



Fortalezas y debilidades de las principales bases de datos de información científica desde una perspectiva bibliométrica

Strengths and weaknesses of the main scientific information databases from a bibliometric perspective

Andrés Vuotto

*Universidad Nacional de Mar del Plata. Facultad de Humanidades.
Departamento de Ciencia de la Información, Argentina*
avuotto@mdp.edu.ar

 <http://orcid.org/0000-0002-1897-9515>

Victoria Di Césare

*Universidad Nacional de Mar del Plata. Facultad de Humanidades.
Departamento de Ciencia de la Información, Argentina*
vdicesare@mdp.edu.ar

 <http://orcid.org/0000-0001-7098-1914>

Natalia Pallotta

*Universidad Nacional de Mar del Plata. Facultad de Humanidades.
Departamento de Ciencia de la Información, Argentina*
npallotta@mdp.edu.ar

 <http://orcid.org/0000-0002-0137-2530>

RESUMEN:

En el presente trabajo se estudian las bases de datos de información científica más relevantes a nivel regional e internacional desde la perspectiva de sus características y funcionalidades. El objetivo principal consiste en identificar las fortalezas y debilidades de estas fuentes secundarias mediante el análisis de sus capacidades de indexación, cobertura, interoperabilidad y búsqueda y recuperación de información científica para la actividad bibliométrica. A partir de la observación y acopio de información desde los portales web de cada base evaluada, se desarrollan dos tablas diseñadas ad hoc que permiten el análisis y la comparación de su desempeño en función de una serie de dimensiones que van desde lo formal (origen, propósito, alcance) hasta lo funcional (acceso, formato, calidad). Se reflexiona sobre el aporte que estas fuentes pueden realizar a la construcción de matrices que permitan elaborar y calcular indicadores para la evaluación de la actividad científica. Se observa una marcada tendencia de los sistemas hacia el servicio de búsqueda y recuperación para la investigación bibliográfica, sin valor añadido para la reutilización de los datos de los registros recuperados. Se concluye que la mayoría de los sistemas estudiados requieren un importante trabajo de reestructuración y enriquecimiento de los registros para que sus servicios de exportación de datos constituyan verdaderas herramientas para la investigación bibliométrica.

PALABRAS CLAVE: Bases de datos, Fuentes de información científica, Bibliometría, Metodología de la investigación.

ABSTRACT:

The present papers studies the most relevant scientific information databases at regional and international level from the perspective of their characteristics and functionalities. The main objective is to identify the strengths and weaknesses of this secondary sources through the analysis of their indexing capabilities, coverage, interoperability and scientific information search and recovery for the bibliometric activity. On the basis of information observation and collection from each database website, two ad hoc design tables are developed in order to analyse and compare their performance according to a series of dimensions that cover from formal aspects (origin, purpose, scope) to functional aspects (access, format, quality). Reflection on the contribution of these sources to the matrix construction that allow the development and calculation of indicators for scientific activity evaluation is made. A strong tendency of the systems towards the search and recovery service for bibliographic research, without added value for the reuse of data, is observed. It is concluded that most of the studied systems require hard work on the restructuring and enrichment of the records so that their data export services constitute true tools for bibliometric research.

KEYWORDS: Databases, Scientific information sources, Bibliometrics, Research methodology.

Recepción: 31 de julio de 2020 | Aceptación: 30 de agosto de 2020 | Publicación: 1 de octubre de 2020

Cita sugerida: Vuotto, A., Di Césare, V. y Pallotta, N. (2020). Fortalezas y debilidades de las principales bases de datos de información científica desde una perspectiva bibliométrica. *Palabra Clave (La Plata)*, 10(1), e101. <https://doi.org/10.24215/18539912e101>



1. INTRODUCCIÓN

Las bases de datos científicas se han convertido en herramientas indispensables para los estudios destinados a analizar la producción intelectual de una institución, disciplina o región. La principal dificultad que presentan es que sus estructuras distan de ser homogéneas entre sí. Este inconveniente, sumado al problema que supone conseguir una buena calidad de los datos, se refleja en una serie de desafíos metodológicos que es preciso resolver de manera consistente para extraer, luego, resultados representativos de la realidad que se desea mostrar y analizar. En el caso de Argentina y la región, otro obstáculo lo constituye la carencia de fuentes de datos que registren su producción científica de forma completa y confiable. Frente a esta situación, muchos estudios de esta naturaleza terminan realizándose recogiendo las publicaciones de bases de datos internacionales (Miguel, de Moya Anegón y Herrero Solana, 2006).

Las bases de datos bibliográficas surgen con el propósito de mejorar la recuperación de información e introducen una dinámica de análisis de los artículos científicos alternativa a las formas de representación basadas en la lingüística y en la indización. Desde su aparición, que data del año 1954 con la creación del *Science Citation Index* (SCI), estas bases han servido para analizar la producción científica a nivel disciplinar, institucional o geográfico, y para generar indicadores bibliométricos para la evaluación científica. Estos indicadores han posicionado a las bases de datos como fuentes de información esenciales para mapear la ciencia, su evolución y dinámica. No obstante, también han recibido críticas por los sesgos que presentan, que favorecen a las ciencias básicas y aplicadas por sobre las ciencias sociales y humanidades (De Filippo, Sanz-Casado, Urbano Salido, Ardanuy y Gómez-Caridad, 2011). Además, existe un marcado desbalance en la representación geográfica e idiomática; en *Web of Science* (WoS) y *Scopus* prevalece una fuerte inclinación por la producción estadounidense y anglosajona en general, mientras que los países centrales no anglófonos se encuentran subrepresentados, al igual que los trabajos de regiones provenientes de la periferia científica (Delgado y Repiso, 2013; Mongeon y Paul-Hus, 2016; Rafols, Chavarró y Ciarli, 2016). Este panorama, sumado a la amplia variedad de fuentes de datos existentes en la actualidad, complejiza la segmentación para los estudios bibliométricos, por la dificultad que conlleva delimitar una muestra representativa.

El movimiento de acceso abierto y la creación de bases de datos regionales han permitido, en parte, hacer frente a esta situación de desventaja al brindar nuevos circuitos de circulación de publicaciones y, con ello, mejorar la capacidad de evaluarlas (Alperin y Fischman, 2015; Beigel y Salatino, 2015). Una de las últimas y más novedosas iniciativas en este sentido es el proyecto *AmeliCA*¹, una plataforma cooperativa creada para fortalecer el modelo de publicación en acceso abierto para Latinoamérica y el Sur Global. Esta iniciativa, que fue generada por RedALyC y cuenta con el apoyo de UNESCO y CLACSO, surge “en respuesta a la crisis de sustentabilidad económica, la falta de reconocimiento ante los sistemas vigentes de evaluación de la ciencia y ante la exclusión de la mayor parte de las revistas de la región, lo cual demanda la formulación de estrategias cooperativas en donde los diversos actores de la comunicación científica apoyen, reconozcan y sostengan el Acceso Abierto” (AmeliCA, 2020).

A las fuentes tradicionales antes mencionadas se suman ahora nuevas alternativas que amplían y resignifican el concepto de “base de datos bibliográfica”. Estas nuevas herramientas han generado una verdadera revolución en términos de la capacidad de ampliar la recuperación y acceso y de medir el comportamiento de mucha producción científica “invisible” en otras fuentes, además de posicionarse como una competencia seria de las bases de corriente principal. Por un lado, la disponibilidad de fuentes alternativas permite corroborar los datos con los que se está trabajando e incrementar su precisión. Por el otro, hacen posible reducir los costos para aquellas investigaciones que no cuentan con los recursos para acceder a los datos de corriente principal. Dos de estas fuentes merecen especial atención por las novedosas dinámicas que introdujeron para el análisis y representación de la actividad científica: *Google Scholar* (GS)² y *Dimensions*.³

Sobre GS es destacable la amplia cobertura que posee, recogiendo publicaciones de una multiplicidad de fuentes académicas y con una variedad enorme de tipologías documentales e idiomas (Orduña Malea,

Martín Martín, Ayllón y Delgado López Cózar, 2016). En este sentido, abarca una serie de documentos y fuentes de información que no son considerados en la corriente principal. Además brinda información sobre la presencia de los autores en la plataforma, mediante la construcción del índice h en sus perfiles personales, y sobre la cantidad de citas que recibieron los documentos. En este sentido, entre los servicios que ofrece se encuentra *Google Scholar Metrics*, una herramienta que elabora rankings de revistas indizadas en GS y permite medir en forma rápida y accesible la visibilidad de artículos recientes en publicaciones académicas. Estas posibilidades que ofrece GS permiten obtener un panorama de la investigación que sería imposible de rastrear en el *mainstream*. Sin embargo, una de las críticas que ha recibido es que la amplia cobertura documental va en detrimento de la precisión en los registros. Otra de las limitaciones encontradas tiene que ver con la accesibilidad a grandes volúmenes de datos para la medición de la actividad científica (Torres Salinas, Ruiz Pérez y Delgado López Cozar, 2009). Frente a esta situación nace en octubre del 2006 el software *Publish or Perish*, una aplicación informática que recupera y analiza citas académicas capturadas de GS (Harzing, 2007).

Respecto a *Dimensions*, se ha consolidado como una herramienta con un gran potencial para los análisis bibliométricos al unificar en su plataforma el acceso a una multiplicidad de datos, que van desde artículos y libros a menciones en las redes sociales, citas, pruebas clínicas, patentes e informes de financiamiento. Los documentos que contempla son un reflejo de todo el proceso de la investigación y no sólo de la instancia final de publicación de los resultados, lo que posibilita obtener un panorama más completo y exhaustivo de la actividad científica.

En líneas generales, existen numerosos esfuerzos realizados para evaluar la producción científico-académica en bases de datos desde una perspectiva bibliométrica, pero poco se ha dicho sobre los alcances y limitaciones de las fuentes de datos para este tipo de estudios. Entre la bibliografía existente se encuentran los trabajos que establecen comparativas entre *WoS* y *Scopus* (Escalona, Cerdá Hernández y Pulgarín, 2010; Fingerman, 2005; Gireesh Kumar, 2013; Gorraiz y Schloegl, 2007; Jiménez Noblejas y Perianes Rodríguez, 2014; LaGuardia, 2010), los que suman a GS como un tercer elemento en esa comparación (Bakkalbasi, Bauer, Glover y Wang 2006; Jacson, 2005; Martín Martín, Orduna Malea, Thelwall y López Cózar, 2018; Meho y Yang, 2007; Mingers y Lipitakis, 2010) y los que se limitan al estudio de otras bases de datos de forma individual por su alcance regional o su cobertura temática o institucional (Alonso Gamboa y Certo, 2015; Michán Aguirre, Calderón Rojas, Nitxin Castañeda Sortibrán y Rodríguez Arnáiz, 2014; Puente Lanzarote, del Campo Hilario y Ruiz de Luzuriaga Peña, 2001). Los inicios de esta clase de estudios se remontan al año 2004 con la creación de *Scopus*, que alentó la generación de estudios que permitieran determinar, entre esta y *WoS*, cuáles bases de datos satisfacían mejor las necesidades de información de los investigadores (Jiménez Noblejas y Perianes Rodríguez, 2014; Torres Salinas, Jiménez Contreras y Delgado López Cózar, 2009). La mayoría de estos trabajos se orientaron a la representatividad de las ciencias básicas y aplicadas, mientras que los estudios sobre las ciencias sociales y humanidades se inclinaron más por analizar su escasa visibilidad en este tipo de fuentes (De La Laurencie y Maddi, 2019; Kulczycki, Engels, Pölönen, Bruun, Duskova, Guns, R., y Zuccala, 2018; Mosbah Natanson y Gingras, 2014). Worwell (2001), por su parte, hace foco en los alcances de la bibliometría para el estudio de las comunidades académicas y, en ese sentido, sugiere que las bases de datos constituyen herramientas fundamentales para el análisis cuantitativo de la producción científica; su principal utilidad en términos bibliométricos es la de brindar un diagnóstico para la toma de decisiones y el establecimiento de políticas. Otro antecedente próximo a los fines de este trabajo lo constituye el estudio de Somoza, Guallar, Rodríguez Gairín y Abadal (2017), que analiza la presencia de la producción científica española en numerosas bases de datos de corriente principal. Sin embargo, el foco de este estudio es la visibilidad de la producción analizada y no tanto las particularidades de las fuentes que la albergan. Cabe destacar también el aporte hecho por Gusenbauer (2018), quien analiza y compara el volumen total de datos almacenado en doce motores de búsqueda académicos y bases de datos bibliográficas utilizadas por los investigadores en general.

El presente trabajo tiene por objetivo analizar las principales bases de datos de información científica, haciendo hincapié en las posibilidades y limitaciones que presentan para los estudios bibliométricos e identificando sus principales características. Esto último adquiere especial relevancia debido a la necesidad de unificar criterios entre estructuras de datos para medir la producción científica de forma fiable y eficaz, lo cual constituye uno de los principales desafíos metodológicos en los estudios bibliométricos.

2. METODOLOGÍA

Se seleccionó un conjunto de 17 bases de datos, tanto especializadas como multidisciplinarias, y de alcance regional como internacional. Al momento de realizar esta selección, se identificó una primera dificultad en la definición de lo que se entiende por “principales bases de datos”. No fue posible localizar publicaciones o autoridades en la temática que permitieran acotar el conjunto, por lo que se optó por generar una síntesis propia para justificar la elección de las bases desde una o varias de las siguientes perspectivas: antigüedad, respaldo institucional o empresarial, caudal de cosecha, comunidad de usuarios, actualización y originalidad. Cabe aclarar que se excluyeron todos aquellos sistemas que funcionan por autoarchivo, como son los repositorios institucionales y temáticos, privilegiando en su lugar la inclusión de herramientas que seleccionan sus fuentes, principalmente revistas científicas, mediante postulación. Las bases de datos elegidas fueron:

- **BASE:** motor de búsqueda de recursos académicos operado desde 2004 por la Universidad de Bielefeld, de Alemania.
- **Dialnet:** portal bibliográfico de literatura científica hispana creado en 2002 por la Universidad de La Rioja, de España. Actualmente pertenece a la Fundación Dialnet.
- *Dimensions:* sistema de información científica lanzado en 2018 por la empresa inglesa Digital Science.
- *DOAJ:* directorio de revistas en acceso abierto lanzado en 2003 por la Universidad de Lund, de Suecia.
- *EBSCOhost:* plataforma de investigación en línea creada en la década de 1990 por la empresa estadounidense *EBSCO Information Services*.
- *Google Scholar:* motor de búsqueda de literatura académica lanzado en 2004 por la empresa estadounidense *Google*.
- *JournalTOCs:* buscador de tablas de contenidos de revistas científicas creado en 2009 por la Universidad Heriot Watt, de Escocia.
- *JSTOR:* biblioteca digital de revistas académicas, libros y fuentes primarias ideada en 1994 por la Fundación Andrew W. Mellon, de Estados Unidos. Actualmente pertenece a la organización sin fines de lucro Ithaka.
- *LILACS:* base de datos desarrollada en 1982 por la Biblioteca Virtual en Salud de BIREME, de la OPS/OMS.
- *OpenAire:* infraestructura de información científica en acceso abierto fundada en 2008 por la Comisión Europea, de la Unión Europea.
- *PubMed:* base de datos creada y mantenida desde 1996 por la Biblioteca Nacional de Medicina de Estados Unidos.
- *RedALyC:* sistema de indización de revistas en acceso abierto fundada en 2003 por la Universidad Autónoma del Estado de México.
- *REDIB:* sistema de información científica creado en 2017 por el Consejo Superior de Investigaciones Científicas de España y la Fundación Universia.

- *SciELO*: biblioteca científica electrónica iniciada en 1997 por la Fundación FAPESP, de Brasil, y BIREME, de la OPS/OMS.
- *ScienceDirect*: base de datos de publicaciones científicas operada desde 1997 por la empresa neerlandesa *Elsevier*.
- *Scopus*: base de datos bibliográfica operada desde 2004 por la empresa neerlandesa *Elsevier*.
- *WOS*: base de datos bibliográfica creada en 1997 por la empresa canadiense *Thomson Reuters*, sobre la base del Institute for *Scientific Information*. Actualmente pertenece a la empresa estadounidense *Clarivate Analytics*.

Seguidamente, se analizó y comparó su desempeño en función de las siguientes dimensiones:

- **Origen:** gubernamental, no gubernamental, empresarial, institucional.
- **Propósito:** científico, educativo, divulgativo, lucrativo, no lucrativo.
- **Alcance geográfico:** diversidad de países o regiones que abarca la base de datos.
- **Cobertura temática:** diversidad de disciplinas que abarca la base de datos.
- **División del conocimiento:** categorías en las que se estructura internamente la base de datos.
- **Cobertura temporal:** período de tiempo que abarca la base de datos.
- **Cobertura idiomática:** diversidad de idiomas de las publicaciones.
- **Tipo de acceso:** acceso abierto o por suscripción, tanto gratuita⁴ como paga.
- **Fuentes cosechadas:** cantidad de revistas y otras fuentes de metadatos desde las que se cosechan los registros.
- **Interfaz de búsqueda:** características de la interfaz de búsqueda respecto de su usabilidad, disponibilidad de ayudas, filtros, texto predictivo, etc. Ponderado como muy completa, completa, incompleta.
- **Presentación de los resultados:** características de las páginas de resultados respecto de su diseño y disposición, opciones de filtrado, disponibilidad de hipervínculos para la navegabilidad, legibilidad de los resúmenes completos, etc. Ponderado como muy completa, completa, incompleta.
- **Exhaustividad de los registros:** variedad de campos que componen los registros.
- **Formatos de exportación:** variedad de formatos en los que es posible realizar exportaciones.
- **Capacidad de exportación:** número máximo de registros que es posible exportar.
- **Campos para la exportación:** variedad de campos que es posible exportar.
- **Modalidad de exportación:** exportación abierta o por suscripción paga o gratuita.
- **Calidad del dato:** normalización o curación de los datos exportados, unicidad e integridad de los registros, estructura del formato, control de autoridades y posibilidad de manipulación en softwares de gestión de datos. Ponderando como calidad alta los casos en los que se observó una adecuada estructura del formato, normalización de campos principales (autor/es, fuente, palabras clave, etc.) y exhaustividad informativa de los registros; calidad media para los casos en los que alguno de los tres aspectos presentó ausencia o deficiencia en su desarrollo; y calidad baja para los casos donde no se identificaron dos o más características.

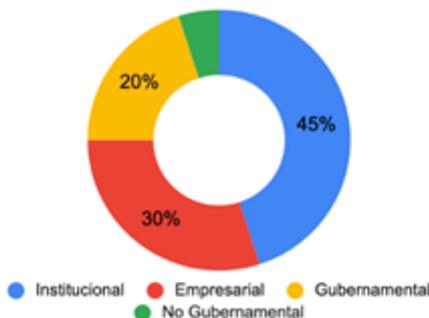
Por último, se confeccionaron dos tablas de doble entrada para caracterizar formal y funcionalmente cada base de datos, y para facilitar su comparación (Anexo). Todas las consultas a los sitios web de las bases fueron efectuadas en julio de 2020.

3. RESULTADOS Y DISCUSIÓN

El análisis formal de las 17 bases de datos arrojó, en una primera instancia y siguiendo un criterio de cuenta completa, que predomina el origen institucional, alcanzando a casi la mitad de los casos

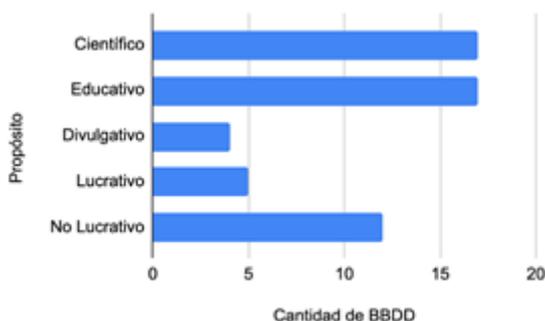
estudiados, seguido de cerca por el origen empresarial. En tres ocasiones se observó una combinación de los orígenes institucional, gubernamental y no gubernamental, lo cual muestra una asociación entre instituciones científicas y educativas con gobiernos y otras organizaciones en la creación y mantenimiento de este tipo de sistemas (Figura 1). Por su parte, los propósitos científico y educativo se encontraron presentes en todas las bases, introduciendo además muy pocas de ellas un objetivo divulgativo orientado a usuarios no especializados. La fuerte presencia del fin no lucrativo (n=12) se identificó ligado a los orígenes mayoritariamente institucionales, gubernamentales y no gubernamentales (Figura 2).

FIGURA 1.
Origen de las BBDD.



Fuente: elaboración propia.

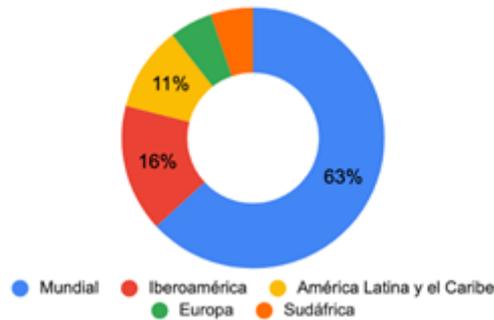
FIGURA 2.
Propósito de las BBDD.



Fuente: elaboración propia.

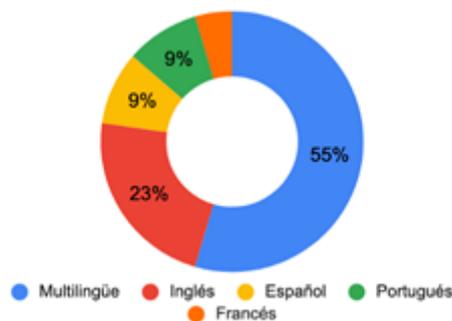
De acuerdo con el alcance geográfico propio del contenido de cada base de datos, se obtuvo que la gran mayoría posee una cobertura mundial (n=12), mientras una minoría se compone de revistas, libros y otros tipos de documentos específicamente iberoamericanos, latinoamericanos, europeos y sudafricanos (Figura 3). Esta dimensión de análisis se vincula directamente con la cobertura idiomática, la cual resulta predominantemente multilingüe (n=12) en consonancia con la cobertura mundial antes señalada (Figura 4).

FIGURA 3.
Alcance geográfico de las BBDD.



Fuente: elaboración propia.

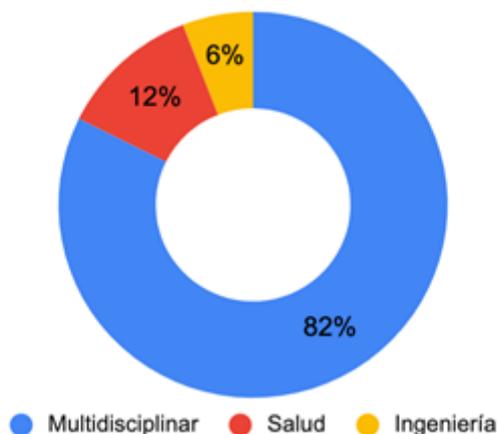
FIGURA 4.
Cobertura idiomática de la BBDD.



Fuente: elaboración propia.

La cobertura temporal fue la dimensión formal sobre la que menos cantidad de información fue posible recabar ya que 11 de las 17 bases no proporcionaron este dato. Para los casos restantes se observó que la mayoría declara su alcance desde los siglos XIX y XX hasta la actualidad y solamente una dice poseer contenidos del siglo XVII en adelante. Seguidamente, a nivel temático fue posible identificar que la gran mayoría de las bases es de tipo multidisciplinar ($n=14$), restando dos de ellas, *LILACS* y *PubMed*, especializadas en Salud, y tan solo una, *BASE*, abocada a la Ingeniería (Figura 5). Resulta de interés destacar que 4 bases, las 3 temáticas antes mencionadas y *DOAJ*, incorporan sistemas de organización del conocimiento, como la Clasificación Decimal de Dewey, la Clasificación de la Biblioteca del Congreso de EE.UU., el Tesouro *MeSH* y su versión traducida y ampliada *DeCS*, a la estructuración de sus colecciones. El resto de las bases, con excepción de *OpenAire*, presentan divisiones del conocimiento diseñadas ad hoc con distintos grados de complejidad y exhaustividad. Se observaron desde grandes categorías como las que aplican *EBSCOhost*, *RedALyC*, *Scopus* y *WoS*, hasta jerarquías con un nivel de especificidad mucho mayor como las de *Dialnet* y *Dimensions*, que dan cuenta de la visión que sus instituciones o empresas productoras tienen de la ciencia.

FIGURA 5.
Cobertura temática de las BBDD.

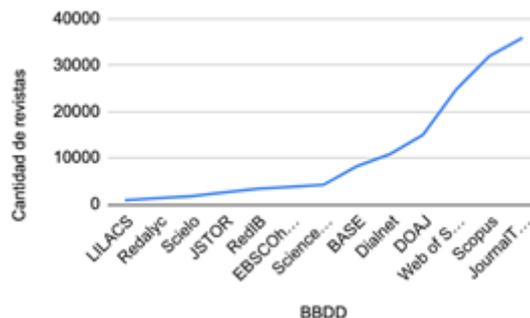


Fuente: elaboración propia.

Si bien no todas las bases estudiadas consignan claramente cuáles ni cuántas son las fuentes desde las que cosechan los registros que componen sus acervos, en 15 casos fue posible identificar o reconstruir esta información. Tal como se muestra en la Figura 6, las cantidades de revistas cosechadas por las 13 bases allí consignadas son muy variables, y van desde un mínimo de 895 para *LILACS*, sistema temático, hasta un máximo de 35.903 para *JournalTOCs*, colección que recoge específicamente tablas de contenido. Llamamos la atención los proyectos de origen regional, *RedALyC* y *SciELO*, por el acotado espectro de fuentes con el que trabajan, lo cual podría deberse a los exigentes requisitos de formato que imponen a las revistas para su postulación. Por su parte, se encontró que las bases de datos *OpenAire* y *PubMed* cosechan a otras bases.

No todos los sistemas integran en sus prestaciones funciones que posibiliten la extracción de metadatos y, en los casos en que lo hacen, existen diferencias técnicas que se deben considerar a priori en cualquier investigación que haga uso de esa información. Sólo 14 de las bases de datos estudiadas permiten obtener exportaciones y muestras de metadatos significativas. Además de *Dialnet*, que no posee servicio de exportación, aquí se excluye a *RedALyC* y a *ScienceDirect* porque sólo permiten descargar de a un registro por vez. En este subgrupo de 14 bases, 3 son exclusivamente de acceso por suscripción paga (*EBSCOhost*, *Scopus* y *WoS*) por lo que significan un costo, ya sea a nivel institucional o personal, para poder acceder al servicio de búsqueda y exportación. El resto de los sistemas son gratuitos aunque para acceder a funciones de exportación en ocasiones solicitan generación de un perfil de usuario. También hay algunos portales, como son *Dimensions*, *JournalTOCs* y *JSTOR*, que ofrecen dos modalidades de trabajo, gratuita y paga, con diferencias en cuanto al servicio de exportación dependiendo del caso. De todas maneras, si se contemplan en conjunto los tipos de acceso abierto y por suscripción gratuita, se obtiene que casi el 80% de las bases poseen alguna forma de consulta no privativa (Figura 7).

FIGURA 6.
Cantidad de revistas cosechadas.



Fuente: elaboración propia.

FIGURA 7.
Tipos de acceso a las BBDD.



Fuente: elaboración propia.

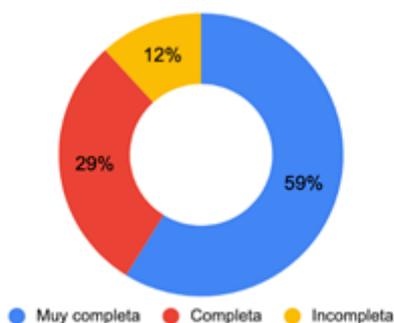
En lo que respecta a las dimensiones funcionales de análisis, se realizó un primer acercamiento a las bases desde esta perspectiva a partir de la observación de sus interfaces de búsqueda y de presentación de resultados, ponderando los productos de dicha observación en función de las categorías muy completa, completa e incompleta. Para ambas dimensiones se concluyó que la mayoría de las bases poseen tanto una interfaz de búsqueda como de presentación de resultados muy completa en el sentido de que proveen al usuario de modalidades de búsqueda simple y avanzada, con una gran variedad de filtros y opciones de refinamiento, navegabilidad por hipervínculos, lectura de los registros acotada y ampliada con resúmenes completos, ayudas didácticas y multimedia, y diseños claros y amigables. Una minoría, sin embargo, no cumplió con alguno o varios de estos criterios, como son los casos de GS y RedALyC, los cuales brindan posibilidades de uso superficiales o poco funcionales. En pocas ocasiones, además, se observó una disparidad en las prestaciones ofrecidas por un mismo sistema. Tales son los casos de Dialnet, que no posee búsqueda avanzada pero sí ofrece un despliegue de resultados con detalles de interés, y RedIB, cuya interfaz de búsqueda presenta todas las características buscadas pero cuya disposición de registros omite los resúmenes (Figuras 8 y 9).

FIGURA 8.
Interfaz de búsqueda de las BBDD.



Fuente: elaboración propia.

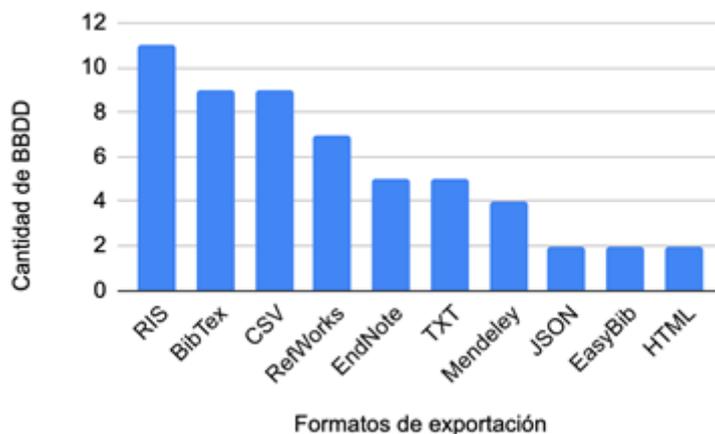
FIGURA 9.
Presentación de resultados de búsqueda en las BBDD.



Fuente: elaboración propia.

Todas las bases que permiten exportar registros incluyen formatos de salida de metadatos potencialmente compatibles con el trabajo bibliométrico y con la actividad de minería de datos. El análisis de esta dimensión consistió en realizar pruebas mediante la selección y exportación de estructuras de archivos que facilitaran el procesamiento y la construcción de bases de datos para el cálculo de indicadores, así como también mediante la importación en programas específicos que interpretaran cuantitativamente la información. Cualquier estructura de datos organizada con igual criterio para todos los registros permite, en principio, su modelado para la investigación estadística, pero aun así existen formatos establecidos que en cierta forma colaboran con la generación de instrumentos de investigación, reduciendo el margen de error y requiriendo una menor labor en el preparado de las matrices para su posterior análisis. Desde esta perspectiva, los formatos que mejor se ajustan a la actividad mencionada son *RefWorks*, *EndNote*, *RIS*, *BibTex*, *CSV*, *XLS* y *Refman*. En un segundo grupo se podrían ubicar los restantes formatos localizados en los sistemas estudiados, como *JSON*, *MARC*, *RDF*, *RTF*, *CITA*, *HTML*, *inCites*, *FECYT CVN* y *TXT* delimitado por *TAB*. Las 16 bases de datos en las que se identificaron servicios de exportación trabajan más predominantemente con los formatos que se muestran en la Figura 10.

FIGURA 10.
Formatos de exportación más predominantes.



Fuente: elaboración propia.

Una dimensión de análisis no menor vinculada a la extracción de datos refiere a la cantidad de registros que es posible obtener. De todo el conjunto solo DOAJ posee la funcionalidad de exportar la totalidad de la colección, en lugar de los resultados de una búsqueda determinada, lo cual puede ser interesante para analizar el contenido de esa base en particular, pero resulta infructífero si lo que se desea es circunscribir el estudio a un dominio temático, disciplinar, institucional, etc. Por su parte, RedALyC y *ScienceDirect* permiten únicamente realizar exportaciones de a un registro por vez, siendo esta alternativa de poca utilidad para la bibliometría. Aún así, la mayoría de las bases permite obtener grandes cantidades de registros mediante selecciones globales (n=11) y, dentro de este subgrupo, apenas 4 sistemas brindan la combinación de opciones ideal para la consecución de un estudio bibliométrico: realizar búsquedas y seleccionar los resultados globalmente, para luego exportar ilimitadamente todos los registros recuperados. Estos son *EBSCOhost* y *JSTOR* por suscripción paga, *GS* por suscripción gratuita y *LILACS* de manera abierta. Si bien las 7 bases restantes de este subgrupo trabajan con cantidades que van desde los 500 hasta los 10.000 registros plausibles de ser seleccionados globalmente y exportados, la sola presencia de un número limitado puede significar un obstáculo para la investigación (Figura 11).

FIGURA 11.
Capacidad de exportación de las BBDD.



Fuente: elaboración propia.

Finalmente, la utilidad de cada herramienta se termina de establecer por la calidad de sus datos exportables, aspecto que puede ser abordado desde distintos criterios y enfoques técnicos. En este estudio se trabajó a partir de una observación simple en cuanto a la normalización o curación de los datos exportados, unicidad e integridad de los registros, estructura del formato, control de autoridades y posibilidad de manipulación en *software* de gestión de datos. La mayoría de los sistemas presentó una correcta estructura de los archivos exportados, aunque respecto de *OpenAire* es necesario remarcar que solo ofrece la extracción de muestras en CSV, siendo este uno de los formatos que más deficiencias presenta en su estructuración. En relación con la normalización de los datos, solo *Dimensions*, *EBSCOhost*, *LILACS*, *PubMed*, RedIB y *SciELO* presentaron una adecuada aplicación de este ítem en las muestras obtenidas. En el caso particular de *JournalTOCs* no fue posible su evaluación debido a que durante la realización de las pruebas la función de exportación arrojó archivos de tipo RIS pero vacíos de datos. Además, se observaron algunas limitaciones en relación con la información que contienen los registros exportados, ya sea por ausencia de metadatos o por las características propias de los formatos de exportación de acuerdo con los cuales, por ejemplo, un formato de cita incluye menos elementos de metadatos que un formato desarrollado para descripción integral del documento, como son RIS o BibTex, o incluso que un formato genérico pero con posibilidad de incluir gran cantidad de información estructurada, como es el caso de CSV. El 70% de los sistemas presentaron registros exhaustivos en sus resultados de búsqueda, lo cual generó la expectativa de obtener exportaciones igualmente completas. Sin embargo, el análisis de los campos para la exportación arrojó ausencias totales o parciales de datos, principalmente de metadatos de descripción de contenido como son las palabras claves y el resumen, en bases como JSTOR, *LILACS*, *OpenAire* y *GS*. En síntesis, analizando la calidad de los datos tomando en consideración la ponderación ya desarrollada en la Metodología, se obtuvo que la mitad de los sistemas con servicio de exportación son de calidad media (n=8), seguidos por un 31,25% de calidad alta (n=5) y apenas un 18,75% de calidad baja (n=3).

CONCLUSIONES

La bibliometría toma como recurso para sus investigaciones los datos que conforman las bases de datos de información científica. Sus objetivos de explorar, describir y evaluar la producción científica de todos los dispositivos que conforman los espacios de investigación, así como también de los distintos países, regiones, temas y disciplinas, sólo pueden concretarse a partir de la información que estos sistemas gestionan. En este sentido, cobra gran relevancia poder conocer y analizar en profundidad la manera en que las bases resuelven su servicio de acopio, procesamiento y exportación de datos para su posterior reutilización. Cada sistema ofrece prestaciones que permiten buscar información científica, obtener datos de calidad sobre la producción recuperada y realizar exportaciones de esos datos, que luego serán sintetizados y representados a partir del cálculo de indicadores bibliométricos. Estos pilares se constituyen como centrales en una base de datos y responden a varios aspectos formales y funcionales que definen al servicio brindado y lo convierten en una herramienta potencialmente útil para el trabajo bibliométrico.

Entre los sistemas estudiados se pudo observar que ninguno responde adecuadamente a la totalidad de las dimensiones observadas. En todos ellos se pudo determinar la existencia de debilidades y fortalezas que se deben considerar en cada etapa de uso (búsqueda, selección, exportación, validación y modelado de la información obtenida). Será necesario dedicar una significativa cantidad de tiempo en el armado de muestras de datos, dependiendo del estudio a desarrollar, a partir de búsquedas en varios sistemas, en la reestructuración de la información atendiendo a la singularidad de cada exportación, y en la conformación de bases de datos enriquecidas luego de una adecuada intervención e integración de los datos fuente. Se plantea la necesidad de que las bases de datos de información científica continúen fortaleciendo su rol como sistemas proveedores de contenido curado en relación a estándares internacionales y nomenclaturas propias del campo científico en el que se desarrollan. Esta función debería ser inherente a sus objetivos; en lugar de que

se constituyan únicamente como sistemas de búsqueda y recuperación que sólo ofrecen acceso al documento, resultaría de gran utilidad que posibilitaran la valoración de sus datos a partir del cálculo de indicadores bibliométricos, colaborando en la reducción del pesado trabajo de minería que en muchos casos se debe llevar a cabo para convertir los datos obtenidos en material útil para la investigación.

Este artículo busca poner de relieve las prestaciones y problemáticas de cada sistema, en la creencia de que es necesario abordar el trabajo bibliométrico desde el conocimiento de los orígenes, objetivos y propiedades de cada herramienta, tanto para aprovechar al máximo sus funcionalidades, como también para evitar arrastrar involuntariamente sus sesgos y limitaciones a los resultados de investigación.

REFERENCIAS

- Alonso Gamboa, J., y Cetto, A. M. (2015). Latindex: revistas científicas iberoamericanas y cooperación regional. *Ciência da informação*, 44(2). Recuperado de <http://revista.ibict.br/ciinf/article/view/1772/2366>
- Alperin, J. P. y Fischman, G. (Eds.) (2015). *Hecho en Latinoamérica: acceso abierto, revistas académicas e innovaciones regionales*. Buenos Aires: CLACSO. Recuperado de <http://biblioteca.clacso.edu.ar/clacso/se/20150722110704/HechoEnLatinoamerica.pdf>
- AmeliCA (2020). *Acerca de AmeliCA*. Recuperado de <http://amelica.org/index.php/que-es-ameli/>
- Beigel, M. F. y Salatino, M. (2015). Circuitos segmentados de consagración académica: las revistas de ciencias sociales y humanas en la Argentina. *Información, cultura y sociedad*, 32, 11-36. <https://doi.org/10.34096/ics.i32.1342>
- Bakkalbasi, N., Bauer, K., Glover, J. y Wang, L. (2006). Three options for citation tracking: Google Scholar, Scopus and Web of Science. *Biomedical digital libraries*, 3(1). <https://doi.org/10.1186/1742-5581-3-7>
- De Filippo, D., Sanz-Casado, E., Urbano Salido, C., Ardanuy, J. y Gómez-Caridad, I. (2011). El papel de las bases de datos institucionales en el análisis de la actividad científica de las universidades. *Revista española de documentación científica*, 34(2), 165-189. <https://doi.org/10.3989/redc.2011.2.797>
- De La Laurencie, A. y Maddi, A. (2019). The dynamics of French publications in social sciences and humanities: a European comparison. In Sapienza University of Rome (Ed.). *17th International Conference on Scientometrics and Informetrics*. Efesto, Italy.
- Delgado, E. y Repiso, R. (2013). El impacto de las revistas de comunicación: comparando Google Scholar Metrics, Web of Science y Scopus. *Comunicar*, 21(41), 45-52. <http://dx.doi.org/10.3916/C41-2013-04>
- Escalona, I., Cerdá Hernández, M. P. y Pulgarín, A. (2010). Web of Science vs. Scopus: un estudio cuantitativo en ingeniería química. *Anales de documentación*, 13, 159-175.
- Fingerman, S. (2005). Web of Science and Scopus: current features and capabilities. *Issues in science and technology librarianship*, 48.
- Gireesh Kumar, T.K. (2013). Comparative analysis of search features of Scopus and Web of Science. In *National Conference on Information Products and Services in the E- environment (NACINPROSE 2013)*, Hyderabad, India. Retrieved from <http://eprints.rclis.org/19731/1/2013%20Gireesh-Scopus.pdf>
- Gorraiz, J., y Schloegl, C. (2007). A bibliometric analysis of pharmacology and pharmacy journals: Scopus versus Web of Science. *Journal of information science*, 10(20), 1-11.
- Harzing, A. W. (2007). *Publish or perish*. Retrieved from <https://harzing.com/resources/publish-or-perish>
- Gusenbauer, M. (2018). Google Scholar to overshadow them all? Comparing the sizes of 12 academic search engines and bibliographic databases. *Scientometrics*, 118, 177-214. <https://doi.org/10.1007/s11192-018-2958-5>
- Jacson, P. (2005). As we may search—Comparison of major features of the Web of Science, Scopus, and Google Scholar citation-based and citation-enhanced databases. *Current science*, 89(9), 1537-1547.
- Jiménez Noblejas, C., y Perianes Rodríguez, A. (2014) Recuperación y visualización de información en Web of Science y Scopus: una aproximación práctica. *Investigación bibliotecológica*, 28(64), 15-31. Recuperado de <http://www.scielo.org.mx/pdf/ib/v28n64/v28n64a2.pdf>

- Kulczycki, E., Engels, T. C. E., Pölönen, J., Bruun, K., Duskova, M., Guns, R., y Zuccala, A. (2018). Publication patterns in the social sciences and humanities: evidence from eight European countries. *Scientometrics*, 116, 463-486. <https://doi.org/10.1007/s11192-018-2711-0>
- LaGuardia, C. (2010). E-Views and Reviews: Scopus vs. Web of Science. *Library journal*. Recuperado de <https://www.libraryjournal.com/?detailStory=e-views-and-reviews-scopus-vs-web-of-science>
- Martín Martín, A., Orduña Malea, E., Thelwall, M. y López Cózar, E. (2018). Google Scholar, Web of Science, and Scopus: A systematic comparison of citations in 252 subject categories. *Journal of informetrics*, 12(4), 1160-1177. <https://doi.org/10.1016/j.joi.2018.09.002>
- Meho, L. I., y Yang, K. (2007). Impact of data sources on citation Counts and rankings of LIS Faculty: Web of Science vs. Scopus and Google Scholar. *Journal of the American Society for Information Science and Technology*. Retrieved from <https://onlinelibrary.wiley.com/doi/epdf/10.1002/asi.20677>
- Michán Aguirre, L., Calderón Rojas, R., Nitxin Castañeda Sortibrán, A. y Rodríguez Arnáiz, R. (2014). Aplicaciones web para recuperación y análisis de literatura de PubMed. *El profesional de la información*, 19(3), 285-291. Recuperado de <http://www.elprofesionaldelainformacion.com/contenidos/2010/mayo/08.pdf>
- Miguel, S. E., de Moya Anegón, F. y Herrero Solana, V. (2006) Aproximación metodológica para la identificación del perfil y patrones de colaboración de dominios científicos universitarios. *Revista española de documentación científica*, 29(1), 36-55. Recuperado de <http://sedici.unlp.edu.ar/handle/10915/89953>
- Mingers, J. and Lipitakis, E. (2010). Counting the citations: a comparison of Web of Science and Google Scholar in the field of business and management. *Scientometrics*, 85(2), 613-625. <https://doi.org/10.1007/s11192-010-0270-0>
- Mongeon, P. and Paul-Hus, A. (2016). The journal coverage of Web of Science and Scopus: a comparative analysis. *Scientometrics*, 106, 213-228. <https://doi.org/10.1007/s11192-015-1765-5>
- Mosbah Natanson, S. and Gingras, Y. (2014). The globalization of social sciences? Evidence from a quantitative analysis of 30 years of production, collaboration and citations in the social sciences (1980–2009). *Current sociology*, 62(5), 626-646. <https://doi.org/10.1177/0011392113498866>
- Orduña Malea, E., Martín Martín, A., Ayllón, J. M., y Delgado López Cózar, E. (2016) *La revolución Google Scholar: destapando la caja de Pandora académica*. Madrid: UNE.
- Puente Lanzarote, L., del Campo Hilario, C. y Ruiz de Luzuriaga Peña, M. (2001). Indicadores de rendimiento para la evaluación de un servicio de bases de datos en línea. *Scire*, 7(1), 89-114. Recuperado de <http://eprints.rclis.org/14991/1/Puente-Indicadores.pdf>
- Rafols, I., Chavarro, D. and Ciarli, T. (2016). Under-representation of research in the global south. Biases in mainstream journal indexing systems. En International Research Conference on Scientometrics, STI Policy and Science Communication, Stellenbosch, Sudáfrica. Retrieved from <http://hdl.handle.net/10261/162452>
- Somoza, M., Guallar, J., Rodríguez Gairín, J. M. y Abadal, E. (2017). Presencia de revistas españolas en bases de datos internacionales. En *Revistas científicas: situación actual y retos de futuro* (pp. 161-178). Barcelona: Universitat de Barcelona.
- Torres Salinas, D., Jiménez Contreras, E. and Delgado López Cózar, E. (2009). Rankings for departments and researchers within a university using two different databases: Web of Science versus SCOPUS. *Scientometrics*, 80, 761–774. <https://doi.org/10.1007/s11192-008-2113-9>
- Torres Salinas, D., Ruiz Pérez, R. y Delgado López Cózar, E. (2009). Google Scholar como herramienta para la evaluación científica. *El profesional de la información*, 18(5), 501-510.
- Worwell, I. (2001). Informetría: explorando bases de datos como instrumentos de análisis. *ACIMED*, 9. Recuperado de <http://eprints.rclis.org/5174/1/sci17100.pdf>

TABLA 1
Análisis formal de las BBDD.

BBDD	Dimensiones de análisis formal								
	Origen	Propósito	Alcance geográfico	Cobertura idiomática	Cobertura temporal	Cobertura temática	División del conocimiento	Tipo de acceso	Fuentes cosechadas
<i>BASE</i>	Institucional	Científico, educativo, no lucrativo	Mundial	Multilingüe	No indica	Ingeniería	Clasificación Decimal de Dewey	Abierto, por suscripción gratuita	8.234 revistas
<i>Dialnet</i>	Institucional	Científico, educativo, divulgativo, no lucrativo	Iberoamérica	Multilingüe	No indica	Multidisciplinar	Ciencias Básicas y Experimentales; Geociencias; Medioambiente; Ciencias Biológicas; Ciencias de la Salud; Agricultura y Alimentación; Tecnologías; Economía y Empresa; Ciencias Jurídicas; Ciencias Sociales; Psicología y Educación; Humanidades; Arte; Filologías	Abierto, por suscripción gratuita	10.836 revistas
<i>Dimensions</i>	Empresarial	Científico, educativo, lucrativo	Mundial	Inglés	1665 al presente	Multidisciplinar	Mathematical Sciences; Physical Sciences; Chemical Sciences; Earth Sciences; Environmental Sciences; Biological Sciences; Agricultural & Veterinary Sciences; Information & Computing Sciences; Engineering; Technology; Medical & Health Sciences; Built Environment & Design; Education, Commerce, Management, Tourism	Abierto, por suscripción gratuita y paga individual e institucional	No indica

Continuación de la Tabla 1.

BBDD	Dimensiones de análisis formal								
	Origen	Propósito	Alcance geográfico	Cobertura idiomática	Cobertura temporal	Cobertura temática	División del conocimiento	Tipo de acceso	Fuentes cosechadas
							& Services; Studies in Human Society; Psychology & Cognitive Sciences; Law & Legal Studies; Studies in Creative Arts & Writing; Language, Communication & Culture; History & Archaeology; Philosophy & Religious Studies		
<i>DOAJ</i>	Institucional	Científico, educativo, no lucrativo	Mundial	Multilingüe	No indica	Multidisciplinar	Clasificación de la Biblioteca del Congreso de Estado Unidos	Abierto, por suscripción gratuita	15.017 revistas
<i>EBSCOhost</i>	Empresarial	Científico, educativo, lucrativo	Mundial	Inglés	No indica	Multidisciplinar	Science; Social Sciences; Humanities; Interdisciplinary	Por suscripción paga individual e institucional	3.800 revistas
<i>Google Scholar</i>	Empresarial	Científico, educativo, no lucrativo	Mundial	Multilingüe	No indica	Multidisciplinar	Economics & Management; Chemical & Material Sciences; Engineering & Computer Science; Health & Medical Sciences; Humanities, Literature & Arts; Life Sciences & Earth Sciences; Physics & Mathematics; Social Sciences	Abierto	No indica
<i>JournalTOCs</i>	Institucional	Científico, educativo, divulgativo	Mundial	Multilingüe	No indica	Multidisciplinar	Medical Sciences; Business & Economics; Biology; Engineering;	Abierto, por suscripción gratuita	35.903 revistas

Continuación de la Tabla 1.

BBDD	Dimensiones de análisis formal								
	Origen	Propósito	Alcance geográfico	Cobertura idiomática	Cobertura temporal	Cobertura temática	División del conocimiento	Tipo de acceso	Fuentes cosechadas
		no lucrativo					Education; Language & Literature; Computer Science; History; Law; Health & Safety; Political Science; Mathematics; Agriculture; Psychology; Humanities	individual y paga institucional	
<i>JSTOR</i>	Institucional, no gubernamental	Científico, educativo, no lucrativo	Mundial	Multilingüe	1870 al presente	Multidisciplinar	Area Studies; Arts; Business & Economics; History; Humanities; Law; Medicine & Allied Health; Science & Mathematics; Social Sciences	Abierto, por suscripción gratuita y paga individual	2.600 revistas
<i>LILACS</i>	Institucional, gubernamental	Científico, educativo, no lucrativo	América Latina y el Caribe	Español, portugués, francés, inglés	1982 al presente	Salud	Tesauro DeCS	Abierto	895 revistas
<i>OpenAire</i>	Gubernamental	Científico, educativo, divulgativo, no lucrativo	Europa	Multilingüe	No indica	Multidisciplinar	No indica	Abierto, por suscripción gratuita	21 bases de datos
<i>PubMed</i>	Gubernamental	Científico, educativo, no lucrativo	Mundial	Multilingüe	1809 al presente	Salud	Tesauro MeSH	Abierto, por suscripción gratuita	2 bases de datos
<i>RedALyC</i>	Institucional	Científico, educativo, no lucrativo	Mundial	Multilingüe	No indica	Multidisciplinar	Ciencias Sociales; Ciencias Naturales y Exactas; Arte y Humanidades	Abierto	1.368 revistas
<i>RedIB</i>	Institucional, gubernamental	Científico, educativo, divulgativo, no lucrativo	Iberoamérica	Multilingüe	No indica	Multidisciplinar	Ciencias de la Salud; Ciencias de la Vida; Ciencias Físicas e Ingeniería; Ciencias	Abierto, por suscripción gratuita	3.385 revistas

Continuación de la Tabla 1.

BBDD	Dimensiones de análisis formal								
	Origen	Propósito	Alcance geográfico	Cobertura idiomática	Cobertura temporal	Cobertura temática	División del conocimiento	Tipo de acceso	Fuentes cosechadas
							Sociales y Humanidades		
<i>SciELO</i>	Institucional	Científico, educativo, no lucrativo	Iberoamérica, el Caribe, Sudáfrica	Español, portugués, inglés	No indica	Multidisciplinar	Ciencias Agrarias; Ciencias Biológicas; Ciencias de la Salud; Ciencias Exactas y de la Tierra; Ciencias Humanas; Ciencias Sociales; Ciencias Aplicadas; Ingeniería; Lingüística y Letras; Artes; Psicoanálisis	Abierto	1.767 revistas
<i>ScienceDirect</i>	Empresarial	Científico, educativo, lucrativo	Mundial	Inglés	No indica	Multidisciplinar	Ciencias Físicas e Ingeniería; Ciencias de la Vida; Ciencias de la Salud; Ciencias Sociales y Humanidades	Abierto, por suscripción paga individual	4.253 revistas
<i>Scopus</i>	Empresarial	Científico, educativo, lucrativo	Mundial	Multilingüe	1960 al presente	Multidisciplinar	Ciencias de la Salud; Ciencias de la Vida; Ciencias Físicas; Ciencias Sociales	Por suscripción paga individual e institucional	31.972 revistas activas
<i>WOS</i>	Empresarial	Científico, educativo, lucrativo	Mundial	Multilingüe	1900 al presente	Multidisciplinar	Ciencias; Ciencias Sociales; Artes y Humanidades	Por suscripción paga individual e institucional	24.779 revistas

TABLA 2.
Análisis funcional de las BBDD.

BBDD	Dimensiones de análisis funcional							
	Interfaz de búsqueda	Presentación de resultados	Exhaustividad de los registros	Formatos de exportación	Capacidad de exportación	Campos para la exportación	Modalidad de exportación	Calidad del dato
BASE	Muy completa	Muy completa	Título, autor/es, resumen, editor, fecha, tipo de documento, proveedor de contenido, palabras clave, idioma, enlace, tipo de acceso	RefWorks, EndNote, RIS, BibTex, MARC, RDF, RTF, JSON, YAML	10 registros por selección individual	Tipo de documento, título, autor/es, fecha, editor, resumen, palabras clave, idioma	Abierta	Media
<i>Dialnet</i>	Incompleta	Completa	Título, autor/es, fuente, ISSN/ISBN, volumen y número, fecha, páginas, idioma, filiación institucional, resumen	No posee	Nula	No posee	No posee	No fue posible evaluar esta dimensión
<i>Dimensions</i>	Muy completa	Muy completa	Título, autor/es, fecha, fuente, tipo de documento, resumen, tipo de acceso, métricas	BibTex, RIS, XLS, CSV	500 registros por selección global	Autor/es, identificador permanente, fuente, palabras clave, volumen y número, páginas, título, enlace, fecha	Por suscripción paga	Alta
<i>DOAJ</i>	Completa	Completa	Título, autor/es, fuente, volumen y número, fecha, identificador persistente, resumen, editor, ISSN, filiación, categorías temáticas, idioma, país, formato/s del texto completo	CSV, JSON	limitada, colección completa sin posibilidad de selección	Título, autor/es, fuente, volumen y número, fecha, identificador persistente, editor, ISSN, categorías temáticas, idioma	Abierta	Media
<i>EBSCOhost</i>	Muy completa	Muy completa	Fuente, título, autor/es, resumen, palabras clave, páginas, fecha, tipo de acceso, enlace	BibTex, CiteULike, EndNote, Mendeley, Zotero, RIS, XML, MARC21, EasyBib	limitada por selección global	Fuente, editor, título, autor/es, filiación, resumen, palabras clave, páginas, fecha, ISSN, idioma	Por suscripción paga	Alta
<i>Google Scholar</i>	Incompleta	Incompleta	Título, autor/es, fuente, fecha, editor, resumen,	BibTex, EndNote, RefMan, CSV	limitada por selección global	Autor/es, título, fuente, volumen y número, páginas, fecha, editor	Por suscripción gratuita	Media

Continuación de la Tabla 2.

BBDD	Dimensiones de análisis funcional							
	Interfaz de búsqueda	Presentación de resultados	Exhaustividad de los registros	Formatos de exportación	Capacidad de exportación	Campos para la exportación	Modalidad de exportación	Calidad del dato
			métricas, documentos relacionados					
<i>JournalTOCs</i>	Muy completa	Muy completa	Título, autor/es, fecha, fuente, volumen y número, páginas, ISSN, editor	RIS	Ilimitada por selección individual	No fue posible evaluar esta dimensión	Por suscripción gratuita	Baja
<i>JSTOR</i>	Muy completa	Muy completa	Tipo de documento, título, autor/es, editor, fecha, tipo de acceso, resumen	RIS, TXT, RefWorks, NoodleTools, EasyBib	Ilimitada por selección global en la versión paga, por selección individual en la versión gratuita	ISSN/ISBN, URL, resumen, autor/es, fuente, número y volumen, páginas, editor, título, fecha	Abierta, por suscripción paga	Media
<i>LILACS</i>	Completa	Muy completa	Título, autor/es, fuente, número y volumen, fecha, tipo de documento, idioma, base de datos, ID, resumen, palabras clave	RIS, CSV, TXT	Ilimitada por selección global	Autor/es, título, fuente, número y volumen, fecha	Abierta	Media
<i>OpenAire</i>	Completa	Completa	Título, tipo de documento, idioma, tipo de acceso, autor/es, fecha, fuente, editor, identificador persistente, palabras clave, resumen, referencias, métricas	CSV	2.000 registros sin posibilidad de selección	Título, autor/es, fecha, identificador persistente, nombre del proyecto de investigación y patrocinador, tipo de acceso	Abierta	Media
<i>PubMed</i>	Muy completa	Muy completa	Título, autor/es, filiación, fuente, fecha, identificador persistente, tipo de acceso, tipo de documento, idioma, resumen, palabras	CSV, TXT	10.000 registros por selección global	ID, título, autor/es, cita, primer autor, fuente, fecha, identificador persistente	Abierta	Alta

Continuación de la Tabla 2.

BBDD	Dimensiones de análisis funcional							
	Interfaz de búsqueda	Presentación de resultados	Exhaustividad de los registros	Formatos de exportación	Capacidad de exportación	Campos para la exportación	Modalidad de exportación	Calidad del dato
			clave, referencias, citas, resultados similares					
<i>RedALyC</i>	Incompleta	Incompleta	Título, autor/es, fuente, fecha, resumen, tipo de acceso	RIS, RefWorks, BibTex	1 registro por selección individual	Tipo de fuente, ID, título, autor/es, palabras clave, fuente, idioma, fecha, enlace	Abierta	Baja
<i>RedIB</i>	Muy completa	Completa	Título, título traducido, autor/es, fuente, fecha, tipo de documento, idioma, enlace, identificador persistente, editor, derechos de uso, palabras clave, resumen, resumen traducido, referencias	RIS, RefWorks, EndNote, Mendeley	Ilimitada por selección individual	Título, fuente, número y volumen, autor, editor, fecha, idioma, enlace, tipo de documento	Por suscripción gratuita	Alta
<i>SciELO</i>	Muy completa	Muy completa	Título, autor/es, fuente, fecha, volumen y número, páginas, resumen, identificador permanente, métricas	RIS, BibTex, CITA, CSV	2.000 registros por selección global	ID, título, autor/es, fuente, idioma, fecha, enlace	Abierta	Alta
<i>ScienceDirect</i>	Muy completa	Completa	Tipo de documento, tipo de acceso, título, fuente, volumen y número, fecha, autor/es, filiación, palabras clave, identificador persistente, licencia de uso, resumen, métricas	RefWorks, Mendeley, RIS, BibTex, TXT	1 registro por selección individual	Autor/es, título, fuente, fecha, ISSN, identificador persistente, enlace, resumen, palabras clave	Abierta	Baja
<i>Scopus</i>	Muy completa	Muy completa	Título, autor/es, filiación, fecha, fuente, número y volumen, páginas, enlace, resumen, palabras clave, referencias, métricas	Mendeley, RefWorks, RIS, CSV, BibTex, HTML	2.000 registros por selección global, 20.000 registros en formato CSV	Autor/es, ID, filiación, título, año, fuente, número y volumen, páginas, tipo de fuente, tipo de documento, identificador	Por suscripción paga	Media

Continuación de la Tabla 2.

BBDD	Dimensiones de análisis funcional							
	Interfaz de búsqueda	Presentación de resultados	Exhaustividad de los registros	Formatos de exportación	Capacidad de exportación	Campos para la exportación	Modalidad de exportación	Calidad del dato
						persistente, tipo de acceso, ISSN, editor, idioma, resumen, palabras clave, patrocinador, financiación, referencias, métricas		
WOS	Muy completa	Muy completa	Título, autor/es, filiación, fuente, número y volumen, páginas, fecha, identificador persistente, tipo de documento, palabras clave, resumen, editor, enlace, métricas	Endnote, RefWorks, inCites, CSV, HTML, FECYT CVN, BibText, TXT	500 registros por selección global, 5.000 registros para cuentas personales	Autor/es, ID, filiación, dirección, editor, resumen, palabras clave, título, fuente, ISSN, idioma, financiación, identificador persistente, referencias, patrocinador, métricas	Por suscripción paga	Media

NOTAS

- 1 <http://amela.org/>
- 2 <https://scholar.google.com/>
- 3 <https://app.dimensions.ai/>
- 4 Por suscripción gratuita se considera al requerimiento de algunos servicios de búsqueda de la generación sin costo de un perfil de usuario para el uso del sistema.