

## Trilogía para la visión científica: las publicaciones científicas, las bases de datos y la bibliometría

Yaniris Rodríguez Sanchez

Instituto de Información científica Tecnológica (Cuba)

yrs201181@gmail.com

---

---

### Resumen

Se presenta un estudio que tiene como propósito incidir en la cultura de los modelos de comunicación científica para la evaluación de la ciencia, y su impacto para la construcción de indicadores bibliométricos. Se analizaron la estructura de las bases de datos LILACS y SCOPUS, para el período 2000 al 2008, desde la visión de usuario de información para determinar indicadores bibliométricos. Las principales dificultades radicaron en el campo autor y filiación. Se establecen características dominantes en la formulación de publicaciones científicas que, posteriormente repercuten en la estructura semántica de las bases de datos bibliográficas. Se identifican limitaciones actuales en la relación existente entre: publicaciones científicas, bases de datos y la bibliometría, al respecto se enuncian consideraciones para una mejor explotación de la praxis en la comunicación científica documentaria, como proceso formal más utilizado para la difusión del nuevo conocimiento.

### Palabras clave

Bibliometría; publicaciones científicas; bases de datos

### Abstract

It is presented a study that has as purpose to influence in the culture of science communication models for the assessment of science, and its impact for the construction of bibliometric indicators. It was analyzed the structure of LILACS and SCOPUS databases for the period 2000 to 2008. The main difficulties were in the author and affiliation field. There are established dominant characteristics in the formulation of scientific publications that subsequently, they will affect the semantic structure of the bibliographic databases. It is identified recent limitations on the relationship among scientific publications, databases and Bibliometric. In this regard, they are enunciated considerations for a better exploitation of the practice in the scientific communication documentary as a formal process more used for the dissemination of new knowledge.

### Keywords

Bibliometrics; scientific publication; databases; LILACS; SCOPUS

## **Introducción**

Establecer modelos que se ajusten y reflejen la realidad en la que se manifiesta la sociedad científica es, actualmente, una de las ocupaciones en las que se sumergen los Sistemas Nacionales de Ciencia Tecnología e innovación (SNCTI). Estos precisan de identificar los paradigmas en los que se desenvuelve la generación del nuevo conocimiento, atribuir variables sociales, económicas, políticas, entre otras (UNESCO, 2005), de esta forma se permite cualificar y argumentar la concepción de indicadores científicos, a partir del entorno en el que se construyen. Los indicadores bibliométricos forman parte de estos, ellos esclarecen el dominio de la actividad científica y dependen de la producción documentaria, suponen que las prácticas en este sentido deben estar pertinentemente establecidas, pues de ahí su alcance y limitaciones para incidir en la formulación de políticas científicas coherentes.

Autores como López, Merton y Price (López-Cózar, 1997; Merton, 1957; Price, 1963) manifestaron que el proceso de investigación científica incluye como salida la certificación de los resultados a partir de su publicación, a lo cual esta autora además de estar de acuerdo con el planteamiento, agrega que de no ser así se estaría negando el carácter dialéctico de la ciencia, con la oportunidad al debate para el mejoramiento de técnicas y métodos, que enriquecen la visión y surgimiento de nuevos propósitos científicos.

Las políticas editoriales deben ser orientadas fundamentalmente para que el proceso de comunicación científica sea responsable de viabilizar el impacto de la producción científica (Bordons, 2004), de lo contrario se condena al estancamiento el esfuerzo de la comunidad que lo genera, para esto se debe revisar tenazmente la conformación de metodologías, que garanticen la visibilidad de las revistas en base de datos.

Las bases de datos constituyen la fuente primaria para las prácticas bibliométricas, como bien es conocido almacenan grandes volúmenes de información; su características para la disposición y recuperación de la información las obligan a estructurar cada campo en función de la explotación de los datos que cubren, debido a esto es importante comprender que es deplorable arrastrar errores de normalización en la formulación de revistas científicas.

## **Materiales y Métodos**

Se analizaron las estructuras de bases de datos LILACS y SCOPUS para la determinación de indicadores bibliométricos de productividad y colaboración científica. Se trabajó con una muestra en LILACS de 4449 registros y en el caso de SCOPUS de 4492. Se analizaron 9 años, del 2000 al 2008. Se analizaron los campos: Autor, Dirección del Autor, Palabras Clave, Fuente de publicación, Resumen, Año, también se confrontó el las vías para exportar la información y los métodos para recuperar la información. Para la organización de la información se trabajó con el gestor bibliográfico de Endnote X. Como método de la investigación científica se utilizó el método empírico y el análisis documental

## **Resultados**

SCOPUS es una base de datos producto de la Elsevier B. V, ofrece datos a los autores citados por temática, actualmente dispone de más 300 usuarios y bibliotecarios

internacionales, cubre más de 15 000 títulos de 4000 editoriales con amplia visibilidad interdisciplinaria, fue creada por iniciativa de cooperación entre el grupo de investigación de la Universidad de Granada, en España: SCImago, con el objetivo de implementar la estructura de los datos para el desarrollo de la plataforma de indicadores bibliométricos, además de los tradicionales el novedoso SCImago Journal Country Rank (SJR) (Moya Anegón, Chinchilla Rodríguez, & Vargas Quesada, 2007)

LILACS es una base de datos producto de Sistema Regional de Salud Pública: BIREME, como alternativa para la comunidad iberoamericana para la difusión de resultados de investigación en el área de ciencias médicas, actualmente cubre más 670 títulos de revistas y aproximadamente más de 350 mil registros (Vera, 2002).

El estudio de estructuras de las bases de datos: LILACS y SCOPUS, para determinar indicadores bibliométricos permitió discernir la envergadura y el impacto del contenido de los campos en las bases de datos bibliográficas, referente a este tipo de indicadores de output. (Ver tabla I)

**Tabla I. Análisis de la estructura general de LILACS y SCOPUS**

<i>Campos</i>	<i>LILACS</i>	<i>SCOPUS</i>
<b>Nombre de los autores</b>	Nombre completo de los autores con los dos apellidos.	<i>Nombre del primer apellido, sigla en el nombre del autor.</i>
<b>Dirección institucional</b>	No presenta.	<i>Reconoce la dirección institucional de todos los autores.</i>
<b>Palabras clave</b>	Presenta.	<i>Presenta</i>
<b>Fuente de publicación</b>	Presenta todos los datos.	<i>Presenta los datos completos.</i>
<b>Exportación de los registros</b>	Por vía de correo electrónico, con salida en formato RIS.	<i>Por vía de correo electrónico y para con salida en formatos para los gestores bibliográficos: Procite y EndNote, hasta 2000 registros.</i>
<b>Recuperación de la información</b>	<i>Presencia de índice, sin embargo los registros que se necesitan deben ser marcados uno por uno por cada página.</i>	<i>Presencia de índice, se permite recuperar rápidamente los registros que se necesitan por páginas o todos sin necesidad de ir consultando las páginas.</i>

Al observar los datos anteriormente, se puede expresar que los problemas esenciales para originar indicadores bibliométricos de producción científica y colaboración, radican en el campo autor y afiliación, con respecto a LILACS también señalar lo engorroso que fue la recuperar la información para exportar los registros, ya que tuvieron que ser señalados uno a uno, además de no especificar el tipo de artículo que se está recuperando, SCOPUS sin embargo jerarquiza la información por tipo de artículo, años, segmentos disciplinarios, fuentes, pero se limita en el campo autor al colocar solamente el primer apellido y la sigla del nombre del autor.

## Discusión

### I. La voluntad de las publicaciones científicas

En el haber de las publicaciones científicas se reconocen aspectos lamentables que provoca la sordera ocular del nuevo conocimiento, por ejemplo: la claridad del lenguaje, la pertinencia del contenido con los objetivos planteados, la exhaustividad desmedida y redundante, el oscurecimiento en los materiales y métodos, entre otros. Al analizar a nivel macro el impacto de esto, no solamente parece la disposición de la comunidad científica en su intento de, en otra oportunidad valorar la perspectiva del autor, sino que se comienza a desarrollar estereotipo alrededor de la publicación científica que se consulta, paulatinamente esta publicación va perdiendo su relevancia y propósito. Ellas surgen como anexo a la investigación científica, por lo tanto la responsabilidad que comprende validar los resultados de investigación para resolver problemas concretos es trascendental, consecuencia de la experticia de un grupo de trabajo y de la alfabetización del investigador respecto a las prácticas para publicar.

Ciertamente el contenido de una publicación científica, independientemente de su formato (electrónico y/o impreso) debe manifestar el carácter científico de la sociedad a la que pertenece, especialmente deben estar diseñadas como espacio para que el investigador pueda estructurar documentalmente su resultado de investigación, dentro del contexto en el que ha sido concebido, en este sentido intervienen variables como el idioma, la cultura de publicación que incluyen los programas de formación científica, la normalización del lenguaje controlado para definir la temática trabajada, el dominio del vocabulario empleado para determinar términos técnico y particulares de cada país, sin embargo debe existir un consenso de los patrones nacionales por contribuir al acervo de la comunidad científica internacional, ya que no se trata de aislamiento, sino de establecer estándares de calidad permisibles para la apertura a la dinámica científica contrastable y comparable.

Desde la ética se podrían enumerar como bien indica Valderrama, Gutiérrez y Toledo (Toledo, 2008; U Gutiérrez Couto, Pérez, & Acción, 2004; Valderrama, 2004) incidencias degenerativas en las publicaciones científicas, que fluctúan desde: el plagio, fraude científico, autoría ficticia, fragmentación y duplicidad de los resultados, además de esas cuestiones, difíciles de detectar, porque precisan del dominio elitista por el comité editorial y el conjunto de árbitros, los cuales no son excluyentes de la complejidad humana imperfecta, existen otras que son más sencillas de manejar, que forman parte de la estructura formal y normativa que debe tener la publicación científica, datos tales como:

1. Presencia de la categoría docente o científica del comité editorial y el comité de redacción
2. Indicación de la internacionalidad de los árbitros
3. La fecha de aceptación y recepción de los artículos
4. Los datos formales como son el año, volumen, número, código ISSN
5. Las instrucciones a los autores
6. La filiación institucional de los autores
7. La normalización de los datos de los autores y coautores, sobre todo de aquellos que pertenecen a la región Latinoamérica, los cuales tienen dos apellidos

Al analizar los parámetros dentro del contenido y circulación no deben excluirse los siguientes:

1. Proporción de artículos originales debe ser mayor o igual al 60%
2. Correspondencia entre título de la publicación y contenido
3. Resúmenes (dos idiomas)
4. Palabras clave (dos idiomas)
5. Bases de datos que procesen la revista (Fleita, Sanchez, & Chacón, 2006)

Tradicionalmente las publicaciones científicas constituye el medio más utilizado para certificar el conocimiento. Se someten a procesos de validación, en las cuales los investigadores buscan el reconocimiento de sus resultados, se perfilan frentes de investigación, se expresa la relación de la inversión en la actividad científica y los resultados con respecto a las prioridades económicas de una región. Muchas de ellas se utilizan como herramientas para evaluar el desempeño de investigador y otorgar estímulos monetarios o concesiones de pasantías, así como financiamiento para nuevos proyectos de investigación, por lo cual es importante formalizar minuciosamente la estructura de las mismas, en pos de describir el esmero de los que contribuyen a través de ellas a la difusión del conocimiento y a los responsables de la visibilidad de estos resultados.

Se conoce del sesgo que sufren las publicaciones científicas por modelos como los del Institute for Information Science (ISI), pero independientemente de esto se atribuyen causas reparables que redundan en los factores anteriormente mencionados, lógicamente el ISI evalúa perfectamente la ciencia del primer mundo y es que fue determinado para tales fines, por lo tanto no es ajustable a condiciones que no cumplan socio-económica, política y científicamente las características de estos sistemas, lo que si no se debe permitir es que actualmente, al hablar de sociedades del conocimiento, la estructura documental del conocimiento aun siga siendo un problema sencillo, pero aun sin resolver.

## **II. El motivo de las bases de datos bibliográficas para evaluar la ciencia**

El auge de las publicaciones científicas cada vez más creciente y el incremento de las tecnologías de la información, propiciaron el marco conceptual idóneo para el surgimiento de las llamadas bases de datos, en sus inicios con el objetivo de almacenar electrónicamente y la recuperación de la información a niveles remotos. Sistemáticamente las bases de datos fueron incrementando el volumen de registros, ganando en organización y procesamiento de los contenidos especializados y multidisciplinarios. Actualmente se reconocen como gestoras para la evaluación del quehacer investigativo de la comunidad científica, a través de la aplicación de las técnicas para construir indicadores bibliométricos (Sánchez, Fleitas, & Álvarez, 2006).

La selección de las fuentes a cubrir por las bases de datos es determinada por distintas metodologías de evaluación, según la intención por las cuales se diseñan será la disposición de sus campos y el nivel de descripción de los datos respectivamente. Existen requerimientos generales que hacen posible la efectividad de una base de datos, tales como: la actualización, la presencia de índices y tesauros, el nivel de indización, la representación de la información, el acceso a la exportación de los datos con diferentes formatos de salidas etc.

Como efecto de causa de, la cada vez más necesaria presencia de indicadores científicos para evaluar los resultados de la actividad de investigación, así como el crecimiento de la información científica, las bases de datos para la obtención de indicadores bibliométricos, precisan de ser validadas rigurosamente en eutimia con el grado de producción científica que ellas cubren, el grado de descripción y estructura de los datos, la exportación de los registros para posteriores análisis matemáticos y la actualización, principalmente.

La omisión de particularidades como: el segundo apellido del autor o que se escriba en siglas el nombre de los autores impide, desde la prácticas métricas: identificar estudios de géneros, atribuir resultados de investigación a autores no responsables de los mismos, incrementar la sinonimia, o lo que es peor no saber quién es el responsable o líder de especialidades disciplinarias, se recomienda para estos casos que se separe por guiones los apellidos de los autores, desde la inclusión de estos datos en las publicaciones científicas, de esta forma arquetipos como los del ISI toman en consideración el derecho de los dos apellidos de los investigadores iberoamericanos, para nada se pretende seguir esquemas como los del ISI, pero es importante hacer lo más visible posible el empeño de los resultados científicos de la comunidad regional.

Otro de los problemas que se evidencian comúnmente es la supresión de la filiación institucional, esto enajena cualquier intento de examinar cuáles son las instituciones que más inciden en el afán científico, impide establecer y comparar núcleos de producción científica, imposibilita el análisis de la colaboración científica y el descubrimiento de los colegios invisibles, exceptúa por los entidades financieras el incremento de la inversión en la investigación.

La recuperación de la información es otra de las causas que limita considerablemente la generación de indicadores bibliométricos, el tener que marcar manualmente uno a uno los registros a procesar posteriormente, en un volumen de más de 8000 registros retarda el tratamiento consecutivo de los datos.

Hasta el momento solamente se ha tratado preocupaciones y limitaciones en el escenario de las bases de datos referente a precisar indicadores bibliométricos de producción y colaboración científica, preexisten aun subsecuentes estudios con respecto a las bases de datos encaminadas al impacto de la producción científica, al respecto señalar las iniciativas del proyecto SciELO y el grupo SCImago (Scimago, 2007), que van rompiendo con la nombrada clasista "corriente principal", añadir modestamente en este pesquis la disparidad socio-económica, política y científica que caracteriza a la sociedad iberoamericana, y la pertinencia de concepciones adecuadas de indicadores bibliométricos que evalúen la ciencia de la región correctamente.

### **III. El argumento de la bibliometría**

El uso de la bibliometría se recomienda cuando se comprende como premisa la significación de las publicaciones científicas para estimar el cometido de la investigación científica, su relación directamente proporcional con las bases de datos bibliográficas, y cuando las bases de datos cumplan como condición la representación del conocimiento estructurado de la sociedad que se pretender valorar científicamente.

No es objetivo de la bibliometría coordinar un sistema matemático disperso y ajeno a la dinámica de la comunidad científica, en el que solamente se registren, en forma de anuario, los tipos de documentos que se generan en un sistema de información (Spinak, 2006), es por esto que metodológicamente definirlos resulta complejo, sobre todo para homogeneizarlos internacionalmente ante entornos peculiares (Okubo, 1997). Principalmente ellos requieren de técnicas devenidas de las ciencias sociales (Hjorland, 2002; Katz, 1999) que permitan reconocer patrones, cualidades y características a partir de indumentarias cuantitativas, es ahí donde se localiza la importancia y complejidad de este tipo de indicadores, encaminados a evaluar los resultados de investigación, en la que implícitamente existen hábitos y comportamientos con desiguales sentidos éticos al respecto de cómo, para qué, por qué se publica y se cita.

La evaluación de las rutinas científicas a partir de la bibliometría necesita de incorporación de estos indicadores, en analogía con otros módulos de fuerte tradición de indicadores científico en los SNCTI, que principalmente han sido establecidos y revisados por entidades como la Organización para la Cooperación y Desarrollo Económicos (OCDE)(OCDE, 1997) ,las publicaciones como: *Science & Engineering Indicators*, del National Science Board de Estados Unidos, los *Science & Technologie Indicateurs* del Observatoire des Sciences et des Techniques de Francia, los *European Reports on S&T Indicators*, editados por la Comisión Europea(CINDOC, 1995); no solo para aprovechar la solidez estadística de los indicadores que en ellos se exponen, sino porque no se puede negar la sinergia holística que entraña apreciar la formulación de políticas científicas en la que intervienen diferentes áreas disciplinarias científicas, conformadas por los propios miembros de la sociedad y hacia ellos, en la cuales se pretende conectar e interpretar cualitativa y cuantitativamente las fortalezas y debilidades de la gremio científico, tecnológico e innovador.

### **Conclusiones**

Determinar las características y la tendencia de los la actividad científica dependen de las herramientas que se utilizan para evaluar los resultados de investigación, previamente se ha de definir conceptualmente qué y cómo identificar “resultados de investigación”, de esta forma se acota el radio de acción para la implementación de técnicas, además de clasificar por áreas del conocimientos, así como establecer parámetros específicos que cualifiquen el impacto de los resultados científico en la sociedad.

Los sistemas de información científico que se generan a partir de los SNCTI, son la base para estructurar el andamiaje estratégico de lineamientos científicos, ellos inciden para definir directivas entorno a la política científica de un país, debido a esto y desde la óptica metrista de la información científica, se manifiesta la importancia de clarificar los mecanismos para la compilación y normalización de la información científica. Anticipadamente se ha de reconocer la necesidad de que exista un conjunto de indicadores científicos, que más allá del aporte numerológico representen cualidades de la actividad científica, porque lamentablemente la propia carencia de modelos nacionales para evaluar resultados de investigación, provoca el reporte de estadísticas fundamentadas en diagnósticos incompletos a causa de no poder controlar la información científica que se genera.

Necesariamente la información científica precisa de estructura documental coherente, con el propósito de reflejar la responsabilidad de los resultados científicos. En el caso de los indicadores bibliométricos es menester que las publicaciones científicas y las bases de datos bibliográficas se ajusten en un modelo comunicacional organizado, y si bien a través de estos indicadores solo se podrá representar una parte del conocimiento científico de la comunidad, mínimo verificarán que existen resultados reales de investigación científica comunicados con la calidad pertinente.

## Bibliografía

- Bordons, M. (2004). Hacia el reconocimiento internacional de las publicaciones científicas españolas. *Revista Española de Cardiología*, 57, 799-802.
- CINDOC. (1995). Proyecto de obtención de indicadores de producción científica de la Comunidad de Madrid (PIPCYT). Madrid: Centro de Información y Documentación Científica.
- Fleita, M. E. M., Sanchez, Y. R., & Chacón, Y. S. (2006). EvaCyT: una metodología alternativa para la evaluación de las revistas científicas en la región. *ACIMED*, 14(5).
- Hjørland, B. (2002). Domain analysis in information science - Eleven approaches - traditional as well as innovative. *Journal of Documentation*, 58 (nº 4), 422-462.
- Katz, J. S. (1999). *Bibliometric Indicators and the Social Sciences*. Retrieved 10 de Mayo, 2008, from <http://www.sussex.ac.uk/Users/sylvank/pubs/ESRC.pdf>.
- López-Cózar, E. D. (1997). Incidencia de la normalización de las revistas científicas en la transferencia y evaluación de la información científica. *Revista de Neurología*, 25 (148), 1942-1946.
- Merton, R. K. (1957). *Social theory and social structure*: Glencoe, IL.
- Moya Anegón, F., Chinchilla Rodríguez, Z., & Vargas Quesada, B. (2007). Coverage análisis of Scopus: A journal metric approach. *Scientometrics*, 73(1), 53-78.
- OCDE. (1997). The Measurement of Scientific and Technological Activities. In OCDE (Ed.), *Manual de Oslo: Proposed guidelines for collecting and interpreting technological innovation data*. Paris.
- Okubo, Y. (1997). *Bibliometric indicators and analysis of research systems: methods and examples*. Paris: OCDE.
- Price, d. S. (1963). *Little science, big science*. , New York. USA.
- Sánchez, Y. R., Fleitas, M. E. M., & Álvarez, E. S. (2006). Cubaciencia y Cumed: dos fuentes para la obtención de indicadores bibliométricos en el área de la salud. *ACIMED*, 14(5).
- Scimago, G. (2007). SCImago Journal Rank (SJR) indicator. Retrieved 7 de Mayo, 2008, from <http://www.scimagojr.com/SCImagoJournalRank.pdf>
- Spinak, E. (2006). Indicadores cientiométricos - Bibliometría y Cienciometría. Retrieved 7 de Mayo 2008, from [http://www.wikilearning.com/articulo/indicadores\\_cientiometricos-bibliometria\\_y\\_cientiometria/8403-1](http://www.wikilearning.com/articulo/indicadores_cientiometricos-bibliometria_y_cientiometria/8403-1)
- Toledo, R. B. (2008, 2000). Aspectos éticos en las publicaciones científicas Retrieved 10 de Mayo, 2008, from <http://www.infodoctor.org/rafabravo/fraude.htm>
- U Gutiérrez Couto, Pérez, A. B., & Acción, B. C. (2004). Cómo realizar una comunicación científica. Estructura de la comunicación científica (I). *Revista Gallega de Terapia Ocupacional TOG*, Diciembre(1).
- UNESCO. (2005). What do bibliometric indicators tell us about world scientific output? . *Bulletin on Science and Technology Statistics*, Septiembre(2).



Valderrama, J. O. (2004). *Asuntos éticamente reprochables en una publicación científica*.

Paper presented at the V Encuentro de editores de Revistas Científicas.

Vera, I. A. (2002). Procedimientos de trabajo para LILACS. *ACIMED*, 4.

#### Datos de la autora

---

Yaniris Rodríguez Sánchez (Cuba)

Licenciada en Bibliotecología y Ciencia de la Información (curso 2003-2004).

Actualmente se desempeña como especialista de información en el IDICT en el grupo de investigaciones bibliométricas. Ha participado en cuatro eventos nacionales, dos internacionales, tres conferencias impartidas por la SOCIT , tres Forum, dos de bases y uno municipal. Tiene cuatro publicaciones, dos como autora principal y dos como coautora. Se encuentra realizando su doctorado en Ciencias Técnicas.

yrs201181@gmail.com