

La ciencia de la ciencia

*Eduardo Alvarez López y
Layla Michán Aguirre*

Resumen

La cienciometría consiste en emplear métodos, herramientas y técnicas cuantitativas basados en el análisis de la información y producción científica, sobre todo de los artículos académicos. Su propósito es analizar, evaluar y visualizar la información para obtener proporciones, tendencias, patrones, relaciones e indicadores. Las aplicaciones de la cienciometría van desde la administración de recursos bibliográficos; la recuperación de información; el mantenimiento y curación de colecciones; la evaluación, diagnóstico, y gestión de la política científica, el estado del arte y revisiones; el análisis, desarrollo, estructura, evolución y relaciones de la dinámica científica y los mapas de ciencia; hasta la obtención de nuevo conocimiento. Debido a la creciente generación de información científica, la cienciometría es un análisis que se realiza frecuentemente; este artículo presenta una introducción general a la cienciometría respecto a sus características, aplicaciones, métodos y ejemplos.

Palabras clave: bases de datos, literatura científica, indicadores, metadatos, revistas.

THE SCIENCE OF SCIENCE

Abstract

Scientometrics consists of employing methods, tools and quantitative techniques based on the analysis of scientific information and production, especially academic articles. Its purpose is to analyze, evaluate and visualize information to obtain proportions, trends, patterns, relationships and indicators. The applications of scientometrics range from administration of bibliographical resources; recovery of information, maintenance and restoration of collections; evaluation, diagnosis, and management of scientific policy, state of the art and reviews; from analysis, development, structure, evolution and relations of scientific dynamics and science maps; to obtaining new knowledge. Due to the growing generation of scientific information, scientometrics is an analysis that is carried out frequently. This article presents a general introduction to scientometrics with respect to its characteristics, applications, methods and examples.

Key words: databases, scientific literature, indicators, metadata, journals.

DOI: <http://doi.org/10.22201/codeic.16076079e.2018.v19n4.a1>



Eduardo Álvarez López

bioeduardo@ciencias.unam.mx

Maestro en Ciencias. Realiza trabajos de consultoría en innovación científica, producción científica, innovación editorial, biotecnología, producción de fármacos y estudio de animales venenosos con fines terapéuticos. Facultad de Ciencias. UNAM.

ORCID iD: <http://orcid.org/0000-0002-1572-164X>

Layla Michán Aguirre

laylamichan@ciencias.unam.mx

Doctora en Ciencias Biológicas por la UNAM, realiza investigación, docencia, difusión e innovación en bases de datos y sistemas de organización de conocimiento en biología, especialista en manejo de información en biología y medicina, así como en la solución de problemas basados en información en ciencias biológicas y médicas.

ORCID iD: <http://orcid.org/0000-0002-5798-662X>

Introducción

La cienciometría consiste en emplear métodos, herramientas y técnicas métricas para investigar el desarrollo, la estructura, la dinámica y la evolución de la ciencia basada en el análisis de la información y la producción científica, en especial los artículos académicos (Leydesdorff, 2001). Su propósito es analizar, evaluar y visualizar la información para obtener proporciones, tendencias, patrones y relaciones. Su origen y desarrollo han sido el resultado de la interrelación entre las ciencias de la información, la política científica, las bases de datos, en especial las bibliográficas, y el análisis cuantitativo. Se le conoce también como estudios métricos de la información científica o estudios cuantitativos de la ciencia.

Durante la primera mitad del siglo XX los acontecimientos políticos y sociales marcaron de manera distintiva la práctica científica, aparecieron instancias nacionales y regionales encargadas de organizar a los científicos, distribuir los fondos para ciencia y tecnología, así como evaluar los resultados de la práctica científica; para lo cual fue necesario integrar la información que generaban y establecer pautas de comparación y evaluación. Simultáneamente aparecían nuevas revistas científicas lo que generó una cantidad de información inmensa que era imposible de procesar por métodos tradicionales. Fue a mediados del siglo XX cuando inició la revolución digital, época en la que se generaron bases de datos electrónicas que fueron sustituyendo a los índices y catálogos bibliográficos impresos, esto facilitó la obtención de proporciones, indicadores y tendencias en la publicación científica. Actualmente [CrossRef](#) es el catálogo bibliográfico formal

“
La cienciometría implica el manejo de cantidades de datos inmensas que permiten el análisis multidimensional de los documentos y la aplicación de análisis estadísticos e indicadores...
”

más extenso, contiene más de 92 millones de documentos científicos, principalmente artículos (CrossRef, 2017).

Es en este contexto en el que se publicó el primer libro con gran impacto escrito por Derek J. de Solla Price [Big Science, Little Science](#) (1963) en el que se presentan análisis métricos de la ciencia de manera analítica y con visualizaciones gráficas de las tendencias y evolución de la ciencia; así como Citation Index of Genetics editada por Garfield y Sher en 1964, la primera base de datos bibliográfica con más de 400 000 citas que provienen de más de 100 000 artículos publicados en 1961. Para 1978 se publicó el primer número de la revista [Scientometrics](#), con esto dio inicio la institucionalización de la disciplinas y para 1993 se fundó la [International Society for Informetrics and Scientometrics](#) (ISSI).

Es así como la cienciometría ha tenido un auge importante durante las últimas décadas, debido a la digitalización y sistematización de la información en bases de datos o colecciones digitales; de tal manera que se han desarrollado nuevas técnicas analíticas, tecnologías de acceso y herramientas de organización del conocimiento como tesauros y ontologías para explorar y organizar las colecciones digitales de manera innovadora (Van Noorden, 2010). Se diseñan a diario nuevas herramientas para realizar búsquedas más eficientes y precisas, así como para hacer análisis mejores y más extensos, ha aumentado la eficacia y la exactitud en la obtención de información y se han generado nuevas oportunidades al incrementarse el espectro de opciones técnicas como los *datos ligados* (*linked data*), la *web semántica* (*semantic web*), así como análisis de redes sociales, la *minería de textos* (*text mining, literature based discovery o knowledge discovery database*) y los *grandes datos* (*big data*). Este último análisis se utiliza para hacer referencia a la generación de nuevo conocimiento interesante, plausible e inteligente a partir de vincular dos o más conceptos de la literatura que hasta ahora no habían sido relacionados (es decir, disjuntos) por medio del uso de programas y algoritmos diseñados para tal fin, con un sustento empírico, cuantificable y comparativo, lo cual aporta elementos interesantes para el análisis y la interpretación de la información científica. Además, recientemente ha surgido la creación y estudio de nuevas mediciones basadas en el análisis de las redes sociales y en información académica usando los portales web de las revistas, bibliotecas digitales, *tweets*, discusiones de blogs, notas, redes de investigadores, entre otras, que se conoce como *Altmetría* (Sugimoto *et al.*, 2017), acrónimo de métricas alternativas (*Alternative Metrics o Altmetrics*).¹

En todas las revistas científicas más prestigiadas aparecen constantemente artículos sobre el tema e incluso se dedican secciones como en [Nature](#). Cuando uno lee un artículo que contiene análisis cienciométricos puede identificar

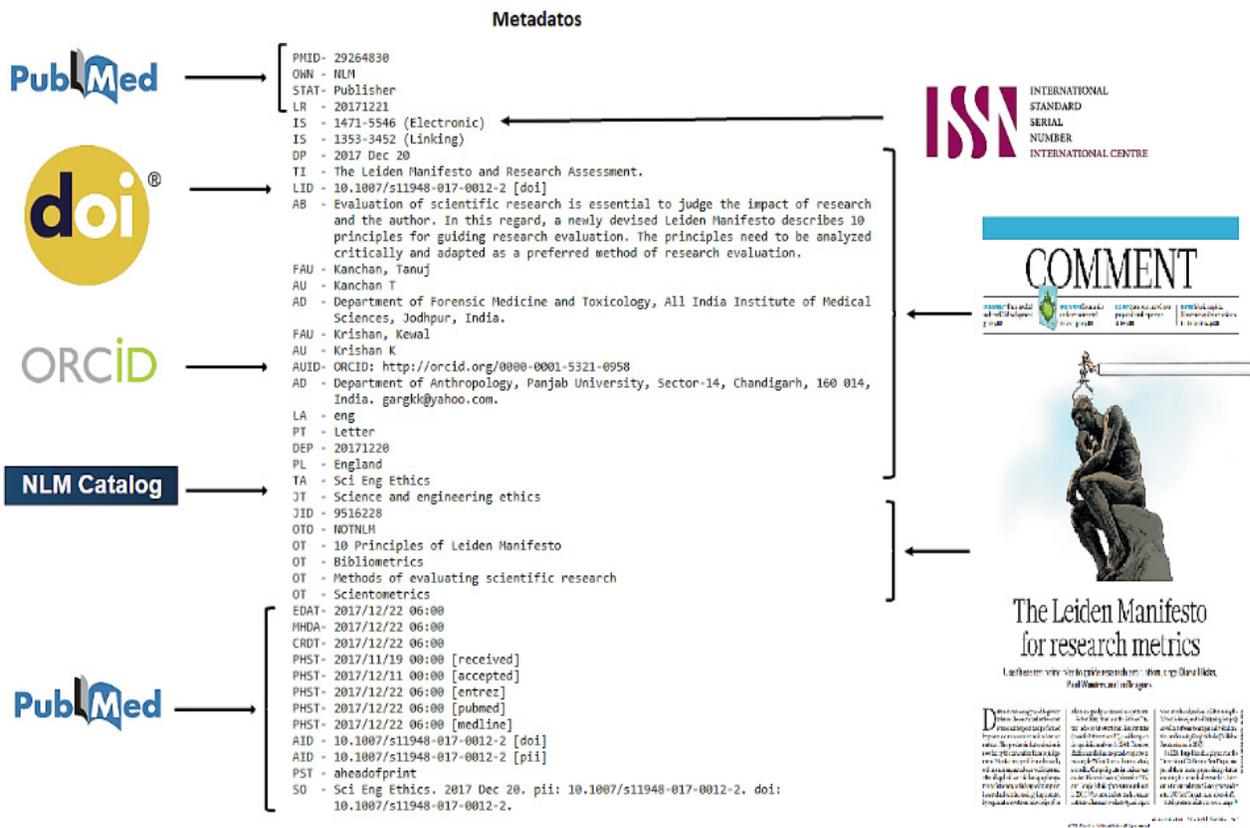
1. Detalles sobre este tema se pueden consultar en el portal <https://www.altmetric.com/> y ejemplos de artículos que contienen este tipo de análisis se pueden leer en <http://collections.plos.org/altmetrics>.

elementos como las preguntas a resolver, la consulta realizada, la colección utilizada, la cantidad de registros procesados, los indicadores utilizados, el objeto de estudio, el propósito y las gráficas que, muchas veces, integran información interesante.

Para realizar los análisis cuantitativos se utiliza información sobre los documentos, o *metadatos*, esto es, datos que permiten describir las características del artículo, capítulo, libro o memoria con detalle en título, tipo de documento, idioma, palabra clave, volumen, número de páginas, año de publicación, resumen, el o los nombres de los autores, institución de inscripción, país y correo electrónico de contacto (figura 1). Otras colecciones además agregan las referencias, las citas de los documentos y asignan descriptores con base en sistemas de organización del conocimiento, como tesauros y ontologías. Algunos tesauros utilizados en las bases de datos bibliográficas académicas son [UNESCO](#), [MESH](#), [IEEE](#), [CAB Thesaurus](#), entre muchos otros.

Figura 1.

Datos de un artículo indexado en PubMed.



Entre los propósitos que tienen los análisis bibliométricos se pueden identificar ocho: la administración de recursos bibliográficos; la recuperación de información; el mantenimiento y curación de colecciones; la evaluación, diagnóstico, gestión y política científica, como lo hace la [OECD](#) y algunos [rankings](#); el estado del arte y revisiones; análisis, desarrollo, estructura y relaciones de la dinámica científica, como los [mapas de ciencia](#); y la obtención de nuevo conocimiento.

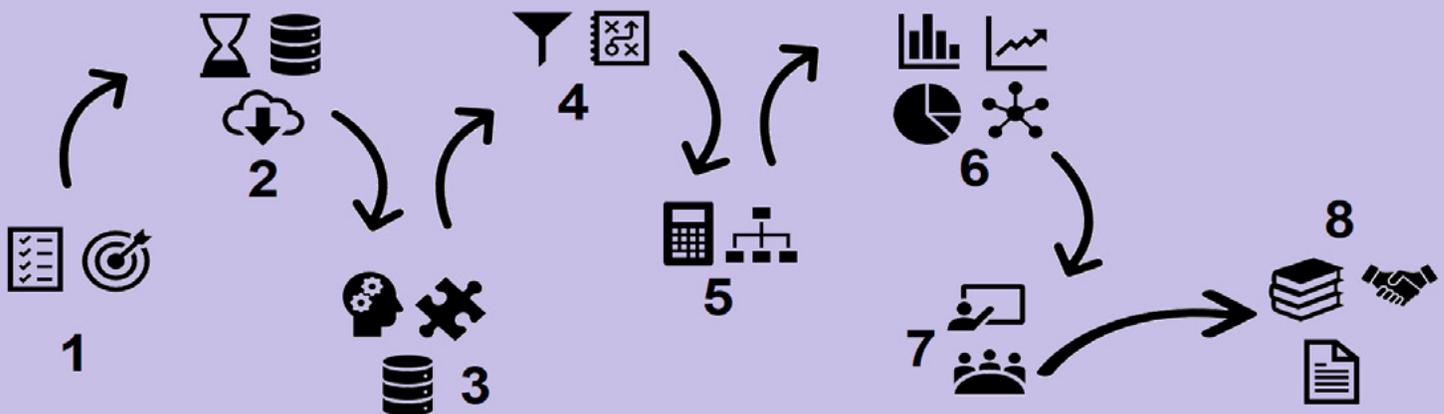
El método cuenciométrico

La cuenciometría implica el manejo de cantidades de datos inmensas (miles y millones de bits) que permiten el análisis multidimensional de los documentos y la aplicación de análisis estadísticos e indicadores, para reconocer las regularidades y tendencias de la producción y la difusión de la ciencia, la tecnología y la innovación con respecto a diversas variables y con las distintas unidades de análisis. Este procedimiento generalmente consiste de ocho etapas (figura 2):

1. **Planeación y diagnóstico**
2. **Recuperación de información.** Esta etapa es crucial porque el resultado depende de la base de datos utilizada, todas son diferentes y tienen características distintivas como la cantidad de registros, las disciplinas que indexa, los criterios de selección de las revistas y los metadatos que presentan
3. **Construcción de la colección.**
4. **Curación.**
5. **Análisis y validación.**
6. **Visualización.**
7. **Interpretación.**
8. **Publicación y difusión del o los productos.**

Respecto a los problemas más comunes para realizar un estudio cuenciométrico se encuentran el costo y la restricción de las bases de datos, la dificultad para extraer los metadatos y la normalización deficiente de los metadatos en las colecciones.

Figura 2.
Protocolo de un análisis cuenciométrico.



Los indicadores cientimétricos

Los indicadores cientimétricos son estándares empleados para medir, cuantificar y evaluar el rendimiento científico de un individuo, grupo, institución, nación o región. Su uso ha permitido tener valores específicos y más objetivos que facilitan la comparación y la toma de decisiones, pero están muy lejos todavía de reflejar el complejo y heterogéneo proceso de la investigación científica.

Estos indicadores pueden clasificarse de entrada como los que mide la OCDE y la UNESCO basados en los insumos de la ciencia (*input*) y los indicadores de salida (*output*); estos últimos son los que comúnmente se utilizan para evaluación científica y son de varios tipos, algunos ejemplos están en la figura 3. Por ejemplo, la producción y difusión del conocimiento científico se puede caracterizar con base en el estudio del impacto de las revistas por medio del factor de impacto que se obtiene de la relación entre la cantidad de artículos publicados y las citas que obtuvieron, o de los artículos académicos a través de la cantidad de citas que reflejan las relaciones entre los grupos de investigación y permiten identificar núcleos significativos de autores y revistas (Pudovkin y Garfield, 2004). El *índice H* representa la relación de las publicaciones realizadas y citadas de un autor o una revista. La cocitación ayuda a determinar la asociación de los investigadores (Trujillo y Long, 2018). Los contenidos, los títulos, los resúmenes, las palabras clave o los descriptores representan el contenido teórico y metodológico de los documentos.



Figura 3.
Indicadores cientimétricos
más usados.

El primer servicio en publicar indicadores cientimétricos para las revistas fue el *Journal Citation Reports* (JCR) de ISI (ahora a cargo de la compañía Clarivate Analytics) que se publica anualmente desde los años setenta, donde se definieron estos indicadores por primera vez para uso público y fueron los que se usaron como referencia para la evaluación científica nacional; entre otras aplicaciones se basan en el artículo, las referencias o las citas a éste (figura 4).

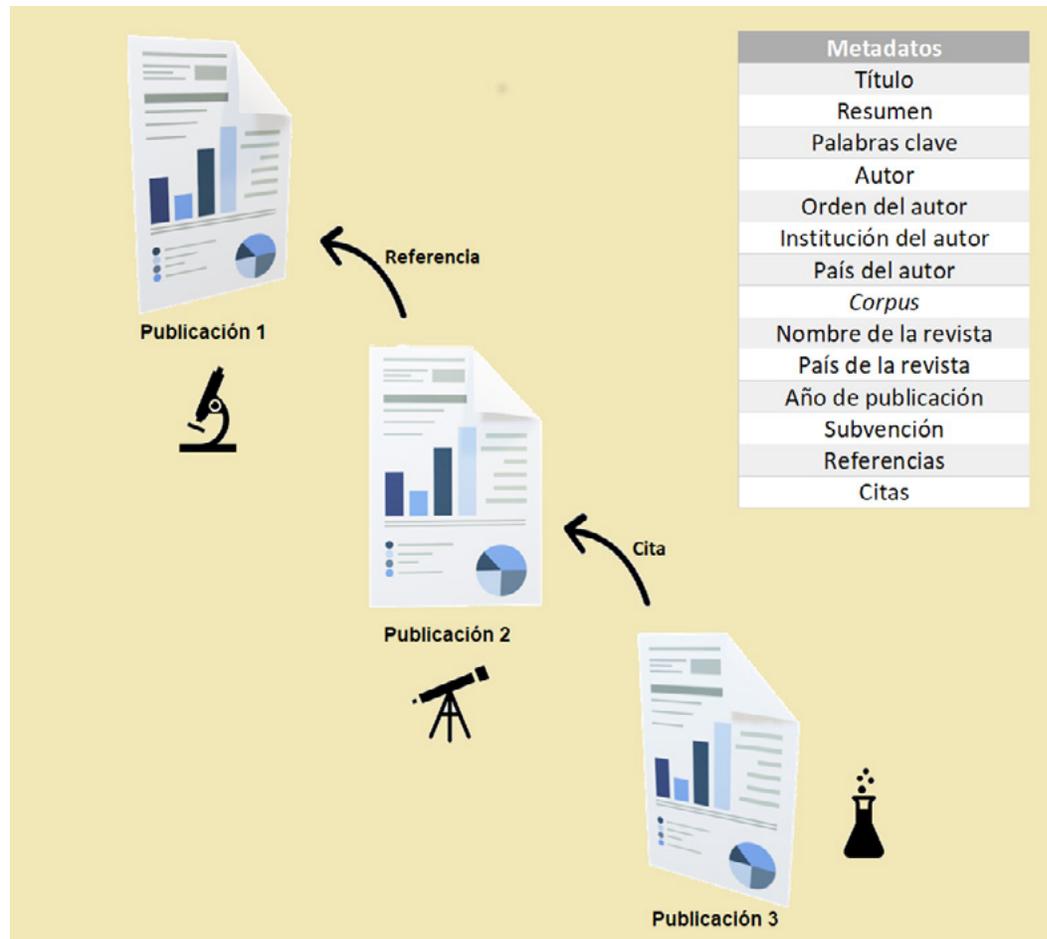


Figura 4.
Algunos elementos de análisis
cientimétricos de los
artículos científicos.

Ahora están disponibles también los indicadores de *Google Académico*, *Eigen-Factor*, *Scopus*, *Scimago*, y para América Latina *Biblat* y *SciELO*. Entre los más utilizados están el total de documentos publicados, la tendencia temporal y la proporción de artículos por revista, temas, autores, instituciones, países, tipo de documento, idioma y descriptores. Entre los indicadores más usados son: índice H, factor de impacto, cantidad de citas, vida media, obsolescencia, índice de instantaneidad o inmediatez (Yue *et al.*, 2007). Asimismo, es común ver en las revistas electrónicas [cifras sobre la cantidad de documentos descargados](#), guardados en manejadores de bibliografía como *Mendeley* o publicados en redes sociales. Generalmente, estos indicadores son más consultados en las bases de datos *WoS* y *Scopus*, en la mayoría de los comités de evaluación científica la referencia es *JCR* y *Master Journal List*.

Entre las ventajas del uso de indicadores están que: 1) son datos cuantificables y estadísticos, 2) constituyen estándares que permiten la comparación, 3) proveen una manera sistemática y objetiva de evaluar variables como el impacto, la influencia y el uso. De esta manera, sintetizan en unos pocos parámetros, las características de distintos grupos de datos, así como en la posibilidad que brindan para establecer pronósticos sobre tendencias futuras y determinar el efecto de diferentes factores en variables de interés. Ofrecen una base sólida para la toma de decisiones prácticas. Es un método objetivo. Permite analizar y evaluar el estado de la ciencia y la tecnología en varias naciones, pues se comparan los resultados obtenidos y se elaboran las tendencias en el tiempo. Son instrumento de evaluación para fomentar y/o estudiar la praxis de la ciencia, definir estrategias y políticas de desarrollo científico nacionales e internacionales.

Tabla 2.

Consideraciones de PenDlebury y el Manifiesto de Leiden para la investigación en análisis métricos.

	Consideraciones de PenDlebury para realizar un análisis cuantitativo	Principios de Leiden para desarrollar un análisis cuantitativo
1	Evaluar si los datos disponibles permiten resolver la interrogante.	La evaluación cuantitativa tiene que apoyar la valoración cualitativa por expertos.
2	Seleccionar la(s) colección(es) idónea(s).	El desempeño debe ser medido de acuerdo con las misiones de investigación de la institución, grupo o investigador.
3	Establecer los límites, términos y operadores correctos para la consulta.	La excelencia en investigación de relevancia local debe ser protegida.
4	Decidir entre la cuenta total o fraccionada.	Los procesos de recopilación y análisis de datos deben ser abiertos, transparentes y simples.
5	Juzgar si los datos requieren de edición para eliminar “artefactos” (normalizar).	Los datos y análisis deben estar abiertos a verificación por los evaluados.
6	Comparar semejanzas.	Las diferencias en las prácticas de publicación y citación entre campos científicos deben tenerse en cuenta.
7	Utilizar medidas relativas, no sólo cuentas absolutas.	La evaluación individual de investigadores debe basarse en la valoración cualitativa de su portafolio de investigación.
8	Obtener múltiples medidas.	Debe evitarse la concreción imprecisa y la falsa precisión.
9	Reconocer la naturaleza sesgada de los datos.	Deben reconocerse los efectos sistémicos de la evaluación y los indicadores.
10	Confirmar que los datos obtenidos sean relevantes a la pregunta.	Los indicadores deben ser examinados y actualizados periódicamente.
11	Revisar si los resultados son razonables.	
12	Contextualizar e integrar los datos.	
13	Presentar los resultados de forma abierta y honesta.	

Entre sus desventajas están que para los indicadores no se consideran elementos de carácter epistemológico, geográfico, histórico, económico, social o político; son multifactoriales y, por lo tanto, difíciles de interpretar, dependen del tiempo, el campo científico, la revista, el orden de publicación, la accesibilidad, visibilidad, internacionalidad, impacto, prestigio, etcétera. Este método puede ser anacrónico, por ello es deseable contextualizar los resultados usando la historia, sociología y etnografía, por ejemplo. Muchas de estas desventajas están relacionadas con el abuso, desconocimiento y descontextualización en su uso. Para contrarrestar lo anterior es importante aplicar pertinentemente el análisis; para ello se puede consultar en la tabla 2 las trece consideraciones de Pendlebury (2009) para realizar un análisis cuantitativo y el *Manifiesto de Leiden*, que en sus diez principios establece recomendaciones para el uso adecuado de los análisis métricos (Hicks *et al.*, 2015).

Aunque el uso de la cuantificación es muy común entre los científicos cuando reportan sus informes o realizan evaluaciones en comités dictaminadores, su uso va más allá de ser cifras útiles para la evaluación de los investigadores, los artículos o las revistas. Los invitamos a revisar artículos de su área de interés o a realizar un análisis cuantitativo de su área de especialidad, pues permite tomar mejores decisiones de investigación basadas en la evidencia científica publicada que es fácilmente visualizable a través de la cuantificación.

Referencias

- ❖ CrossRe. (2017). *Status*. Recuperado de: <https://data.crossref.org/reports/status-Report.html>
- ❖ de Solla Price, Derek J. (1963). *Little Science, Big Science*. Nueva York: Columbia University Press.
- ❖ Garfield, E., Sher, I. H. (1964). Genetics Citation Index [from the reviewers]. *Science Fortnightly*, 1 (14), 4.
- ❖ Hicks, Diana, Wouters, Paul, Waltman, Ludo, de Rijcke, Sara y Rafols, Ismael (2015). Bibliometrics: The Leiden Manifesto for research metrics. *Nature*, 520, 429-431.
- ❖ Leydesdorff, L. (2001). *The Challenge of Scientometrics*. Irvine, California: Universal-Publishers.
- ❖ Pendlebury, D. A. (2009). The use and misuse of journal metrics and other citation indicators. *Archivum Immunologiae et Therapiae Experimentalis (Warsz)*, 57 (1), 1-11. DOI: <https://doi.org/10.1007/s00005-009-0008-y>
- ❖ Pudovkin, Alexander I., Garfield, Eugene. (2004). Rank-normalized impact factor: A way to compare journal performance across subject categories. *Proceedings of the Association for Information Science and Technology* 41 (1), 507-515. DOI: <https://doi.org/10.1002/meet.1450410159>

- ❖ Sugimoto, Cassidy. R., Work, Sam, Lariviere, Vincent y Haustein, Stefanie (2017). Scholarly use of social media and altmetrics: A review of the literature. *Journal of Association for Information Science and Technology*, 68 (9), 2037-2062. DOI: <https://doi.org/10.1002/asi.23833>
- ❖ Trujillo, Caleb M., Long, Tammy M. (2018). Document co-citation analysis to enhance transdisciplinary research. *Science Advances*, 4 (1), 1-9. DOI: <https://doi.org/10.1126/sciadv.1701130>
- ❖ Van Noorden, Richard (2010). Metrics: a profusion of measures. *Nature*, 465, 864-866. DOI: <https://doi.org/10.1038/465864a>
- ❖ Yue, W. y W., Concepcion y Rousseau, R. (2007). The immediacy index and the journal impact factor: Two highly correlated derived measures. *Canadian Journal of Information and Library Science*, 28 (1), 33-48. Recuperado de: https://www.researchgate.net/publication/278738130_The_immediacy_index_and_the_journal_impact_factor_Two_highly_correlated_derived_measures

Cómo citar este artículo

- ❖ Álvarez López, Eduardo y Michán Aguirre, Layla (2018). La ciencia de la ciencia. *Revista Digital Universitaria* (RDU). Vol. 19, núm. 3 julio-agosto. DOI: <http://doi.org/10.22201/codeic.16076079e.2018.v19n4.a1>.