a las instituciones de investigación vinculadas al sistema nacional de ciencia y tecnología de cada país.
- *Idioma de indización*: inglés
- *Idioma de búsqueda*: inglés
- *Tipo de contenido*: *Índices de resúmenes* (no muestra el texto completo del artículo). En los casos de artículos de acceso abierto, puede contener el enlace al texto completo publicado por la revista.
- *Enlace*: <https://tinyurl.com/2p8tbkvw>

Se trata de una de las bases de datos corporativas más relevantes del siglo XX. Fue creada por Eugene Garfield en el marco del Institute for Scientific Information (ISI). La primera versión de *Science Citation Index* (SCI) fue

publicada en 1963 en formato impreso (Garfield, 1963). Luego, en 1972, se publica el *Social Science Citation Index* (SSCI) (Weinstock, 1972; Garfield, 1975) y, algunos años más tarde, en 1978, el *Arts & Humanities Citation Index* (AHCI) (Garfield, 1977). Varias décadas después, estos registros se transformaron en grandes sistemas de información con decenas de millones de registros a los que actualmente se accede a través de la plataforma Web of Science (WoS), que nuclea diversas bases de datos multidisciplinares (ver [Apéndice](#)). En la década de 1990, ISI fue comprado por el grupo canadiense Thomson que, a su vez, en abril de 2008, adquirió la agencia de noticias Reuters, por lo que la firma pasó a ser conocida como Thomson Reuters. En 2016, las compañías Onex Corporation y Baring Private Equity Asia adquieren los productos de Thomson Reuters, por lo que Web of Science pasó a manos de estas empresas bajo el nombre corporativo de Clarivate Analytics. Esta adquisición por parte de fondos asiáticos ha generado acuerdos de colaboración con el National Center for Science and Technology Evaluation de China, para acrecentar la participación de ese país en las bases de datos que integran la plataforma Web of Sciences (Clarivate Analytics, 2017a).

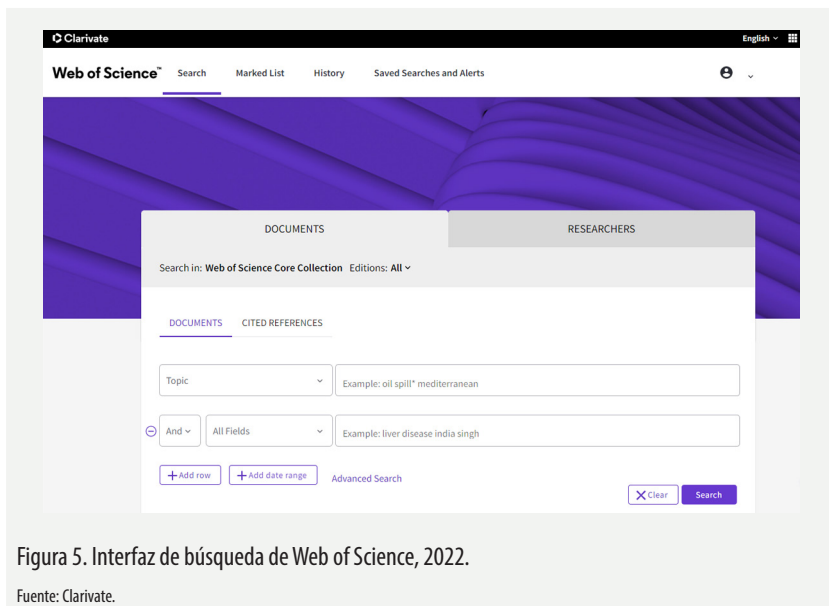


Figura 5. Interfaz de búsqueda de Web of Science, 2022.

Fuente: Clarivate.

Su finalidad no es mostrar el contenido a texto completo, sino documentar la existencia de las investigaciones incluidas en la base, por lo que en muchos casos solo se podrá obtener el título y el resumen del artículo, los nombres de los autores, las filiaciones institucionales, el nombre de la revista, los datos de publicación, la institución que financió la investigación, etc., pero no el texto completo del artículo. En 2004, luego del surgimiento de Scopus, WoS comenzó una fase de expansión: las grandes bases creadas por Garfield se reagruparon en *Web of Science Core Collection* y se incorporaron, además, Chinese Science Citation Database; Russian Science Citation Index; SciELO Citation Index, entre otros recursos (ver [Apéndice](#)). Todas estas nuevas bases pueden ser consultadas a través de la plataforma Web of Science, desde la misma interfaz (Figura 5).

El acceso a la plataforma WoS es pago, y los organismos estatales de ciencia y técnica de cada país son los encargados de firmar los contratos para dar acceso a las instituciones de investigación asociadas al programa. En el caso de Argentina, no forma parte de los recursos que la Biblioteca Electrónica de Ciencia y Tecnología, dependiente del Ministerio de Ciencia, Tecnología e Innovación (MINCYT) pone a disposición de las instituciones asociadas, por lo que las instituciones argentinas no tienen acceso a esta base de datos.

En cuanto a la cobertura temporal retrospectiva, por la diversidad de colecciones que hoy forman parte de WoS, los años varían. Como se puede ver en el [Apéndice](#), tanto en *Science Citation Index* como en *Social Sciences Citation Index* la cobertura llega hasta 1900, pero en *Arts & Humanities Citation Index* llega hasta 1975 y en las restantes es aún menor. Estas variaciones deberían tenerse en cuenta a la hora de definir el recorte temporal de la investigación.

Capítulo 5

¿Cómo buscar en una base de datos?

Interfaz de búsqueda

Toda base de datos cuenta con una interfaz que permite acceder a los registros existentes en esa base a través de un motor de búsqueda, que será el encargado de rastrear la solicitud realizada en la interfaz. Los elementos básicos de esa interfaz de búsqueda son:

- *Término de búsqueda:* es la expresión que se va a buscar, sea un concepto, un término, un descriptor, el nombre de un autor o una autora, el título de una revista, etc.
- *Campo de búsqueda:* indica en qué campo el motor debe ir a buscar el término de búsqueda elegido. La forma de nominar los campos varía según la base de datos utilizada.
- *Operador lógico:* permiten ampliar (*or*) o acotar (*and, not*) el número de artículos recuperados.

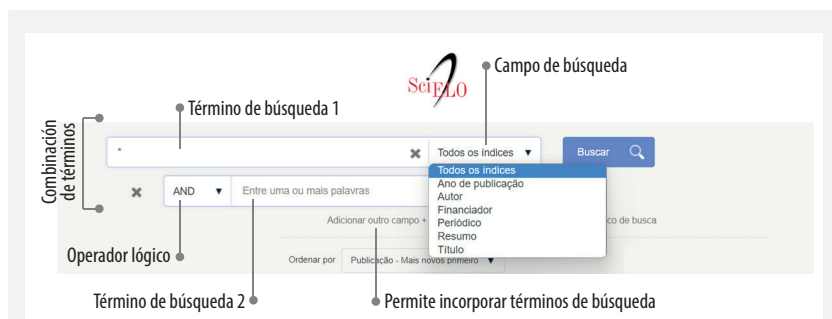


Figura 6. Interfaz de búsqueda avanzada de SciELO.

Fuente: Elaboración propia con base en SciELO.org.

La Figura 6 muestra los elementos de la interfaz de búsqueda avanzada de la biblioteca electrónica SciELO que, al igual que otras bases de datos, permite combinar dos o más términos de búsqueda. Esta condición es imprescindible para recortar la búsqueda. Por ejemplo, si quisiéramos recuperar trabajos sobre “desigualdades y acceso a los servicios de salud”, y la interfaz no nos brindara la posibilidad de combinar dos o más términos de búsqueda, probablemente utilizaríamos solo el término “desigualdades”. Al realizar en SciELO la búsqueda con el término “*inequalities*”, el motor de búsqueda recupera más de 5.000 registros, con una pertinencia temática muy reducida. Sin embargo, si combinamos los tres elementos que componen nuestro tema, utilizando el operador lógico “y” (“*inequalities*” AND “*access*” AND “*health services*”) los resultados no llegan a 400 registros con una pertinencia temática mucho mayor.

Idiomas de búsqueda

Si retomamos el ejemplo anterior, cabría preguntarse ¿por qué utilizar términos en inglés para buscar en una base de datos como SciELO, que cuenta con un alto porcentaje de revistas de América Latina? Esta pregunta encierra algunos presupuestos desacertados como, por ejemplo, que en las revistas de América Latina se publican solo trabajos de la comunidad científica latinoamericana o que esa comunidad publica sus trabajos mayoritariamente en español. Ambas afirmaciones son erróneas.

Desde hace siglos, una de las prácticas generalizadas en la comunidad científica es que cada equipo de investigación elige la revista en la que desea publicar. Más allá de que esa elección esté mediada por múltiples condicionamientos⁶, las revistas científicas de América Latina han sido históricamente —y continúan siéndolo en la actualidad— uno de los canales de circulación de la producción científica del mundo. Por estas revistas transitan una gran diversidad de lenguas, no solo porque reciben contribuciones de muchos otros países del mundo, sino porque la propia comunidad científica latinoamericana escribe y publica en diversas lenguas.

Por la propia estandarización idiomática que promovieron desde sus inicios los sistemas de información, y que se acentuó aún más en la era digital, las revistas deben incluir una versión del título y del resumen en inglés, más allá de la lengua en la que esté publicado cada artículo. De este modo, hacer

⁶Nos referiremos a los condicionamientos generados por los sistemas de evaluación de la producción científica escrita, los intereses corporativos, las dimensiones simbólicas, etc. (Salud Colectiva, 2011; Martinovich *et al.*, 2014; Martinovich, 2015, 2016, 2020, 2021).

las búsquedas en inglés permite recuperar trabajos publicados, por ejemplo, en francés, portugués, español, etc.

En el caso de SciELO, si bien existen ciertas diferencias en el número de registros recuperados al hacer las búsquedas en inglés, esas diferencias varían según los términos utilizados. Por ejemplo, al buscar en la interfaz con el término “desigualdades”, el motor de búsqueda recupera 4.912 registros mientras que, con “*inequalities*”, recupera 5.065 registros. Sin embargo, con otros términos, las diferencias son mucho mayores: por ejemplo, con el término “depresión”, el motor de búsqueda recupera 4.329 registros mientras que, con “*depression*”, recupera 8.059 registros. Esto nos muestra que detrás de una herramienta de búsqueda hay una gran diversidad de usos del lenguaje y prácticas de publicación que deberían ser consideradas al momento de analizar los registros obtenidos.

Si realizamos la misma prueba en PubMed, las diferencias son mayores: con el término “*inequalities*”, recupera 26.637 registros mientras que, con “desigualdades”, solo 91 registros. ¿Esto significa que en PubMed hay menos producción de América Latina? Si retomamos los párrafos anteriores, nos daríamos cuenta que la pregunta está mal formulada, dado que el idioma en el que están los registros en una base de datos no se relaciona necesariamente con el idioma de publicación de los artículos. Incluso, el idioma de publicación no tiene necesariamente que ver con la región de la que provenga el artículo. Un dato muy importante para poder comprender los resultados obtenidos en las búsquedas es que uno de los requisitos para que una revista científica ingrese a PubMed —más allá del país de publicación de la revista y de la lengua en la que esté publicado cada artículo— es que todas las revistas deben enviar a PubMed una versión del título y del resumen en inglés. Por lo tanto, de los datos que incluimos al inicio de este párrafo, solo podríamos inferir que PubMed comenzó a aceptar resúmenes en español y que, por el momento, al realizar búsquedas con el término “desigualdades” en el campo *Title/abstract* (título y resumen) el motor de búsqueda recupera 91 registros.

Términos de búsqueda: descriptores y términos libres

Para poder organizar millones de registros al interior de una base de datos, ciertas bibliotecas electrónicas desarrollan listados estructurados de términos para que las personas que busquen trabajos sobre un tema específico puedan recuperarlos. Por lo tanto, en las bases de datos que operan con descriptores, lo primero que deberíamos hacer es identificar qué términos vinculados a nuestro problema de investigación se encuentran en el listado de

descriptores. Estos términos se organizan en colecciones estructuradas denominadas *tesauros*.

Tesaurus de la UNESCO

Creado para recuperar documentos en la biblioteca digital de la UNESCO. Áreas: educación, cultura, ciencias naturales, ciencias sociales y humanas, comunicación e información. Idiomas: árabe, inglés, francés, ruso y español. Acceso a la biblioteca digital UNESDOC: <https://tinyurl.com/yur5rshh>
Acceso al Tesaurus de la UNESCO: <https://tinyurl.com/3byuxna3>

Tesaurus de la OIT

Compilación de términos relacionados con el mundo del trabajo, creado por la Organización Internacional del Trabajo (OIT), para recuperar documentos en la plataforma “Labordoc”, repositorio digital de la OIT. Áreas: trabajo. Idiomas: inglés, francés y español. Acceso a Labordoc: <https://tinyurl.com/muupdre5>
Acceso al Tesaurus de la OIT: <https://metadata.ilo.org/thesaurus.html>

Descriptores de Ciencias de la Salud (DeCS)

Creados para recuperar documentos en la Biblioteca Virtual en Salud (BVS). Áreas: ciencias de la salud. Idiomas: español, francés, inglés y portugués. Acceso a la BVS: <https://pesquisa.bvsalud.org/portal/advanced>
Acceso a los DeCS: <https://decs.bvsalud.org>

Medical Subject Headings (MeSH)

Creados para recuperar documentos en PubMed. Áreas: biomedicina y ciencias de la salud. Idiomas: inglés. Acceso a PubMed: <https://pubmed.ncbi.nlm.nih.gov/advanced>
Acceso a los MeSH: <https://www.ncbi.nlm.nih.gov/mesh>

Eric Thesaurus

El Tesauro Europeo de la Educación fue creado para recuperar documentos en la base de datos Education Resources Information Center (ERIC). Áreas: educación. Idioma: inglés.

Acceso a ERIC: <https://eric.ed.gov>

Acceso a ERIC Thesaurus: <https://eric.ed.gov/?ti=all>

Es importante aclarar, que estos términos son utilizados para clasificar documentos al interior de las bases de datos para las cuales fueron creados, pero no serán de utilidad para los actuales motores de búsqueda que rastrean términos en la totalidad de texto. Si decidimos realizar búsquedas bibliográficas con descriptores deberemos tener en cuenta que:

- Son efectivos cuando se utilizan términos unívocos, por ejemplo, nombres de enfermedades, materiales, técnicas y métodos tipificados internacionalmente, y que la propia comunidad científica acordó como denominación común.
- No deberían utilizarse para prácticas, acciones o sentidos sociales, que no responden a un lenguaje unívoco (ver el capítulo ¿Cómo crear un mapa de términos de búsqueda?)
- Para que las bases de datos vinculen un documento a un determinado descriptor, es necesario que previamente el documento haya sido identificado con un descriptor. De lo contrario, si la base de datos no cuenta con personal para asignar descriptores y quienes editan los documentos no reemplazan las expresiones libres que autoras y autores colocan como “palabras claves”, ese documento no podrá ser recuperado en las búsquedas realizadas con descriptores.

Campos de búsqueda

La elección de los términos es tan relevante como la elección de los campos de búsqueda. Imaginemos que estamos reconstruyendo la bibliografía publicada sobre “formulación de políticas” y al buscar cuál es el *Mesh Term* que existe para poder comenzar a rastrear algunos trabajos, recuperamos el descriptor *policy making*. Pero al momento de hacer la búsqueda en PubMed, en vez de seleccionar el campo “*Mesh Term*”, no seleccionamos ningún campo

en particular y dejamos el campo que viene por defecto que es “*All fields*”, por lo que el motor de búsqueda va a recuperar todos los trabajos que en alguna parte del registro tengan los dos elementos juntos (*policy making*) o cada uno por separado⁷, por lo que va a recuperar artículos que contengan algunas de esas palabras, incluso, en el nombre de la revista en el que fueron publicados (por ejemplo, va a traernos todos los artículos publicados en la revista *Health Policy* aunque no aborden la noción de “formulación de políticas”); también va a recuperar artículos que en el título o en el resumen tengan algunas de esas palabras, por lo que puede recuperar estudios sobre cocinar o preparar comidas (*making meals*) o tomar decisiones (*decision making*) que poco tienen que ver con el tema que queremos reconstruir.

En algunas bases de datos, la interfaz nos muestra muchos campos de búsqueda. Por ejemplo, PubMed muestra más de 40 campos de búsqueda (National Center for Biotechnology, 2005), mientras SciELO muestra solo siete: todos los índices, año de publicación, autor, financiador, periódico, resumen y título. ¿Eso significa que la base de datos de SciELO tiene solo siete campos? No, de hecho, tiene más, pero quizás en solo siete campos puede garantizar que la información esté estandarizada en todos los registros. ¿Qué significa esto?

Para que una búsqueda sea efectiva se requieren dos niveles de normalización: por un lado, la base de datos debe definir campos de búsqueda y, por otro, los documentos que ingresan a esa base de datos deben estar procesados para que esos campos puedan ser leídos por máquinas de manera que el motor de búsqueda puede recuperarlos. En el caso de las bases de datos basadas en descriptores, por un lado, debe existir un listado controlado y jerarquizado y, por otro, ese listado controlado y jerarquizado debe aplicarse a los documentos que ingresan a la base de datos para clasificarlos, de manera de que el motor de búsqueda puede identificarlos.

Para graficar este problema, imaginemos el siguiente ejemplo: durante años acumulamos fotocopias de libros y artículos y decidimos organizarlos en cajas, para lo cual elaboramos un listado con los temas que permitan agrupar nuestras fotocopias. Pero no alcanza con tener el listado, sino que es necesario rotular las cajas y, lo más importante, poner los documentos en las cajas correspondientes porque, de lo contrario, el listado quedará desvinculado de las fotocopias.

⁷La expresión de búsqueda que genera el motor al seleccionar la opción “*All fields*” es la siguiente: “*policy making*”[MeSH Terms] OR (“*policy*”[All Fields] AND “*making*”[All Fields]) OR “*policy making*”[All Fields]. Esto muestra que busca los términos juntos y por separado en todos los campos.

Dentro de una base de datos, este tipo de estructura requiere, por un lado, de un gran número de personas capaces de analizar los documentos para asignarles los descriptores y, por otro, criterios consensuados y estandarizados para asignar esos descriptores a los documentos. En el caso de PubMed, durante décadas contó con una gran cantidad de personas dedicadas a asignar MeSH a todo el material que ingresaba, y para ello utilizaban procedimientos estandarizados (U.S. National Library of Medicine, 2016b). Pero en 2021, anunciaron la transición hacia una indización automatizada a través del algoritmo MTIA que está en un proceso de continuo aprendizaje para mejorar la aplicación de términos MeSH en forma automatizada. El trabajo humano está centrado en garantizar que los términos MeSH asignados por el algoritmo MTIA sean correctos.

Sin embargo, si la base de datos, por su propia estructura descentralizada no cuenta con personal dedicado a rotular los documentos que ingresan y no cuentan con algoritmos capaces de asignarlos, el listado de términos controlado y jerarquizado queda desvinculado de los documentos. Por ejemplo, cuando el autor elige qué “palabras claves” describirán el contenido de su artículo, muy pocos autores asocian la elección de esos términos con la necesidad de utilizar un lenguaje estructurado (descriptores) que permita una correcta indización y posterior localización de su trabajo, y muchas revistas no lo exigen. Los descriptores no dan cuenta de la *lógica descriptiva del autor*, sino de la *lógica clasificatoria de cada base de datos*.

Por otra parte, como ya hemos mencionado, los índices de términos controlados tienden a quedar rezagados por lo que el descriptor tiende a quedarse obsoleto frente a los cambios que se producen en el lenguaje (Clarivate Analytics, 2017b). Desde esta perspectiva, es más recomendable reconstruir un mapa de términos posibles y buscarlos como términos libres en el campo título y resumen.

Filtros

Una vez realizada una búsqueda, los filtros permiten recortar el total de registros recuperados sobre la base de algún parámetro. Si tomamos como ejemplo el tema de “mujeres en prisión”, imaginemos que deseáramos abordar cuál es la perspectiva de la comunidad científica sobre el tema, pero en el período 1980-2015. En ese caso, se aplica el filtro “año” o “*publication dates*”, según la base de datos, y filtra los resultados según el año de publicación del artículo, independientemente del alcance temporal de la investigación desarrollada en el artículo. Es decir, entre 1980 y 2015 pueden haberse publicado

estudios que aborden el período 1920-1940, que quedarán dentro del recorte porque el filtro no lee el contenido, sino que solo toma el año en el que ese trabajo fue publicado, y esta distinción entre año de publicación y período abordado en los estudios recuperados es importante definirla.

En todos los casos, la elección de los filtros debe responder a los propios objetivos de la investigación. Por ejemplo, el filtro “tipo de estudio” o “*article types*” puede ser muy útil para seleccionar ensayos clínicos, estudios de cohorte, o incluso retractaciones (*retraction of publication*) o artículos históricos (*historical articles*). Sin embargo, un filtro como “*text availability*”, que recorta o excluye en función de la disponibilidad del resumen, del texto completo, o del texto completo abierto, solo podría utilizarse en el caso de que tenga una estrecha relación con el tema a investigar, por ejemplo, si se analiza la producción en acceso abierto presente en determinadas bases de datos porque, de lo contrario, produciría un recorte arbitrario. Lo mismo sucede con otros filtros como “*languages*”, que recortan según el idioma de publicación del texto. En el caso de SciELO, los países que figuran en el filtro “colección” corresponden a las diversas colecciones de revistas por país, por lo que, al seleccionar, por ejemplo, Argentina no estaremos seleccionando estudios *sobre* Argentina o realizados *en* Argentina, sino simplemente revistas científicas editadas en Argentina y presentes en la colección SciELO Argentina, que pueden contener artículos provenientes de diversos países y que aborden problemas situados en otras regiones del mundo.

Los filtros suelen variar de una base de datos a otra, por lo que es recomendable recurrir a la información de soporte sobre los filtros que provee la propia base de datos.

Recorte temporal

Como se puede ver en el [Apéndice](#), el alcance temporal retrospectivo de los registros de las distintas bases de datos es muy variado, por lo que es necesario tenerlos en cuenta a la hora de definir un período de estudio que logre una representatividad temporal adecuada. Por ejemplo, imaginemos que alguien se propone analizar la evolución histórica de los estudios sobre tuberculosis y para ello propone un recorte temporal de cien años, tomando el período 1900-2000 y la siguiente selección de bases de datos: Scopus, PubMed y SciELO. Esa decisión va a generar una inconsistencia en los resultados, dado que Scopus cuenta con registros retrospectivos solo hasta 1970, PubMed hasta 1946 y, en el caso de SciELO, depende del año de inicio de cada revista y de la digitalización retrospectiva que cada revista haya logrado hacer. Por lo tanto,

los resultados obtenidos podrían ser leído erróneamente y concluir que el inicio de las investigaciones sobre el tema se produjo en la década de 1940, cuando en realidad existen estudios muy anteriores sobre tuberculosis.

Por otra parte, la decisión de utilizar distintas bases de datos para lograr cierta representatividad geográfica puede generar cierta inconsistencia temporal si se proponen períodos muy grandes de tiempo. Siguiendo con el ejemplo anterior, en 1970 se visualizaría un aumento importante en el número de contribuciones, lo cual podría ser leído como el surgimiento de un interés renovado por la tuberculosis, cuando en realidad se debe exclusivamente a la capacidad de digitalización retrospectiva de Scopus. Por lo tanto, es muy importante tener en cuenta el alcance retrospectivo de cada base de datos a la hora de plantear el recorte temporal de la investigación.

Recorte geográfico

En la etapa exploratoria es importante analizar cuál es el recorte geográfico más apropiado para cada tema particular. En muchas ocasiones, se plantea como recorte “América Latina” sin tener en cuenta el alcance real del término. ¿Cuántos países imaginamos cuando mencionamos América Latina?, ¿diez?, ¿veinte? ¿Cuántos idiomas de publicación incluiremos? ¿Podremos recuperar producción escrita de todos los países que integran la región sobre el tema que estamos intentando abordar? Estas son algunas de las preguntas que deberíamos formularnos.

Otro tema a tener en cuenta es qué entendemos por recorte geográfico: ¿el lugar de producción de los textos? ¿el ámbito de circulación de los textos? ¿la temática abordada en las investigaciones? Una de las características de la comunidad científica es que el país de quien investiga queda asociado a la institución en la que trabaja, más allá de su propia nacionalidad. Por ejemplo, si una autora brasileña se va a trabajar a una institución española, las investigaciones que realice dentro de ese marco institucional no serán consideradas dentro de la producción de América Latina, ni la producción brasileña, porque su nueva filiación institucional se inscribe en España. Sin embargo, si consideramos el recorte geográfico asociado a la temática del propio texto, es decir, investigaciones que analicen aspectos, por ejemplo, de Brasil, más allá de la nacionalidad de las autoras y los autores, podremos incluir trabajos realizados por equipos de investigación de cualquier parte del mundo, que publiquen sus trabajos en idiomas muy diversos y en revistas editadas en cualquier región del mundo, siempre y cuando el recorte geográfico de

la propia investigación coincide con nuestra elección. Esta distinción es de suma importancia para utilizar los filtros de forma correcta.

Extracción de registros

Mientras transitamos la etapa exploratoria es recomendable no exportar los resultados de las búsquedas, porque nos llenaremos de registros que probablemente no volveremos a utilizar. Pero una vez que finalicemos la etapa exploratoria y comencemos a realizar las búsquedas definitivas, es importante definir de qué modo se guardarán los artículos recuperados.

Ciertas bases de datos ofrecen la opción de exportar los registros obtenidos en las búsquedas en valores separados por coma o archivos CVS (*comma-separated values*) que pueden ser fácilmente interpretadas por Excel, a través de la función que transforma “texto en columnas”. El límite de registros a exportar depende de cada base de datos (por ejemplo, PubMed, nos permite exportar un máximo de 10.000 registros, mientras SciELO, un máximo de 2.000).

Una vez exportados los registros de las distintas bases de datos, podremos agruparlos en un solo archivo, lo que permite eliminar duplicaciones, y conformar un primer cuerpo documental al cual aplicar, a partir de una lectura inicial de todos los títulos y resúmenes, los criterios de inclusión y exclusión creados en la etapa exploratoria.

Primeros pasos: ¿cómo optimizar las búsquedas exploratorias iniciales?

Tal como menciona Maria Cecília de Souza Minayo “la fase exploratoria de una investigación es tan importante que en sí puede ser considerada una investigación exploratoria” (Minayo, 2013, p. 141). En esta etapa debemos delimitar el tema, optar por el marco teórico-conceptual, definir el objeto de estudio, explicitar un abordaje metodológico y las técnicas de análisis, entre muchas otras cosas. Pero la realidad es que, tal vez, al iniciar este proceso, ni siquiera tengamos en claro qué problema abordaremos sobre el tema elegido. Entonces, ¿cómo adentrarnos en algo que aún desconocemos? Para eso proponemos una serie de pasos que pueden servir como disparador de un proceso recursivo, que vaya abriendo nuevos caminos y permita adentrarnos en las distintas miradas precedentes sobre el problema que estamos intentando abordar.

- 1) En primer lugar, elegir un tema inicial, aunque aún no tengamos en claro qué problema abordaremos sobre ese tema.
- 2) Luego deberíamos seleccionar una base de datos que, por el recorte geográfico, la accesibilidad, o el recorte temático, se adapte mejor a nuestro tema de investigación.
- 3) Partiendo de un término inicial, realizar una búsqueda exploratoria inicial. El término a utilizar puede ser una condición (como la marginalización), o la población afectada por esa condición (por ejemplo, “jóvenes en situación de calle”). También puede ser una institución, como el Estado, y cruzar ese término con una práctica propia del Estado (por ejemplo, “legislación laboral”).
- 4) En esos resultados iniciales, deberíamos identificar los diversos términos (en español y en inglés) que la comunidad científica utiliza para nombrar el tema que nos interesa. En los casos que se aborden temáticas relacionadas con procesos sociales, no deberíamos preocuparnos por identificar “descriptores” —Descriptores en Ciencias de la Salud (DeCS) o Medical Subject Headings (MeSH)—, sino los propios términos que aparecen en los títulos y resúmenes de los artículos que vayamos recuperando. En este sentido, es importante evitar imponer ciertas denominaciones que se presuponen como válidas sin cotejar si el espectro terminológico utilizado es más amplio o distinto al imaginado.
- 5) Realizar nuevas búsquedas con los “descriptores” o los “términos libres” recuperados en la búsqueda exploratoria inicial.
- 6) Dentro de los artículos recuperados, deberíamos comenzar a identificar los criterios por los cuales consideramos que ciertos artículos se aproximan al problema que quisiéramos abordar y los criterios por los cuales un artículo se aleja de nuestro problema, y lograr describir y categorizar esos criterios lo más detalladamente posible. Este procedimiento nos va a permitir cincelar el tema y el problema, ir decidiendo, delimitando y, a la vez, profundizando los límites de nuestro objeto de estudio.
- 7) En este proceso inicial, deberíamos poder recuperar al menos cuatro artículos sobre el problema que nos interesaría abordar, que se aproximan al abordaje que nos interesa.

8) Leer los artículos e identificar:

- el problema que abordan sobre ese tema;
- cómo justifican social y académicamente el problema y qué fuentes utilizan;
- si explicitan algún marco conceptual desde el cual abordar el problema;
- qué metodología implementaron;
- cuál es el objeto de estudio y cómo lo delimitaron geográfica, temporal y conceptualmente;
- del listado de referencias bibliográficas, recuperar algunas fuentes que puedan ser de interés para alguno de los grandes apartados de la investigación (justificación del problema, marco teórico-conceptual, metodología, etc.)

Estos pasos deberían reiterarse hasta lograr moldear el tema, el problema, los conceptos teóricos desde los cuales se van a conformar categorías y dimensiones para analizar un objeto de estudio delimitado. En este proceso, es importante poder interrogar a esos textos, no solo desde un plano descriptivo que nos permita ver qué hicieron y qué decisiones tomaron, sino también pensar por qué tomaron esas decisiones, por qué optaron por esa metodología para abordar ese tema; si no explicitan un marco conceptual, por qué no lo hacen, cuál es el contexto histórico, político, económico de esa producción, qué actores sociales están involucrados, qué tipo de tensiones están presentes... Cuantas más preguntas logremos formularle a esos textos, más enriqueceremos nuestra mirada sobre el tema, lo que nos va a permitir que nuestro problema se asiente en un territorio más firme, pero también más complejo.

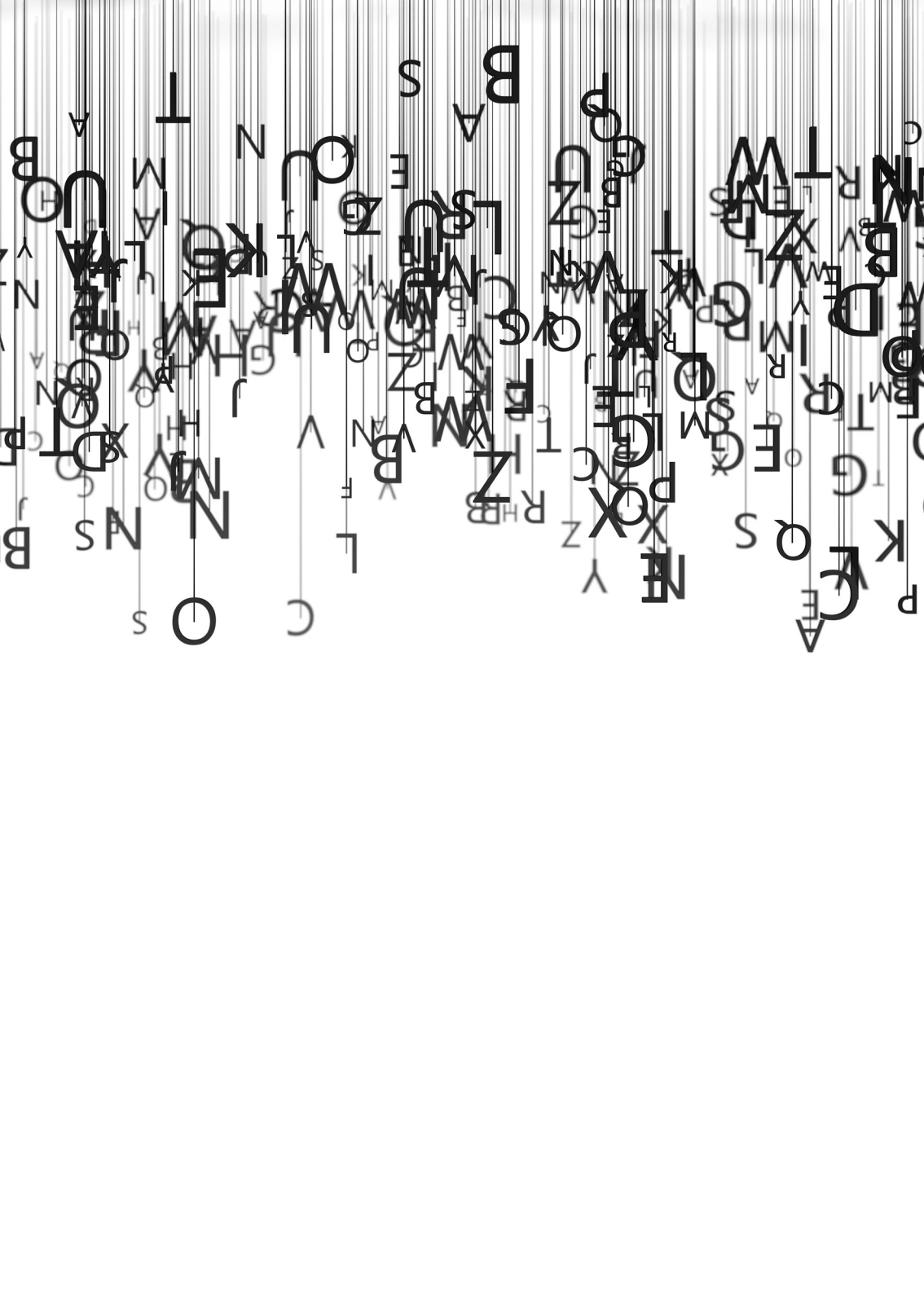
Si bien la etapa exploratoria no requiere de un diseño de búsqueda, dado que estamos “explorando” los recursos, comprendiendo la lógica de funcionamiento de los motores de búsqueda, el tipo de registro que contiene cada base de datos, etc., esta etapa será un insumo clave para la elaboración posterior del diseño. Por lo tanto, es importante conservar información acerca de dónde, cómo y cuándo se realizaron las búsquedas iniciales y qué resultados se obtuvieron.

Registrar cada una de estas búsquedas exploratorias permite analizar cuáles son las mejores estrategias a seguir y, sobre todo, comenzar a delinear los criterios de inclusión y exclusión que aplicaremos al universo de registros encontrados. Estos criterios funcionarán posteriormente como un filtro analítico muy potente, por lo tanto, esta etapa resultará muy útil para clarificar no solo esos criterios y los distintos procedimientos de búsqueda, sino para recortar y redefinir el proyecto de investigación.

Las búsquedas exploratorias, al permitirnos el acceso a una multiplicidad de perspectivas sobre un tema, suelen tentarnos a desviar nuestra atención y a querer modificar constantemente nuestro recorte. Si bien esto resulta muy enriquecedor en cuanto a la apertura de nuestro universo referencial, debemos realizar un esfuerzo significativo para elegir un camino a seguir y sostener esa elección para poder avanzar y llegar a buen puerto en los plazos establecidos.

Un recurso para estos casos es armar una carpeta de “futuras investigaciones” donde podamos guardar todo aquel material que nos pareció sumamente interesante, pero que no necesariamente se vincula con el recorte o el abordaje particular que elegimos en esta investigación.

Según Magariños de Morentin (1991), la cultura, como sustituto de la experiencia, le proporciona a cada persona una multitud de sistemas en los que quedan significativamente interpretados los interrogantes que pueda encontrar en su entorno; es decir, cada persona no tiene frente a sí un mundo, sino un espejo de sus propios sistemas semióticos de identificación, de modo que no seleccionamos lo que existe en el mundo sino aquello que percibimos como conocido. Este proceso, al que el autor denomina “clausura simbólica”, es el que debemos intentar desandar, a través de una ruptura que nos permita formularnos preguntas sobre aquello que hasta el momento dábamos por sentado o nos pasaba desapercibido.



Capítulo 6

¿Cómo crear un mapa de términos de búsqueda?

Una vez que atravesamos esa instancia exploratoria inicial que nos permitió tomar las definiciones iniciales respecto del problema a ser abordado, si decidimos hacer una investigación que tome como objeto de estudio la producción escrita de la comunidad científica, debemos elaborar alguna estrategia para componer un “mapa de términos”.

En diversas áreas de conocimiento existen términos que han sido consensuados como lenguaje unívoco para denominar algunos eventos determinados y que pasan a formar parte de tesauros, es decir, índices de términos controlados y organizados jerárquicamente, que sirven como lenguaje unívoco para clasificar los artículos al interior de una base de datos, y para la búsqueda y recuperación de artículos al interior de esas mismas bases de datos. Hay una gran variedad de índices de términos controlados como los Medical Subject Headings, los Descriptores en Ciencias de la Salud, el Tesauro de la OIT, el Tesauro de la UNESCO, etc., en los que se pueden encontrar términos que nominan procesos, procedimientos, microorganismos, componentes celulares, elementos químicos, etc. Por lo tanto, si se quisieran recuperar investigaciones vinculadas a algunos de estos términos, en teoría, se debería utilizar el término controlado o “descriptor” que nomina el tema elegido.

Sin embargo, incluso en las áreas donde se suele utilizar un lenguaje unívoco, como es la medicina, la dinámica del uso social del lenguaje va a una velocidad mucho mayor que la inclusión de los nuevos términos en los índices, lo que demanda realizar búsquedas exploratorias en las bases de datos solo para recuperar el término con que la comunidad de pares nomina una condición específica. En una de las tantas clases de búsqueda bibliográfica, una estudiante estaba interesada en recuperar estudios sobre los nacimientos prematuros, pero no todos los nacimientos prematuros, sino los prematuros tardíos. Siguiendo la lógica de los descriptores, buscó en el listado de MeSH Terms y localizó el descriptor que más se acercaba: “*premature birth*”. Pero al realizar las búsquedas en PubMed con el descriptor, no lograba encontrar los

artículos que estaba buscando. No había un descriptor que indicara específicamente el tipo de nacimiento prematuro que ella estaba buscando, y si bien por su propia práctica laboral sabía que existía bibliografía sobre el tema, no estaba pudiendo acceder a esos artículos. Le propuse que se olvidara de los descriptores y comenzara a explorar la base de datos utilizando términos libres, hasta que luego de varios intentos probó con la expresión “*late-preterm*” (con guion) que aparecía en uno de los artículos y la base de datos le mostró ese recorte que tanto había buscado y al cual no lograba acceder.

Uno de los muchos ejemplos de disociación entre los descriptores y los usos terminológicos por parte de la comunidad científica es el término “vulnerabilidad social”. Mientras que en el listado de Descriptores de Ciencias de la Salud está definido como “Grado de susceptibilidad o de riesgo a que está expuesta una población de sufrir daño ante un desastre natural”, si hacemos una búsqueda con ese descriptor al interior de la Biblioteca Virtual en Salud, el resultado nos muestra muchos otros usos del término “vulnerabilidad social” que no están vinculados a los desastres naturales. Esto muestra que el lenguaje es una expresión viva, incluso en aquellas áreas que suelen utilizar nomenclaturas unívocas.

Las diferentes denominaciones no pueden ser consideradas como “sinónimos” porque nominan cosas muy distintas. En los casos de eventos y prácticas sociales, los distintos términos responden a diferencias conceptuales, políticas, culturales, etc. Según el lingüista John Lyons (1984), no existen palabras con el mismo significado en el sistema de la lengua, dado que para ser sinónimos deberían compartir el mismo significado descriptivo, expresivo y social (sinonimia completa) y en todos los contextos de enunciación (sinonimia absoluta). En el campo científico, quizá sea el ámbito en el que esta definición tenga mayor relevancia. El término traumatismo craneoencefálico es un evento definido que cumple una serie de condiciones, por lo que no admitiría ser sustituido por otro que guarde cierta relación sinonímica. Lo mismo ocurre con los conceptos desarrollados por distintas corrientes de pensamiento: no son sustituibles, se puede ampliar su definición, pero no son intercambiables. Por ejemplo, el concepto de “campo” de Pierre Bourdieu podrá describirse utilizando recursos de sinonimia (terreno de disputas, espacio de lucha, etc.) pero el concepto en sí no es sustituible.

El uso indistinto y variado en un mismo texto de diversos términos para designar un mismo concepto, evento, entidad o grupo (por ejemplo, adulto mayor, anciano, tercera edad), sin definir conceptualmente qué se entiende por cada uno de ellos, se considera una inconsistencia o falta de claridad terminológica. Sin embargo, esa variedad terminológica que responde a diversas denominaciones creadas por distintos grupos sociales, propuestas

conceptuales, etc., está presente en las distintas investigaciones existentes sobre el tema, y para intentar recuperar esa riqueza de perspectivas es necesario reconstruir los términos presentes en los diversos textos. A estas expresiones las denominaremos *términos libres*, en contraposición a los *descriptores* o *términos controlados* que forman parte de los listados de MeSH, DeCS, etc.

En áreas donde la terminología cambia muy rápidamente, o que investigan fenómenos muy recientes, las búsquedas de bibliografía a través de descriptores o términos controlados solo permiten reconstruir eventos, entidades o procesos desde una perspectiva anterior, dado que los índices de términos controlados se mueven muy lentamente, tienden a quedarse rezagados respecto de la expansión de los campos y del desarrollo de un nuevo lenguaje, por lo que el descriptor tiende a quedarse obsoleto frente al nuevo lenguaje (Clarivate Analytics, 2017b).

Ejemplos de “mapas de términos”

En el trabajo titulado “*Climate change research in view of bibliometrics*”, en la primera parte del apartado metodológico, los autores realizan una interesante descripción acerca de la complejidad de recuperar artículos sobre el amplio campo de investigación del “cambio climático”, dado que es un tema abordado desde diversas disciplinas y que no cuenta con una clasificación exhaustiva a través de palabras claves, descriptores o tesauros (Haunschild et al., 2016). Como método de selección de los términos de búsqueda recuperan la propuesta de cuatro pasos de Wang *et al.* (2014), en la que realizan una búsqueda inicial que da origen a un proceso de reformulación de consultas, basado en una variante encontrada en la búsqueda anterior, es decir, realizan una primera búsqueda con “cambio climático” (*climate change*) con el propósito de recuperar nuevos términos como:

climat chang* (climate, climatic change, changes, changing)*

climat warming (climate, climatic)*

global temperature (temperatura, temperatures)*

global warming

greenhouse gas (gas, gases)*

greenhouse effect (effect, effects)*

greenhouse warming

Estos términos fueron utilizados para recuperar otras denominaciones propuestas por investigadores de distintas regiones y áreas de conocimiento para

aludir a aquello que los autores inicialmente consideraban como “cambio climático”. El asterisco que se ve en los ejemplos es un símbolo que se utiliza para truncar un término e indicarle al motor de búsqueda que recupere todos los artículos que en su título y resumen contengan alguna variación de, por ejemplo, el término *climat* chang**, por lo cual va a recuperar tanto *climate* como *climatic* en sus combinaciones con *change*, *changes* o *changing*.

En algunas bases de datos, por ejemplo en Web of Science, el asterisco también puede utilizarse al inicio de la palabra (Clarivate, 2021), por ejemplo: **fertility* permitiría recuperar tanto *fertility* como *infertility* y, si bien *fertil** incluye términos como *fertility*, *fertilization*, *fertilized*, *fertilizability*, al utilizar **fertil**, deberíamos poder recuperar todas las variaciones anteriores y también términos como *self-fertilization*. Sin embargo, el motor de búsqueda de PubMed, por ejemplo, no reconoce este truncamiento inicial y solo admite el truncamiento final (U.S. National Library of Medicine, 2006).

Pasos para crear un “mapa de términos”

Cada concepto, proceso, acción, evento, entidad, organismo, etc., tiene su propio espectro terminológico que se transforma al ritmo de los movimientos que se produzcan en cada área de estudio en particular, por lo que es necesario adaptar y justificar, en cada caso, las estrategias a seguir. A modo de ejemplo, mostraremos uno de los primeros pasos que debería realizarse en la *fase exploratoria* para reconstruir un *mapa de términos*. En este caso, proponemos el tema *mujeres en prisión*. Como ya mencionamos, existen bases de datos cuya estructura organizativa y relacional se centra en los “*subject headings*” (tema principal), mientras otras basan su estructura relacional en las citas, de manera de reconstruir redes de “conversaciones” (Clarivate Analytics, 2017b). Como en este caso estamos tratando de ejemplificar la diferencia entre *términos controlados* (descriptores) y *términos libres*, vamos a iniciar el proceso de mapeo de términos en la base de datos PubMed, organizada a partir de una estructura temática.

El primer paso fue buscar el descriptor (*Mesh Terms*) que diera cuenta del ámbito “prisión”, o de su condición, es decir, “encarcelada”, y recuperamos el término *prison* y *prisoner*. Un punto clave a tener en cuenta en esta instancia de reconstrucción del espectro terminológico, es que no deberíamos imponerle denominaciones al tema. Si bien en términos personales puedo considerar que no deberían utilizarse expresiones como “mujeres presas”, que transforman una condición transitoria en un adjetivo que califica a esa persona,

estamos intentando reconstruir las distintas perspectivas sobre el tema por lo que se debería recuperar la mayor variedad terminológica.

Como el término *women* no es un descriptor, se buscó como término libre en el campo *Title/Abstract*, y se utilizaron los términos *prison* y *prisoner* para realizar dos búsquedas diferentes:

Pubmed | Búsqueda 1 (ámbito)

Línea 1 = Campo de búsqueda: *MeSH Terms* | Término: *prison*

Operador lógico = AND

Línea 2 = Campo de búsqueda: *Title/Abstract* | Término: *women*

Número de registros: 829

Fecha: 07/08/2022

Pubmed | Búsqueda 2 (condición)

Línea 1 = Campo de búsqueda: *MeSH Terms* | Término: *prisoner*

Operador lógico = AND

Línea 2 = Campo de búsqueda: *Title/Abstract* | Término: *women*

Número de registros: 1.817

Fecha: 07/08/2022

Al leer los títulos de los artículos recuperados en estas primeras búsquedas en PubMed surgieron las diversas formas en que la comunidad científica denomina el ámbito “prisión” (*prison; prisons; jail; jails; penitentiary institutions*) y la condición de “mujer en prisión” (*incarcerated women; imprisoned women; female prison inmates; women prisoners; female prisoners; female detainees; women in prison; jailed women*). Si bien se podría haber realizado una sola búsqueda truncando el término *prison**, tendríamos una noción mucho más acotada de lo que sucede con cada uno de los términos que nos impediría ver, por ejemplo, que al buscar con el término *prison* no solo recupera muchos menos registros, sino que un gran porcentaje de los registros que recupera hacen mención a la condición de “encarcelada” más que al ámbito “prisión”. No olvidemos que estamos en la fase exploratoria, haciendo una reconstrucción inicial del tema, por lo que el interés no debería centrarse en hacer búsquedas definitivas ni extraer los resultados obtenidos, sino lograr un primer acercamiento al tema para reconstruir un mapa de términos que posteriormente podamos plasmar en el diseño de nuestro instrumento de búsqueda.

El segundo paso fue utilizar este primer mapa de términos para realizar búsquedas en bases de datos cuya estructura relacional no se centre en temas, como es el caso de SciELO, por ejemplo.

SciELO | Búsqueda 3 (término libre truncado)

Línea 1 = Campo de búsqueda: *resumen* | Término: *prison**

Operador lógico: AND

Línea 2 = Campo de búsqueda: *resumen* | Término: *women*

Esquema de búsqueda = (*resumen: prison**) AND (*resumen: women*)

Número de registros: 365

Fecha 07/08/2022

El campo de búsqueda equivalente a *title/abstract* en PubMed y al campo *título, resumen, asunto* de la BVS, no está disponible entre las pocas opciones que ofrece la interfaz de búsqueda de SciELO, por lo que adoptamos *resumen* como el campo que se asemeja más. El término truncado *prison** incluye tanto *prison* como *prisons*, *prisoner* y *prisoners*. Luego de realizar búsquedas con los restantes términos, tanto en SciELO como en las restantes bases de datos, a partir de la lectura de los títulos de los artículos, el mapa quedó conformado de la siguiente manera:

*prison** (*prison, prisons, prisoner, prisoners*)

*jail** (*jail, jails, jailed, jailing*)

*penitentiary** (*penitentiary, penitentiaries*)

*incarcerat** (*incarcerate, incarcerated, incarcerating*)

*imprison** (*imprison, imprisoned, imprisoning*)

*inmate** (*inmate, inmates*)

*detaine** (*detainee, detainees, detained*)

En todos los casos, los términos truncados se combinaron con los términos *women* y *female*. Si bien se encontraron otras denominaciones, ya quedaban incluidas dentro del mapa de términos seleccionado como, por ejemplo, la expresión “mujeres en situación de cárcel” (traducido como *women experimentation in jail situation*) que contiene el término *jail*, o “mujeres cumpliendo pena privativa de libertad” (traducido como *women are serving prison sentences*) que contiene el término *prison*.

Este proceso de reconstrucción de un mapa de términos a partir de la lectura de los títulos que van surgiendo en las distintas búsquedas es uno de los primeros pasos que deberían realizarse en la *fase exploratoria*. Por un lado, ayuda a ampliar la mirada sobre el tema de investigación elegido, y evitar imponer ciertas denominaciones que se presuponen como válidas y abarcativas sin cotejar si el espectro terminológico es más amplio que lo imaginado. Por otro lado, en esa lectura inicial de los títulos se pueden detectar diversos actores, espacios, perspectivas, elecciones metodológicas, etc., desde las cuales se aborda el tema en las diversas regiones del mundo.

Capítulo 7

Diseño de un instrumento de búsqueda bibliográfica

Luego de atravesar la fase exploratoria, es necesario comenzar a plasmar las decisiones adoptadas sobre el mapa de términos, la selección de bases de datos, los campos de búsqueda, el recorte conceptual, temporal y geográfico del tema, el tipo de estudios, los criterios de exclusión e inclusión de los documentos, etc. Todas esas decisiones forman parte del apartado metodológico, en el que se describen y justifican las elecciones realizadas sobre los “materiales” y las “técnicas”. La propuesta de explicitar el instrumento de búsqueda bibliográfica no se vincula, en este caso, a la idea de “replicabilidad”. No se trata solo de mostrar que puede ser reproducido y, por lo tanto, validado. El instrumento en este caso debe ser ajustado cuantas veces sea necesario de manera de que cada una de las elecciones sean pertinentes con relación a los objetivos propuestos. La descripción detallada del instrumento es lo que permite que “otros” (tutoras y tutores, integrantes del jurado, lectoras y lectores) evalúen la pertinencia, consistencia y representatividad del corpus de investigación sobre el que se realiza el análisis.

Por ejemplo, en el estudio *“Digital presence of Norwegian scholars on academic network sites—Where and who are they?”* (Mikki et al., 2015), basados en la idea de que las tecnologías emergentes aumentan la visibilidad de los investigadores y el intercambio de ideas, los autores comparan la presencia de los investigadores noruegos en sitios como ResearchGate, Academia.edu, Google Scholar Citations, ResearcherID y ORCID, con los datos extraídos del sistema noruego de investigación. En el apartado metodológico describen cómo obtuvieron los datos de las y los investigadores, qué datos extrajeron, cuál fue el número total de registros y qué periodo abarcaron. Luego detallan las características de las distintas fuentes de datos (redes sociales, plataformas, motores de búsqueda, etc.), describen la forma de extracción de los datos de cada una de las bases y qué datos se extrajeron de cada una de ellas. También explican las limitaciones y los problemas que tuvieron en la extracción de

datos para luego detallar los procesos de validación y limpieza de los datos para su posterior análisis.

La forma que adopte el instrumento de búsqueda, recuperación, extracción, validación y depuración de los datos depende de las particularidades que plantee cada problema de investigación. La riqueza descriptiva de las decisiones adoptadas, en muchos casos, es producto de la capacidad de explorar, observar y analizar críticamente ciertos escenarios que hasta el momento de la investigación resultaban totalmente ajenos.

Elementos de un instrumento de búsqueda bibliográfica

Si bien cada investigación tiene sus particularidades, en líneas generales, los diversos apartados metodológicos analizados proponen una estructura basada en los siguientes elementos:

- *Pregunta de investigación*: incluir la pregunta dentro del primer párrafo del apartado metodológico permite cotejar con mayor facilidad la coherencia vertical de las distintas elecciones adoptadas.
- *Recorte geográfico y temporal*: ambos recortes permiten situar histórica, política y socialmente el recorte temático propuesto, por lo que no solo es importante explicitar el ámbito geográfico y el alcance temporal propuesto sino justificar esas elecciones.
- *Tipos de documentos seleccionados y soportes*: definir y describir qué tipos de documentos van a conformar el corpus de investigación (por ejemplo, estudios de cohorte, etnografías, estudios epidemiológicos, estudios de caso, triangulación de métodos, etc.) y si, en función del recorte temporal, se incluirán artículos en soporte impreso o solo en soporte electrónico.
- *Selección de bases de datos de bibliografía científica*: enumerar las bases de datos seleccionadas (PubMed, Scopus, SciELO, BVS, etc.) y justificar su inclusión sobre la base del criterio de *representatividad*: el diseño del instrumento de búsqueda debe permitir recuperar el espectro más abarcativo y representativo del recorte propuesto en la pregunta de investigación, teniendo en cuenta que las bases de datos presentan distintos recortes temáticos, geográficos y conceptuales.

- *Proceso de selección de descriptores y términos libres*: describir el proceso de selección de los términos de búsqueda y de sus combinaciones, realizado en la fase exploratoria.
- *Cantidad de artículos recuperados*: explicitar las combinaciones de descriptores, términos libres, campos de búsqueda y filtros utilizados en cada base de datos en una tabla que incluye los resultados preliminares obtenidos en cada base de datos (Tabla 2).

Tabla 2. Cantidad de artículos recuperados, según bases de datos seleccionadas. Junio, 2018.

Bases de datos	Términos de búsqueda			Campos de búsqueda	Filtro 1 (Período)	Filtro 2 (Tipo de documento)	Cantidad de artículos
	Término 1	Operador lógico	Término 2				
PubMed	prison*	AND	Argentina	Title/Abstract	1970-2017	Journal article	19
	jail*	AND	Argentina	Title/Abstract	1970-2017	Journal article	1
	incarcerat*	AND	Argentina	Title/Abstract	1970-2017	Journal article	3
	imprison*	AND	Argentina	Title/Abstract	1970-2017	Journal article	11
BVS	prison*	AND	Argentina	Título, resumen, asunto	1970-2017	Artículo	62
	jail*	AND	Argentina	Título, resumen, asunto	1970-2017	Artículo	28
	incarcerat*	AND	Argentina	Título, resumen, asunto	1970-2017	Artículo	4
	imprison*	AND	Argentina	Título, resumen, asunto	1970-2017	Artículo	15
SciELO	prison*	AND	Argentina	-	1970-2017	Artículo	47
	jail*	AND	Argentina	-	1970-2017	Artículo	12
	incarcerat*	AND	Argentina	-	1970-2017	Artículo	2
	imprison*	AND	Argentina	-	1970-2017	Artículo	18
Scopus	prison*	AND	Argentina	Article title, Abstract, Keywords	1970-2017		112
	jail*	AND	Argentina	Article title, Abstract, Keywords	1970-2017		9
	incarcerat*	AND	Argentina	Article title, Abstract, Keywords	1970-2017		12
	imprison*	AND	Argentina	Article title, Abstract, Keywords	1970-2017		31
Total							386

Fuente: Elaboración propia a partir de datos de las bibliotecas electrónicas PubMed, BVS, SciELO y Scopus.

Nota: El asterisco (*) es un símbolo de truncamiento que se utiliza para recuperar todos los términos derivados de una raíz común (prison* recupera prison, prisons, prisoner, prisoners).

- *Criterios de inclusión y exclusión*: estos criterios operan como filtros que se aplican al momento de la lectura de los títulos y resúmenes del material recuperado inicialmente en las bases de datos, con el propósito de eliminar aquellos que no se ajusten al recorte geográfico y temporal propuesto, al tipo de documentos y soportes definidos previamente y, sobre todo, al recorte conceptual, el cual debería explicitarse con claridad para delimitar qué tipo de artículos serán incluidos y cuáles excluidos (desde un punto de vista conceptual o desde algún parámetro no definido previamente). Por ejemplo, la búsqueda realizada en la Tabla 2 sobre “salud en cárceles argentinas”, podría haber sido diseñada para identificar las diversas dimensiones desde las cuales se aborda la salud de las personas detenidas en cárceles argentinas. Sin embargo, si no se explicita desde qué perspectiva teórica se aborda el concepto de “salud” y el alcance del concepto, se incluirán o excluirán trabajos sin un criterio formal que interpele la selección. El diseño de una búsqueda amplia como la propuesta en la Tabla 2, permite recuperar trabajos como “*Cárcel de mujeres: Una mirada etnográfica sobre las relaciones afectivas en un establecimiento carcelario de mediana seguridad en Argentina*”, en cuyo resumen tampoco se menciona la palabra “salud”. Solo la definición conceptual de los criterios de inclusión y exclusión basados en el marco teórico-conceptual elegido permitirá definir si las “relaciones afectivas” forman parte de la concepción de salud o no y, por lo tanto, si el artículo será incluido o excluido.

- *Estrategias adicionales de búsqueda*: en ciertos casos en los que el alcance temporal de la investigación incluya períodos anteriores a 1990, en los que gran parte de la producción no se encuentra digitalizada y sea necesario recuperar la producción en soporte impreso, o en trabajos que se propongan reconstruir y analizar propuestas teóricas en las que el libro y el capítulo de libro prevalezcan por sobre el artículo, se pueden utilizar las referencias bibliográficas de los trabajos recuperados en las búsquedas como punto de partida para recuperar otros trabajos y reconstruir un corpus más ajustado a la propuesta. Estas decisiones suelen ser explicitadas y descriptas como “estrategias adicionales de búsqueda”.

- *Diagrama de selección de corpus de investigación*: el procedimiento de conformación del corpus a ser analizado, en el que se especifica el número de registros resultantes de la eliminación por duplicaciones y por la aplicación de criterios de inclusión y exclusión y las tablas asociadas, suele graficarse en un diagrama (Figura 7) y puede ir acompañado por

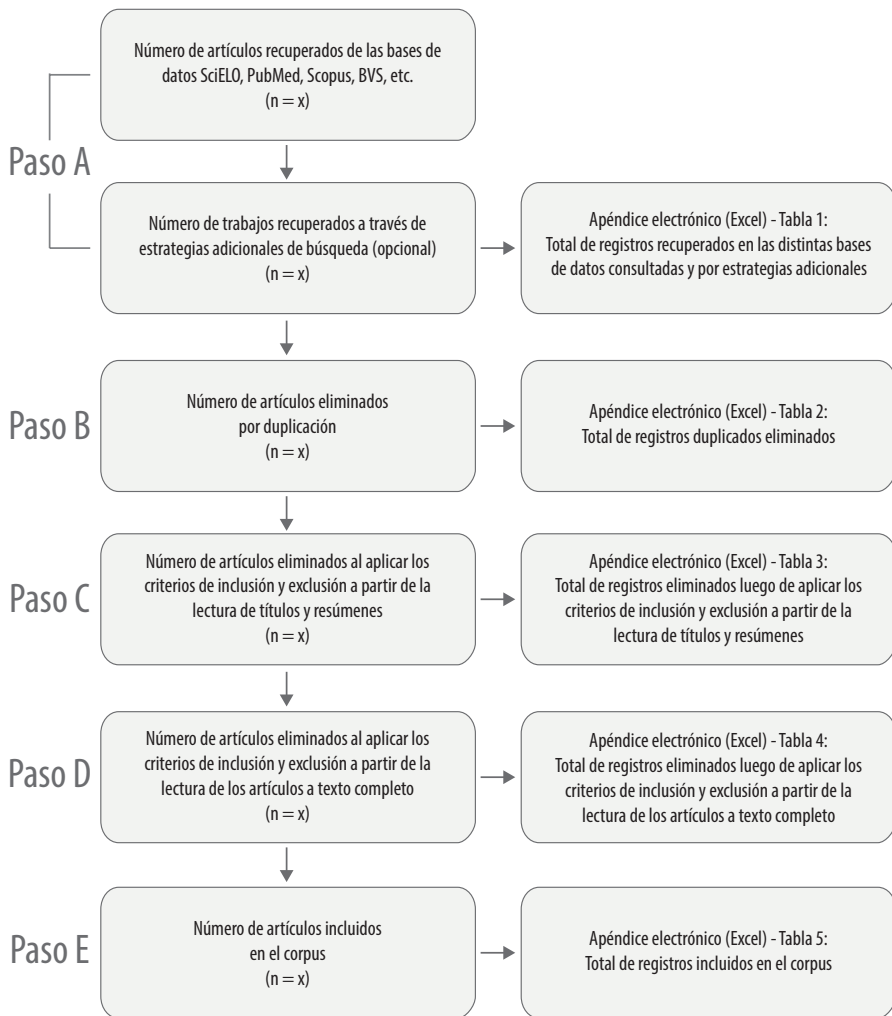


Figura 7. Ejemplo del diagrama del proceso de selección de las referencias bibliográficas que conforman el corpus de investigación.

Fuente: Elaboración propia.

las tablas con los registros consolidados. Las estrategias adicionales de búsqueda se sugieren, sobre todo, para los casos en que el alcance temporal de la investigación incluya períodos anteriores a 1990, o trabajos que se propongan reconstruir y analizar propuestas teóricas.

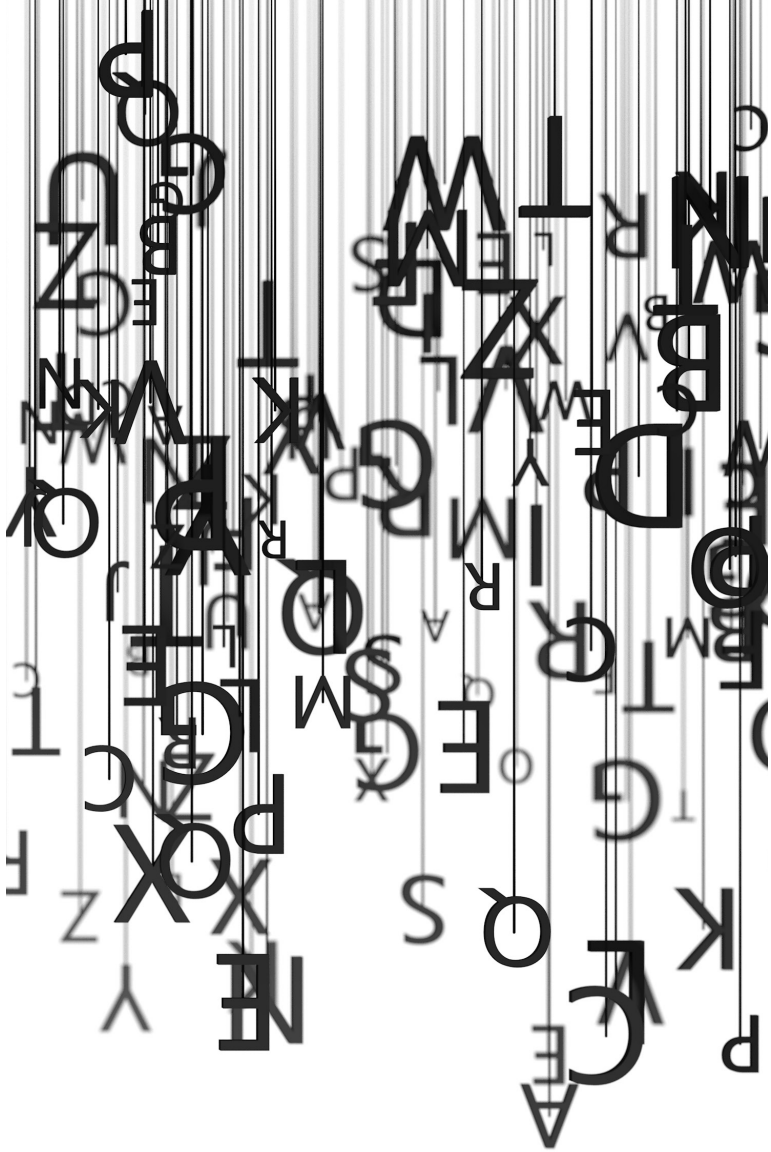
- *Categorías o dimensiones*: enumerar las categorías o dimensiones de análisis (definidas previamente en el marco conceptual) que se utilizarán para agrupar y analizar los textos.
- *Cuerpo documental*: el proyecto de investigación suele ir acompañado de un apéndice electrónico con las cuatro tablas enumeradas, como se muestra en la Figura 7 (Pasos A, B, C y D). Adicionalmente, dentro del proyecto, se suele incorporar la tabla con el total de registros incluidos en el *corpus* (Paso E). Las distintas tablas que muestran el proceso de selección de trabajos son de gran utilidad para transparentar el proceso de selección del material.

Todos los elementos que componen el instrumento están estrechamente vinculados entre sí y, a su vez, con la pregunta de investigación, los conceptos teóricos que orientaron la pregunta, los objetivos y el tipo de análisis que proponga el trabajo. Por lo tanto, cualquier modificación que se realice en alguno de los ítems repercute necesariamente, tanto en los otros componentes del instrumento, como en la coherencia vertical de todo el proyecto.

A modo de cierre

En este texto nos propusimos mostrar, por un lado, los diversos usos y tipos de fuentes que pueden formar parte de una investigación y, por otro, que lograr una amplia investigación bibliográfica requiere de múltiples búsquedas exploratorias que nos van a permitir ingresar a esa conversación previa en la que se inscribe todo problema de investigación. El acceso a esa conversación es paulatino. Es como cuando ingresamos a un grupo que ya se conocen hace años: al principio nos quedamos afuera de los chistes, de las anécdotas, no conocemos las historias previas, ni siquiera lo que les pasó en el último encuentro, nos sentimos fuera de tono cuando hablamos porque todavía nos cuesta apropiarnos de la manera que tienen de decir las cosas. Solo si mantenemos nuestro interés y nos abrimos a esa dialéctica interrogativa de pregunta y respuesta que para Gadamer implica el comprender, comenzaremos a reconocer quiénes dialogan, de qué hablan, qué plantean, cuánta historia tienen ya transitada. Y cuando entendamos un chiste y logremos reírnos junto con el resto, recién ahí es el momento en que nos comenzamos a sentir parte.

Eso mismo sucede con un tema de investigación: para poder reconstruir las historias y los recorridos previos, no alcanza con mirar desde afuera y quedarnos con los textos que ya conocíamos, es necesario abrirnos a buscar y recuperar nuevas lecturas desconocidas hasta ese momento, lo que demanda una actitud exploratoria que mantenga siempre abierta la curiosidad.



Bibliografía

- Acuña, H. R. (1979). El Index Medicus latinoamericano: Una aspiración se hace realidad. *Educación Médica y Salud*, 13(3), 209-211.
- Aguilar, P. L., Gluzman, M., Grondona, A., & Haidar, V. (2014). ¿Qué es un corpus? *Revista de la Carrera de Sociología*, 4(4), 35-64.
- Almeida-Filho, N. (2020). Desigualdades en salud: Nuevas perspectivas teóricas. *Salud Colectiva*, 16, e2751. doi: [10.18294/sc.2020.2751](https://doi.org/10.18294/sc.2020.2751).
- Alonso, J. P. (2016). Derechos en el final de la vida: El rechazo de tratamientos médicos en los tribunales argentinos. *Salud Colectiva*, 12(3), 383-396. doi: [10.18294/sc.2016.989](https://doi.org/10.18294/sc.2016.989).
- Angell, M. (2006). *La verdad acerca de la industria farmacéutica: Cómo nos engaña y qué hacer al respecto*. Buenos Aires: Grupo Editorial Norma.
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology*, 8(1), 19-32. doi: [10.1080/1364557032000119616](https://doi.org/10.1080/1364557032000119616).
- Bae, J.-M. (2014). Narrative reviews. *Epidemiology and Health*, 36, e2014018. doi: [10.4178/epih/e2014018](https://doi.org/10.4178/epih/e2014018).
- Baiadera, A. P. A. (2018). ¿Articulación imposible? Debates parlamentarios sobre el voto joven y el voto extranjero en Argentina. *Diarios del Terruño. Reflexiones sobre Migración y Movilidad*, 3(5), 29-44.
- Bajtín, M. (2011). El problema de los géneros discursivos. En: *Estética de la creación verbal* (2ª ed.). Buenos Aires: Siglo XXI Editores.
- Banco Mundial. (2022). *Gasto en investigación y desarrollo (% del PIB)—Korea, Rep., Japan*. Recuperado de: <https://tinyurl.com/5n73h265>.
- Bireme, OPS, & OMS. (2008). *Metología LILACS: Guía de selección de documentos para la base de datos LILACS*. Bireme, OPS, OMS. Recuperado de: <https://tinyurl.com/2p8u4bh7>.
- Bolívar, A. (2013). La definición del corpus en los estudios del discurso. *Revista Latino-Americana de Estudios del Discurso*, 13(1), 3-8.
- Bradford, S. C. (1934). Sources of information on specific subjects. *Engineering: An Illustrated Weekly Journal*, 137(3550), 85-86.
- Cabezas-Clavijo, Á., Robinson-García, N., Escabias, M., & Jiménez-Contreras, E. (2013). Reviewers' ratings and bibliometric indicators: Hand in hand when assessing over research proposals? *PLoS ONE*, 8(6), e68258. doi: [10.1371/journal.pone.0068258](https://doi.org/10.1371/journal.pone.0068258).
- Caravedo, L. (1999). *Linguística del corpus: Cuestiones teórico-metodológicas aplicadas al español*. Salamanca: Ediciones Universidad Salamanca.
- Castillo Blasco, L. (2002). Trabajo intelectual y científico (Normas de estilo). En: *Introducción a la información científica y técnica*. Valencia: Universidad de Valencia.

- Clarivate Analytics. (2017a). *NCSTE of China and Clarivate Analytics partner on international scientific research collaboration report*. MediaRoom. Recuperado de: <https://tinyurl.com/5fh637fx>.
- Clarivate Analytics. (2017b). *Uncovering the right research*. Recuperado de: <https://youtu.be/C4hiRh9ydvI>.
- Clarivate (2021). *Web of Science Help*. Recuperado de: <https://tinyurl.com/5n7t9uvs>.
- Creighton, H. (2013). (De)construir la otredad: Las mujeres inmigrantes en la prensa escrita española | Revista de Paz y Conflictos. *Revista de Paz y Conflictos*, 6, 78-106. doi: [10.30827/revpaz.v6i0.810](https://doi.org/10.30827/revpaz.v6i0.810).
- Day, R. A. (2005). *Cómo escribir y publicar trabajos científicos* (3ª ed.). Organización Panamericana de la Salud. Recuperado de: <https://tinyurl.com/544sb8wx>.
- De Solla Price, D. J. (1963). *Little Science, Big Science*. New York: Columbia University Press.
- Del Vicario, M., Zollo, F., Caldarelli, G., Scala, A., & Quattrociocchi, W. (2017). Mapping social dynamics on Facebook: The Brexit debate. *Social Networks*, 50, 6-16. doi: [10.1016/j.socnet.2017.02.002](https://doi.org/10.1016/j.socnet.2017.02.002).
- Demonte, F. (2011). La construcción de la malnutrición infantil en la prensa escrita argentina durante la crisis de 2001. *Salud Colectiva*, 7(1), 53-71. doi: [10.18294/sc.2011.278](https://doi.org/10.18294/sc.2011.278).
- Dixon-Woods, M., Bonas, S., Booth, A., Jones, D. R., Miller, T., Sutton, A. J., Shaw, R. L., Smith, J. A., & Young, B. (2006). How can systematic reviews incorporate qualitative research? A critical perspective. *Qualitative Research*, 6(1), 27-44. doi: [10.1177/1468794106058867](https://doi.org/10.1177/1468794106058867).
- Eco, U. (1992). *Cómo se hace una tesis: Técnicas y procedimientos de investigación, estudio y escritura*. México: Gedisa.
- Elsevier. (2020). *Scopus: Content Coverage Guide*. Elsevier. Recuperado de: <https://tinyurl.com/4auwhnmy>.
- Elsevier. (2018). *Scopus*. Recuperado de: <https://tinyurl.com/bde4ez4p>.
- Farji Neer, A. (2015). Cuerpo, derechos y salud integral: Análisis de los debates parlamentarios de las leyes de Identidad de Género y Fertilización Asistida (Argentina, 2011-2013). *Salud Colectiva*, 11(3), 351-365. doi: [10.18294/sc.2015.721](https://doi.org/10.18294/sc.2015.721).
- Fingerman, S. (2006). Web of Science and Scopus: Current Features and Capabilities. *Issues in Science and Technology Librarianship*. doi: [10.5062/F4G44N7B](https://doi.org/10.5062/F4G44N7B).
- Foucault, M. (2002). Las regularidades discursivas. En *La arqueología del saber*. Buenos Aires: Siglo XXI Editores Argentina.
- García, Y. B. (2019). ¿Libertad de enseñanza? ¿Atentado contra la libertad de conciencia? Discusiones parlamentarias en torno al uso de los manuales del Hermano Damasceno en las aulas laicas del Uruguay de primera mitad del siglo XX. *Estudios digital*, 42, 31-50. doi: [10.31050/re.vi42.25129](https://doi.org/10.31050/re.vi42.25129).
- Garfield, E. (1955). Citation indexes for science: A new dimension in documentation through association of Ideas. *Science*, 122(3159), 108-111. doi: [10.1126/science.122.3159.108](https://doi.org/10.1126/science.122.3159.108).
- Garfield, E. (1963). Science Citation Index. En: *Science Citation Index 1961* (p. V-XVI). Philadelphia: Institute of Scientific Information.

- Garfield, E. (1975). The Social Science Citation Index, more than tool. *Current Contents*, 12, 6–9.
- Garfield, E. (1977). Will ISI'S Arts & Humanities Citation Index revolutionize scholarship? *Current Contents*, 32, 5–9.
- Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *Educational Researcher*, 5(10), 3–8. JSTOR. doi: [10.2307/1174772](https://doi.org/10.2307/1174772).
- Grant, M. J., & Booth, A. (2009). A typology of reviews: An analysis of 14 review types and associated methodologies. *Health Information and Libraries Journal*, 26(2), 91–108. doi: [10.1111/j.1471-1842.2009.00848.x](https://doi.org/10.1111/j.1471-1842.2009.00848.x).
- Grieneisen, M. L., & Zhang, M. (2012). A comprehensive survey of retracted articles from the scholarly literature. *PLoS ONE*, 7(10), e44118. doi: [10.1371/journal.pone.0044118](https://doi.org/10.1371/journal.pone.0044118).
- Guitelman, P. (2006). *La infancia en dictadura: Modernidad y conservadurismo en el mundo de Billiken*. Buenos Aires: Prometeo.
- Habermas, J. (2007). *Ciencia y técnica como "ideología"* (5ª ed.). Madrid: Tecnos.
- Haunschild, R., Bornmann, L., & Marx, W. (2016). Climate Change Research in View of Bibliometrics. *PLoS ONE*, 11(7), e0160393. doi: [10.1371/journal.pone.0160393](https://doi.org/10.1371/journal.pone.0160393).
- Hulme, E. W. (1923). *Statistical bibliography in relation to the growth of modern civilization: Two lectures delivered in the University of Cambridge in May, 1922*. Butler & Tanner Grafton. Recuperado de: <https://archive.org/details/statisticalbibli00hulmuoft>.
- IBM. (2021). *Acerca de la minería de textos*. Recuperado de: <https://tinyurl.com/28nudrs9>.
- Iosa, M. P. (2021). Vida digna y derecho a la salud en los fallos de la Corte Interamericana de Derechos Humanos. *Forum*, 12, 105–130.
- Kearns, C. E., Glantz, S. A., & Schmidt, L. A. (2015). Sugar Industry Influence on the Scientific Agenda of the National Institute of Dental Research's 1971 National Caries Program: A Historical Analysis of Internal Documents. *PLoS Medicine*, 12(3), e1001798. doi: [10.1371/journal.pmed.1001798](https://doi.org/10.1371/journal.pmed.1001798).
- Linne, J., & Angilletta, M. F. (2016). Violencia en la red social: Una indagación de expresiones online en adolescentes de sectores populares marginalizados del Área Metropolitana de Buenos Aires. *Salud Colectiva*, 12(2), 279–294. doi: [10.18294/sc.2016.741](https://doi.org/10.18294/sc.2016.741).
- Lotka, A. A. (1926). The frequency distribution of scientific productivity. *Journal of the Washington Academy of Sciences*, 16(12), 317–323.
- Lyons, J. (1984). *Introducción al lenguaje y a la lingüística*. Barcelona: Teide.
- Magariños de Morentin, J. A. (1991). *El mensaje publicitario*. Buenos Aires: Edicial.
- Marcuse, H. (1969). Industrialización y capitalismo en Max Weber. En *La sociedad industrial y el marxismo*. Buenos Aires: Quintana.
- Martinhago, F., Lavagnino, N. J., Folguera, G., & Caponi, S. (2019). Factores de riesgo y bases genéticas: El caso del trastorno por déficit de atención e hiperactividad. *Salud Colectiva*, 15, e1952. doi: [10.18294/sc.2019.1952](https://doi.org/10.18294/sc.2019.1952).
- Martinovich, V. (2015). La dimensión política del acceso abierto: ¿el conocimiento como bien público o como mercancía? *Salud Colectiva*, 11(3), 297. doi: [10.18294/sc.2015.718](https://doi.org/10.18294/sc.2015.718).

- Martinovich, V. (2016). Salud Colectiva: Una apuesta al acceso abierto. *Salud Colectiva*, 12(2), 145. doi: [10.18294/sc.2016.1005](https://doi.org/10.18294/sc.2016.1005).
- Martinovich, V. (2020). Indicadores de citación y relevancia científica: Genealogía de una representación. *Dados*, 63(2), e20190094. doi: [10.1590/001152582020218](https://doi.org/10.1590/001152582020218).
- Martinovich, V. (2021). Ciencia abierta en América Latina: Repensar la interdependencia dinámica entre las ciencias y la sociedad. *Revista Espacios*, 42(24), 1–14. doi: [10.48082/espacios-a21v42n24p01](https://doi.org/10.48082/espacios-a21v42n24p01).
- Martinovich, V., Arakaki, J., & Spinelli, H. (2014). Diez años de Salud Colectiva: Una aproximación a las reglas de juego del campo editorial científico. *Salud Colectiva*, 10(1), 5. doi: [10.18294/sc.2014.205](https://doi.org/10.18294/sc.2014.205).
- Medina-López, C., Marin-García, J. A., & Alfalla-Luque, R. (2010). Una propuesta metodológica para la realización de búsquedas sistemáticas de bibliografía (A methodological proposal for the systematic literature review). *Working Papers on Operations Management*, 1(2), 13. doi: [10.4995/wpom.v1i2.786](https://doi.org/10.4995/wpom.v1i2.786).
- Menéndez, E. L., & Di Pardo, R. B. (2008). La representación social negativa de los procesos de salud/enfermedad/atención en la prensa escrita. *Salud Colectiva*, 4(1), 9–30. doi: [10.18294/sc.2008.314](https://doi.org/10.18294/sc.2008.314).
- Mialon, M., Sêrodio, P., & Baeza Scagliusia, F. (2018). Criticism against the NOVA classification: Who are the protagonists? *World Nutrition*, 9(3), 176–240. doi: [10.26596/wn.201893176-240](https://doi.org/10.26596/wn.201893176-240).
- Mikki, S., Zygmuntowska, M., Gjesdal, Ø. L., & Ruwehy, H. A. A. (2015). Digital Presence of Norwegian Scholars on Academic Network Sites—Where and Who Are They? *PLOS ONE*, 10(11), e0142709. doi: [10.1371/journal.pone.0142709](https://doi.org/10.1371/journal.pone.0142709).
- Minaverry, C., Cáceres, V., Minaverry, C., & Cáceres, V. (2016). La problemática del arsénico en el servicio de agua en la provincia de Buenos Aires, Argentina. Análisis de casos jurisprudenciales. *Revista internacional de contaminación ambiental*, 32(1), 69–76.
- Minayo, M. C. de S. (2013). *La artesanía de la investigación cualitativa*. Buenos Aires: Lugar Editorial.
- Minayo, M. C. de S. (Org.) (2012). *Investigación social: Teoría, método y creatividad*. Buenos Aires: Lugar Editorial.
- Moscovici, S. (1979). *El psicoanálisis, su imagen y su público*. Buenos Aires: Editorial Huemul.
- Mota, A. M.; Bedrikow, R. (2021). Sufrimiento de mujeres en una ocupación urbana de São Paulo, Brasil: un desafío para el trabajo en atención primaria de la salud. *Salud Colectiva*, 17, e3358. doi: [10.18294/sc.2021.3358](https://doi.org/10.18294/sc.2021.3358).
- National Center for Biotechnology. (2005). Search Field Descriptions and Tags. En *PubMed Help*. National Center for Biotechnology Information (US). Recuperado de: <https://tinyurl.com/55f2au2b>.
- Noblit, G. W., & Hare, D. (1988). *Meta-ethnography: Synthesizing qualitative studies*. SAGE.
- O'Connor, M. (1978). Standardisation of bibliographical reference systems. *British Medical Journal*, 1(6104), 31–32. doi: [10.1136/bmj.1.6104.31](https://doi.org/10.1136/bmj.1.6104.31).

- Oddone, M. J. (2013). La vejez en los textos de lectura de la escuela primaria: Un recorrido entre fines del siglo XIX y los inicios del siglo XXI en Argentina. *Salud Colectiva*, 9(1), 27–40.
- OECD. (1963). *Proposed standard practice for surveys of Rresearch and development*. Recuperado de: <http://www.oecd.org/sti/innno/Frascati-1963.pdf>.
- OECD. (1992). *Oslo Manual: The measurement of scientific and technological activities*. European Commission. Recuperado de: <http://www.oecd.org/science/innno/2367580.pdf>.
- OECD. (1995). *Measurement of Scientific and Technological Activities (Canberra Manual)*. OECD Publishing. doi: [10.1787/9789264065581-en](https://doi.org/10.1787/9789264065581-en).
- Ojeda, M. N., Pastor, E. M., & Pisonero, C. G. (2015). El humor en la publicidad gráfica de la Dirección General de Tráfico en España (1960-2009). *Revista Latina de Comunicación Social*, 70, 765–781.
- Packer, A. L., Prat, A. M., Luccisano, A., Montanari, F., Santos, S., & Meneghini, R. (2006). El modelo SciELO de publicación científica de calidad en acceso abierto. En D. Babiní & J. Fraga, *Edición electrónica, bibliotecas virtuales y portales para las ciencias sociales en América Latina y El Caribe*. Buenos Aires: CLACSO. p. 191–208. Recuperado de: <https://tinyurl.com/4nsvb9v>.
- Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J., Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., ... Alonso-Fernández, S. (2021). Declaración PRISMA 2020: Una guía actualizada para la publicación de revisiones sistemáticas. *Revista Española de Cardiología*, 74(9), 790–799. doi: [10.1016/j.recesp.2021.06.016](https://doi.org/10.1016/j.recesp.2021.06.016).
- Parada, T. C. (2018). Subjetividad y autonomía: Significados y narrativas sobre la discontinuación de fármacos psiquiátricos. *Salud Colectiva*, 14(3), 513–529. doi: [10.18294/sc.2018.1861](https://doi.org/10.18294/sc.2018.1861).
- Parker, L., Williams, J., & Bero, L. (2018). Ethical drug marketing criteria for the 21st century. *BMJ*, k1809. doi: [10.1136/bmj.k1809](https://doi.org/10.1136/bmj.k1809).
- Patiño-Londoño, S. Y., Mignone, J., Castro-Arroyave, D. M., Gómez Valencia, N., & Rojas Arbeláez, C. A. (2016). Guías bilingües: Una estrategia para disminuir las barreras culturales en el acceso y la atención en salud de las comunidades wayuu de Maicao, Colombia. *Salud Colectiva*, 12(3), 415. doi: [10.18294/sc.2016.883](https://doi.org/10.18294/sc.2016.883).
- Pochintesta, P. A., & Mansinho, M. (2014). *Modelos de envejecimiento en la publicidad gráfica: Un análisis de género*. 3(3), 94–115.
- PRISMA. (2009). *Preferred Reporting Items for Systematic Reviews and Meta-Analyses*. Recuperado de: <http://www.prisma-statement.org/Default.aspx>.
- Ricœur, P. (1986). *Del texto a la acción*. México: Fondo de Cultura Económica.
- Rodríguez-Garrido, P. (2022). Geografías para un buen nacer: Una aproximación geo(-corpo)gráfica del parto en domicilio en Chile. *Salud Colectiva*, 18, e3848. doi: [10.18294/sc.2022.3848](https://doi.org/10.18294/sc.2022.3848).

- Rubenich, G. B., Heck, S. T., Hellmann, F., & Schlemper Junior, R. B. (2015). El uso de placebo en ensayos clínicos de fase III en Brasil. *Salud colectiva*, 11(1), 99–114. doi: [10.18294/sc.2015.418](https://doi.org/10.18294/sc.2015.418).
- Salud Colectiva. (2011). El “impacto” de las políticas nacionales de ciencia y técnica sobre las publicaciones científicas. *Salud Colectiva*, 7(2), 259–263. doi: [10.18294/sc.2011.384](https://doi.org/10.18294/sc.2011.384).
- Samaja, J. (2007). El análisis del proceso de investigación. En: *Epistemología y metodología*. Buenos Aires: Eudeba. p. 199–305.
- Slimovich, A. (2016). Política y redes sociales en Argentina. El caso de los candidatos presidenciales de 2011 en Twitter. *Signo y Pensamiento*, 35(68), 86–100. doi: [10.11144/Javeriana.syp35-68.prsa](https://doi.org/10.11144/Javeriana.syp35-68.prsa).
- Smythe, E., & Spence, D. (2012). Re-viewing literature in hermeneutic research. *International Journal of Qualitative Methods*, 11(1), 12–25. doi: [10.1177/160940691201100102](https://doi.org/10.1177/160940691201100102).
- U.S. National Library of Medicine. (2006). *Skill Kit: Truncation in PubMed®* [Newsletters]. Recuperado de: <https://tinyurl.com/2kmh6bxm>.
- U.S. National Library of Medicine. (2011). *NLM 175th Anniversary | Our Milestones*. Recuperado de: <https://apps.nlm.nih.gov/175/milestones.cfm>.
- U.S. National Library of Medicine. (2016a). *FAQ: Index Medicus Chronology* [FAQs, Help Files, Pocket Cards]. Recuperado de: <https://tinyurl.com/45dxb3jx>.
- U.S. National Library of Medicine. (2016b). *The MEDLINE Indexing Process: Determining Subject Content* [Training Material and Manuals]. Recuperado de: <https://tinyurl.com/mrwc7p3j>.
- U.S. National Library of Medicine. (2017a). *History of MeSH* [Other]. Recuperado de: https://www.nlm.nih.gov/mesh/intro_hist.html.
- U.S. National Library of Medicine. (2017b). *OLDMEDLINE Data*. Recuperado de: <https://tinyurl.com/ycxhathx>.
- Vallone, F. (2009). Pequeños grandes clientes: La publicidad de sucedáneos de la leche materna en dos revistas pediátricas de Argentina entre 1977 y 2006. *Salud Colectiva*, 5(1), 87–105. doi: [10.18294/sc.2009.232](https://doi.org/10.18294/sc.2009.232).
- Vanadia, L. N., & Rodríguez, M. L. (2019). El saber ver y el saber hacer en el hospital-asilo de Oliva (1914-1950): El caso Adela. *Salud Colectiva*, 15, e2211. doi: [10.18294/sc.2019.2211](https://doi.org/10.18294/sc.2019.2211).
- Wang, B., Pan, S.-Y., Ke, R.-Y., Wang, K., & Wei, Y.-M. (2014). An overview of climate change vulnerability: A bibliometric analysis based on Web of Science database. *Natural Hazards*, 74(3), 1649–1666. doi: [10.1007/s11069-014-1260-y](https://doi.org/10.1007/s11069-014-1260-y).
- Web of Science Training (Director). (2017). *Web of Science—Citation Sources* [Video]. Recuperado de: <https://tinyurl.com/yck4y5pp>.
- Weinstock, M. (1972). The new Social Sciences Citation Index (SSCI): Will add a new dimension to research on man and society. *Current Contents*, 21, 317–319.

Desde diversas áreas de investigación, la palabra “bibliografía” suele ser utilizada para designar el listado de referencias consultadas o el listado de textos propuestos para un curso, pero lo cierto es que la bibliografía es un campo de estudios en sí mismo, una disciplina con tradiciones, rupturas y discusiones internas, de la cual los distintos tipos de trabajos de “revisión bibliográfica” tomaron históricamente prestado muchas de sus características esenciales. Sin embargo, mientras que, en la década de 1980, la bibliografía como disciplina rompió con sus propios moldes descriptivos y enumerativos, los distintos tipos de estudios que toman la bibliografía como insumo siguen atrapados en su impronta de contabilizar, enumerar y describir existencias.

Entonces ¿por qué no intentar recuperar las nuevas perspectivas de la bibliografía, de la historia y de la sociología de los textos para enriquecer las formas en que las diversas áreas del conocimiento analizan las producciones textuales?; ¿por qué no abordar la propia narrativa científica con la riqueza con que ciertas áreas de las ciencias analizan otras producciones discursivas?; ¿es posible hacer este proceso sin repensar las formas de recuperar e interpretar esa bibliografía? Entendemos que enriquecer las formas de interpretar los textos nos demanda repensar las formas en que buscamos e interpretamos ese recorte de fuentes escritas que componemos y este es el camino que les proponemos transitar.

