
CoARA y la gobernanza de la evaluación

CoARA and the governance of evaluation

Jordi Molas Gallart

(jormoga@ingenio.upv.es)

Instituto de Gestión de la Innovación y del Conocimiento
(INGENIO), CSIC-UPV, Valencia

Recibido: 29-05-2023; Revisado: 07-06-2023, 12-06-2023; Publicado: 29-06-2023

Resumen: La reforma de la evaluación de la investigación emprendida por muchas de las mayores instituciones científicas europeas responde a problemas profundos y requiere cambios que van mucho más allá de la incorporación de nuevos indicadores y la adaptación de las prácticas actuales.

Palabras clave: evaluación de la investigación; gobernanza; CoARA

Abstract: The reform of research evaluation undertaken by many of the largest European scientific institutions responds to deep-seated problems and requires changes that go far beyond the incorporation of new indicators and the adaptation of current practices.

Keywords: research evaluation; governance; COARA

Como citar este artículo/Citation: Molas Gallart, J. (2023). CoARA y la gobernanza de la evaluación. *Enredadera: revista de la Red de Bibliotecas y Archivos del CSIC*, (39), 33-37. <https://doi.org/10.20350/digitalCSIC/15378>

El 20 de julio de 2022 se publicó el Acuerdo sobre la Reforma de la Evaluación de la Investigación (Agreement on Reforming Research Assessment¹). Impulsado por la Comisión Europea, en la elaboración del documento participaron más de 350 organizaciones. La redacción y publicación del Acuerdo ha sido un paso de un proceso que continúa con el establecimiento de una “Coalición” cuyo objetivo es implementarlo (Coalition on the Reform of Research Assessment – CoARA²). A finales de mayo de 2023, 553 organizaciones de todo el mundo han firmado el acuerdo, de las cuales 481 se han unido a la Coalición para participar activamente en un proceso colaborativo de reforma.

El Acuerdo identifica una serie de problemas con los que se enfrenta la evaluación de la ciencia y propone abordarlos de forma coordinada con la

¹ <https://coara.eu/agreement/the-agreement-full-text/>

² <https://coara.eu/>



participación voluntaria de instituciones científicas de todo tipo (universidades, centros de investigación, agencias de evaluación, organismos financiadores, infraestructuras de investigación y asociaciones científicas). En su Anexo I el Acuerdo señala que

“La comunidad investigadora coincide en que, para mejorar la calidad de la investigación y el atractivo de los entornos de investigación, es necesario reformar las prácticas de evaluación de la investigación... Se sabe que los procesos de evaluación que se basan predominantemente en métricas basadas en revistas y publicaciones dan lugar a una cultura de ‘publicar o perecer’ que no reconoce la diversidad de enfoques y podría ir en detrimento de la calidad... Por ejemplo, este predominio puede: fomentar la cantidad y la rapidez en detrimento de la calidad y el rigor; propiciar la aparición de revistas y congresos depredadores; fomentar la publicación en revistas de pago debido a sus elevados factores de impacto, a pesar de la disponibilidad de alternativas de acceso abierto; inducir a la aversión al riesgo, ya que asumir riesgos puede reducir las posibilidades de publicación; generar una atención excesiva a los rankings que obstaculice la colaboración; y malgastar esfuerzos, tiempo y recursos mediante la duplicación del trabajo, ya que los hallazgos ‘negativos’ quedan en gran medida sin comunicar.” (mi traducción)

Este diagnóstico se alinea con los problemas identificados por documentos anteriores como la San Francisco Declaration On Research Assessment (DORA³) y el Leiden Manifesto (Hicks et al., 2015), que coinciden en la necesidad de cambiar las prácticas de evaluación. Todos estos documentos se focalizan en el uso inadecuado de indicadores bibliométricos para la evaluación del desempeño de los científicos y de las organizaciones en las que trabajan. El argumento central es que las prácticas que convierten al indicador bibliométrico en el criterio dominante de evaluación son incapaces de recoger la diversidad de contribuciones a la generación y aplicación del conocimiento que se hacen desde la ciencia. Para ello es necesario recurrir a las valoraciones de expertos, informadas por, pero no subordinadas a, métricas relevantes. Siendo esto así, ¿no sería suficiente con ajustar las técnicas y protocolos de evaluación para alinearlos con el consenso que estos documentos reflejan? Cuidar de que los paneles de evaluación hagan un uso responsable de indicadores cuantitativos e identificar indicadores que ayuden de forma efectiva a los evaluadores parecería no requerir una coalición con más de 400 participantes institucionales. ¿Por qué es necesario organizar una coalición del alcance de CoARA? La respuesta se puede encontrar entre los objetivos que el Acuerdo plantea: “los firmantes del Acuerdo colaborarán para hacer posible una *reforma sistémica* basada en los principios comunes...” (mi énfasis). Los cambios necesarios no son sólo un ajuste de procedimientos, sino que requieren un cambio profundo en el sistema de gobernanza de la ciencia.

³ <https://sfdora.org/read/>

El creciente uso indiscriminado de métricas basadas en publicaciones científicas afecta, no sólo a los procesos de evaluación, sino también a la forma en qué se hace y se comunica la ciencia. En primer lugar, ha ocasionado un “desplazamiento de objetivos” (Rijcke et al., 2015): ante un sistema de evaluación que gira en torno a las publicaciones y las citas, el objetivo central de la actividad científica ha pasado de la generación y aplicación de conocimiento a la publicación en medios específicos (“revistas de impacto”) y la obtención de citas. Los problemas citados en el Anexo I del Acuerdo son la expresión de este cambio, pero sus efectos van más allá. La forma en que se transmite el conocimiento se ve también afectada cuando los investigadores buscan incrementar “sus” indicadores de publicación e impacto favoreciendo, por ejemplo, las auto citas, o dividiendo un resultado entre varias publicaciones (“salami publishing”) (Huth, 1986). Más allá de estos intentos se han expandido comportamientos fraudulentos como la venta de autorías o las más sofisticadas “círculos de citas” (por los que se establecen un grupo de revistas interconectadas que generan citas las unas a las otras) (Lockwood, 2020). Estas prácticas tienen el efecto de desvirtuar el indicador que constituye su objetivo: el efecto que se conoce como “Ley de Goodhart” (Chrystal & Mizen, 2003). La “Ley de Goodhart” sugiere que cuando un indicador se usa como objetivo para evaluar el desempeño, los evaluados pueden alterar su comportamiento de manera que distorsionan el significado del indicador. La cita deja de transmitir la información que transmitía en la ausencia del sistema de evaluación que la encumbra como medida de desempeño. Aún más importante es el efecto sobre la selección de temas de investigación. En disciplinas con alta producción de artículos, los investigadores se ven forzados a orientar su investigación hacia líneas que han recibido ya un alto número de citas (Chu & Evans, 2021). Por otro lado, en áreas con un número relativamente bajo de investigadores (como por ejemplo la taxonomía) emergen dificultades que impiden la valorización de sus contribuciones a través del análisis de citas (Zeppelini et al., 2021).

El hecho de que la publicación científica haya pasado de ser, ante todo, un mecanismo de comunicación a convertirse primordialmente en un instrumento para el reconocimiento y la evaluación de los científicos y las organizaciones en las que trabajan, se ha combinado en la última década con un cambio profundo en el mercado editorial: el paso de un modelo de “pagar por leer” a uno de “pagar por publicar”. Este cambio, apoyado por grandes grupos editoriales especializados en estos nuevos modelos de negocio, ha sido, hasta el momento, el principal resultado de las políticas de “acceso abierto”. El efecto ha sido una combinación explosiva de un modelo de negocio editorial en que ingresos y beneficios se incrementan con el número de publicaciones, con una práctica de gobernanza científica en que la evaluación de los investigadores se basa también en el número de publicaciones y citas. Un modelo en el que se generan más publicaciones, se generan también más citas y más beneficios para los grupos editoriales líderes en este nicho de mercado. Tal ha sido el éxito de este modelo que ha convertido a las editoriales especializadas en la producción de mega- revistas de pago en preferidas del sistema académico español (López-Cózar & Martín-Martín, 2022), todo ello con un efecto económico no desdeñable y que no ha pasado desapercibido entre los organismos que financian la investigación académica.

Finalmente, la práctica de hacer de los indicadores bibliométricos la piedra de toque de la evaluación científica se ha incrustado en la gestión de los recursos destinados a la investigación y en la imagen popular que se ha construido del desempeño científico. Por una parte, el uso de indicadores bibliométricos permite la comparación rápida cuando existen muchos sujetos a evaluar y poca capacidad evaluadora. También, al ofrecer medidas de desempeño que podría decirse no dependen de una valoración subjetiva, resulta muy útil para justificar decisiones en contextos donde la litigación contra las decisiones de comités de evaluación es común. El recurso a indicadores cuantitativos simples puede presentarse entonces como una garantía de “seguridad jurídica” a pesar de sus efectos nocivos sobre la actividad científica. Por otra parte, la publicación de cualquier ranking de investigadores u organizaciones basados en indicadores bibliométricos recibe de inmediato la atención de una prensa normalmente desinteresada en la gestión de la ciencia, y es publicada con ostentación por las instituciones cuya imagen se refuerza por el hecho de que sus investigadores hayan destacado por el impacto de sus publicaciones científicas.

En resumen, el recurso a indicadores bibliométricos como métrica determinante en procesos de evaluación forma parte de un sistema fuertemente acoplado que define la gobernanza de la ciencia. Por tanto, la reforma que persigue CoARA no sólo exige un cambio en las técnicas y procesos de evaluación, sino que requiere un profundo cambio sistémico. El problema es de formulación fácil, pero de solución difícil. CoARA reconoce el alcance de los cambios necesarios y el Acuerdo recoge aspectos como la necesidad de educar y apoyar a los investigadores y de proporcionar recursos para lograr los cambios organizativos necesarios. Todo ello se requiere para conseguir un sistema de evaluación flexible capaz de reconocer la diversidad de contribuciones a la investigación y de las carreras investigadoras.

Referencias

- Chrystal, K. A., & Mizen, P. D. (2003). Goodhart's Law: its origins, meaning and implications for monetary policy. En P. Mizen (Ed.), *Central banking, monetary theory and practice: Essays in honour of Charles Goodhart* (Vol. 1, pp. 221-243). Edward Elgar. <https://doi.org/10.4337/9781781950777.00022>
- Chu, J. S., & Evans, J. A. (2021). Slowed canonical progress in large fields of science. *Proceedings of the National Academy of Sciences*, 118(41), e2021636118. <https://doi.org/10.1073/pnas.2021636118>
- Hicks, D., Wouters, P., Waltman, L., Rijcke, S. de, & Rafols, I. (2015). The Leiden Manifesto for research metrics. *Nature*, (520), 429-431. <https://doi.org/10.1038/520429>
- Huth, E. J. (1986). Irresponsible authorship and wasteful publication. *Annals of Internal Medicine*, 104(2), 257-259. <https://doi.org/10.7326/0003-4819-104-2-257>
- Lockwood, M. (2020). Editorial: citation malpractice. *Proceedings of The Royal Society A*, 476(2242), 20200746. <https://doi.org/10.1098/rspa.2020.0746>
- López-Cózar, D., & Martín-Martín, A. (2022). Detectando patrones anómalos de publicación científica en España: Más sobre el impacto del sistema de evaluación científica. [Preprint] https://www.researchgate.net/publication/363535388_Detectando_patrones_anomalos_de_publicacion_cientifica_en_Espana_Mas_sobre_el_impacto_del_sistema_de_evaluacion_cientifica

- Rijcke, S. de, Wouters, P. F., Rushforth, A. D., Franssen, T. P., & Hammarfelt, B. (2015). Evaluation practices and effects of indicator use—a literature review. *Research Evaluation*, 25(2), 161-169. <http://dx.doi.org/10.1093/reseval/rvv038>
- Zeppelini, D., Dal Molin, A., Lamas, C. J. E., Sarmiento, C., Rheims, C. A., Fernandes, D. R. R., Lima, E. F. B., Silva, E. N., Carvalho-Filho, F., Kovác, L., Montoya-Lerma, J., Moldovan, O.T., Souza-Dias, P. G. B., Demite, P. R., Feitosa, R. M., Boyer, S. L., Weiner, W. M., & Rodrigues, W.C. (2021). The dilemma of self-citation in taxonomy. *Nature Ecology & Evolution*, 5(1), 2. <https://doi.org/10.1038/s41559-020-01359-y>