



Editorial

Como organizar la información de ciencia y tecnología en Latinoamérica?

How to organize science and technology information in Latin America?

Mauricio Palacios

Editor asociado, Revista Colombia Médica, Universidad del Valle. Cali, Valle, Colombia.

Palacios M. How to organize science and technology information in Latin America?. Colomb Med (Cali). 2016; 47(3): 131-32.

© 2016 Universidad del Valle. Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Attribution License, que permite el uso ilimitado, distribución y reproducción en cualquier medio, siempre que el autor original y la fuente se acreditan.

Historia: Recibido: 15 Sep 2016, Revisado: 26 Sept 2016, Aceptado: 27 Sept 2016.

Para orientar mejor las políticas de ciencia y tecnología se requiere información de alta calidad y actualizada sobre las organizaciones, investigadores, proyectos y productos. El crecimiento del uso del internet en investigación aportó mayor información sobre estos aspectos; pero, el volumen de datos dificultó la forma de procesarlos y organizarlos en forma útil para comprenderlos y tomar decisiones sustentadas. Problemas como la duplicación de información, dificultades en el seguimiento de procesos (autores y proyectos con productos), falta de identificación de temáticas de investigación y de redes de conocimiento han aumentado en los últimos veinte años. En todo esto, el factor más importante que dificulta la organización de los datos ha sido la necesidad de identificación de cada componente.

Tal vez, el elemento del ecosistema de la ciencia más complicado de identificar son los investigadores debido a que existe mucha variabilidad en la escritura del nombre, la afiliación institucional y el correo de contacto. Adicionalmente, los homónimos, los errores tipográficos y de citación afectan de forma importante cualquier base de datos^{1,2}. Por esta razón, muchas propuestas para identificar personas se han generado en el mundo desde los años 90. Latinoamérica realizó también una propuesta de identificación de los investigadores denominada Proyecto del *Currículum Vitae* en Ciencia y Tecnología de América Latina y el Caribe, más conocida como CVLAc. Este proyecto adoptó la tecnología y metodología Currículum Lattes desarrollada por el Consejo Nacional de Desarrollo Científico y Tecnológico de Brasil (CNPq). Fue ambicioso en sus objetivos, con acceso abierto para compartir información y se implementó con el apoyo de la Organización Panamericana de la Salud (OPS) y el Centro Regional de Información de Ciencias de la Salud (BIREME) a comienzos del presente siglo³. CVLAc no alcanzó todos los objetivos propuestos; aunque ha sido el sistema de identificación de investigadores que

más convocó en Latinoamérica y con él se apoyó la búsqueda de pares evaluadores de las revistas. Además, inspiró a otros desarrollos propios al quehacer investigativo de la región como las bases de datos de grupos y de revistas. La herramienta fue exitosa en la primera década del presente siglo.

Actualmente, CVLAc presenta varios problemas derivados del escaso desarrollo posterior a su implementación. El más prominente es el aislamiento respecto a los otros sistemas de identificación internacionales. Sin desarrollo en la adquisición de metadatos, integración de la información y armonización de los identificadores, el interés es escaso para una iniciativa recíproca con alguno de los sistemas de mayor cobertura como Researcher ID, Research Gate, Google Scholar, o incluso de la región como el de Autores Redalyc. El segundo problema es la transparencia en la información. CVLAc no tiene un sistema de validación sobre el ingreso de los datos que detecte errores o duplicación de la información. Todavía se realizan procesos de verificación manual por convocatoria y esto ha permitido algunos casos de conducta cuestionable. Existen otros aspectos (actualización manual, almacenamiento de datos, etc) que le restan atractivo y competitividad respecto a otras bases de datos. Por estas razones se puede considerar que CVLAc y los desarrollos derivados de él no tienen información con la calidad para sustentar políticas y acciones en caminadas al desarrollo de científico de un país.

Estas dificultades con los sistemas de información en ciencia se presentan en diferente medida en muchos de los países del mundo⁴, y las alternativas de solución deben tener la infraestructura de interoperabilidad para ofrecer consultas con tres atributos: identificación, correferencia y semántica⁵. La gestión de la información en investigación requiere programas de acceso abierto para evitar sobre costos a las instituciones o investigadores,

Autor de correspondencia:

Mauricio Palacios. Editor Revista Colombia Médica, Universidad del Valle. Telephone: +57 2 558 1939, Cali, Valle, Colombia. E-mail: mauricio.palacios@correounivalle.edu.co.

con semántica normalizada para poder enlazar información de varias bases de datos entre sí, con sistemas de identificación universales que eviten errores de duplicación y fragmentación de la información, y con código abierto con el fin de realizar los cambios necesarios que requiere las formas particulares de países o regiones en la forma de investigar, asociarse y publicar⁶. Con estos requisitos, se destacan algunos programas que son la base de la información en países con liderazgo en investigación.

El sistema de identificación de objetos (DOI) implementado y respaldado por Data Cite ha sido el más claro avance y con resultados calculables después de su implementación. Para el caso de artículos de revista, capítulos de libro y otros documentos resultados de investigación, la marcación está normalizada y la semántica se realiza con lenguaje controlado. En la práctica, el uso del DOI evita la duplicación de productos, facilita los análisis bibliométricos y la interoperabilidad con otra base de datos se realiza sin dificultad⁷. La implementación es institucional o editorial con costos bajos.

Inspirado en el DOI, se desarrolló el sistema de identificación de investigadores ORCID (Open Researcher and Contributor ID). Es un sistema internacional de acceso abierto creado y sustentado por Consortia Advancing Standards in Research Administration Information. ORCID tiene interoperabilidad con otros sistemas de identificación como researcher ID y Autores Redalyc (para autores de revistas latinoamericanas). Puede importar información de productos donde el autor use su identificador ORCID y tiene auditorías sobre la solicitud de un producto identificado con el DOI por parte de más de un autor homónimo. En el mundo se le reconoce porque es el sistema que más aporta a la desambiguación de autores y mejora la transparencia en la forma de publicar. Algunos países lo han adoptado como iniciativa oficial para sus sistemas de información, entre ellos Australia e Italia⁸. En Italia, la implementación del ORCID logró una vinculación superior al 80% de los investigadores en los primeros 6 meses, mejoró el flujo de información en los centros de investigación y en el organismo estatal de ciencia y tecnología⁹.

Institucionalmente, se requiere un gestor de información que interactúe con los dos sistemas anteriores. Hasta el momento, el software DSpace, desarrollado por Massachusetts Institute Technologic, ha mostrado mejor posibilidades de adaptación a las particularidades de cada centro de investigación o región. La adopción generalizada para un país, no depende de una iniciativa gubernamental. La Organización de Universidades Italianas, ANDU, definió su uso como un emprendimiento académico, que luego fue respaldado por el ministerio de Educación e Investigación italiano⁹. Para las instituciones de investigación resolvería en parte la invisibilidad de la ciencia latinoamericana¹⁰.

Un poco menos definido, por la complejidad de la tarea y el volumen de información para procesar, son los sistemas nacionales de información de ciencia e investigación. En este nivel, no hay consenso en el programa a usar; pero, se puede analizar los desarrollos a partir de la integración de los sistemas DSpace institucionales¹¹.

Todos estos sistemas requieren la voluntad de los investigadores, las instituciones y la organización de ciencia y tecnología gubernamental. Pero, convencer a cada uno de los participantes

puede ser más fácil porque estos sistemas retribuyen la participación en formas de consultas temáticas, por procesos, bibliométricas, entre otras. Podrían generarse discusiones durante la implementación sobre problemas de persistencia de información, de autoridad y de confianza; válidas y necesarias para la sociedad; pero, que no contradicen la utilidad para los sistemas de investigación.

Con reconocimiento por el aporte realizado por el CVLAC y sus bases de datos derivadas es hora de abandonarlas; estamos en el tiempo en el que Latinoamérica tenga mayor difusión y diseminación de su ciencia y pensamiento crítico para el mundo con identidad y lenguajes universal.

Conflicto de interés:

No hay

Referencias

- Potthast M, Braun S, Buz T, Duffhauss F, Friedrich F, Gülzow JM, *et al.* Who Wrote the Web? Revisiting Influential Author Identification Research Applicable to Information Retrieval. Ferro N, Crestani F, Mothe J, Silvestri F, Di Nunzio GM, Moens M-F. *Advances in Information Retrieval*. Chapter 9626 Switzerland: Springer International Publishing; p. 393–407. 2016.
- Lerchenmueller MJ, Sorenson O. Author disambiguation in PubMed: evidence on the precision and recall of authority among NIH-Funded Scientists. *PLoS One*. 2016;11(7): e0158731.
- de los Rios R, de Assis SPH. El espacio virtual de intercambio de información sobre recursos humanos en Ciencia y Tecnología de América Latina y el Caribe Del CV Lattes al CvLAC. *Ciência Informação*. 2001; 30(3): 42–7.
- Mazov N, Gureev V. The role of unique identifiers in bibliographic information systems. *Scient Tech Information Processing*. 2014; 41(3): 4.
- Bazzanella B, Bouquet P. An Interoperability Infrastructure for Digital Identifiers in e-Science. De Nart D, Tasso C, Calvanese, Diego. *Digital libraries on the move*. Chapter 612 Switzerland: Springer International Publishing; p. 167–178. 2016.
- Delgado JE. *Revistas Científicas de Universidades de Chile, Colombia y Venezuela: Actores y Roles*. *Education Policy Analysis Arch*. 2014; 22(34): 1–23.
- Neumann J, Brase J. DataCite and DOI names for research data. *J Computer-Aided Molecular Design*. 2014; 28(10): 1035–41.
- Boon W-M, Leves F. NHMRC initiatives to improve access to research outputs and findings Wee-Ming Boon. *Medical J Australia*. 2016; 202(11): 558.
- Galimbertia P Susanna, Mornatib S. The Italian model of distributed research information management systems: a case study. 13th International Conference on Current Research Information Systems; Scotland, UK2016.
- Orduña-Malea E, López-Cózar ED. The dark side of open access in Google and Google Scholar: the case of Latin-American repositories. *Scientometrics*. 2015; 102(1): 829–46.
- Amorim R, Castro J, Rocha dSJ, Riveiro C. A comparison of research data management platforms: architecture, flexible metadata and interoperability. *Universal Access in the Information Society*. 2016; 1–12.