



# Herramienta para Evaluación de Recursos Académicos

Estado actual y perspectivas a futuro

Rucci, Enzo

*III-LIDI, Facultad de Informática, Universidad Nacional de La Plata - Comisión de Investigaciones Científicas*

[erucci@lidi.info.unlp.edu.ar](mailto:erucci@lidi.info.unlp.edu.ar)

<https://orcid.org/0000-0001-6736-7358>

Villarreal, Gonzalo Luján

*PREBI-SEDICI, Universidad Nacional de La Plata - CESGI, Comisión de Investigaciones Científicas*

[gonzalo@prebi.unlp.edu.ar](mailto:gonzalo@prebi.unlp.edu.ar)

<https://orcid.org/0000-0002-3602-8211>

Última actualización: 21 de septiembre de 2023



# HERA - Contexto General

La rapidez con la que la cantidad y disponibilidad de información aumenta ha tenido fuerte impacto en el modo en que los investigadores la utilizan. Es habitual que un investigador lea, acceda y utilice información extraída de artículos de otros colegas. Hasta hace algunas décadas atrás, el requisito de originalidad y el proceso de revisión por pares eran elementos suficientes para certificar la calidad del material con el que se trabajaba. Sin embargo, esto cambió de manera radical con el surgimiento de la Internet y de sistemas con mayor grado de automatización para la publicación y evaluación de documentos científicos [1]. Un estudio realizado a finales del 2018 sugiere que al menos 3 millones de artículos fueron publicados en ese año, en aproximadamente 33100 revistas científicas [2].

Una característica muy buscada y asociada estrechamente con la calidad tanto de una revista como de un artículo es su impacto. De hecho, es un error común confundir calidad con impacto y tomarlos como sinónimos. La calidad de una revista se asocia a los procesos que esta sigue y a los requisitos que impone para que un artículo sea finalmente publicado. Por su parte, el impacto se asocia a la repercusión que la revista o el artículo tienen en la comunidad científica, usualmente medido en número de citas [3].

La necesidad de determinar calidad e impacto de recursos científicos se traduce en lo que hoy vemos como sistemas de evaluación, indicadores y métricas de revistas y artículos. Si bien no son una herramienta automatizada en su totalidad, ni mucho menos estandarizada, estos elementos permiten tener valores concretos o numéricos que pueden dar (o aproximar) respuestas a preguntas como: este recurso<sup>1</sup> ¿es de buena calidad?, ¿es de alto impacto?, ¿es útil?, ¿es válido?

Lamentablemente, los sistemas, métricas e indicadores mencionados anteriormente no están exentos de subjetividad ni de sesgos. En consecuencia, actualmente los miembros de la comunidad académica y científica se enfrentan a una tarea ardua y engorrosa cuando deben determinar la calidad y el impacto de una publicación científica. Esta situación se debe a dos factores. En primer lugar, al crecimiento exponencial en la disponibilidad de recursos que tuvo el área debido al desarrollo tecnológico. En segundo lugar, la ausencia de estándares y la consecuente existencia de múltiples métricas y sistemas de evaluación que, aunque comparten objetivos, no siempre lo llevan a cabo de la misma manera.

Es en este contexto donde surge la Herramienta para Evaluación de Recursos Académicos (HERA), una herramienta web que apunta a simplificar, agilizar y apoyar el proceso de determinar la calidad y el impacto de un recurso académico. Para ello, recopila información proveniente de diferentes bases de datos académicas (principalmente indicadores/métricas de calidad e impacto) en forma rápida, para luego ser exhibidas de forma integrada y amigable al usuario.

## Propósito y Versiones

HERA apunta a simplificar, agilizar y dar apoyo al proceso de evaluación de recursos académicos:

---

<sup>1</sup> Se entiende como recurso a cualquier publicación individual o seriada que resulte habitual en la ciencia.

- La simplificación está dada por dos factores: 1) ante la multiplicidad y diversidad de bases de datos académicas existentes, HERA realiza una selección basada en criterio de expertos, logrando que el usuario final tenga una visión representativa del recurso a bajo costo; y 2) el funcionamiento no requiere más que ingresar el identificador del recurso en un campo de búsqueda y hacer clic en un botón.
- La agilidad proviene del proceso de recopilación automático de información que realiza HERA para un recurso determinado, lo que exime al usuario de la tarea de realizar búsquedas en sitios externos de forma individual.
- El apoyo al proceso de evaluación se basa en la recopilación y presentación conjunta de métricas e indicadores diversos asociado al recurso de interés. Rápidamente se podrán conocer números de citas (tradicionales y alternativas), factores de impacto, información sobre licencias y acceso abierto, e inclusión en bases de datos reconocidas. De esta manera, HERA provee una visión integral del recurso para que el usuario pueda realizar un análisis propio según sus objetivos e intereses.

La [versión 1.0 de HERA](#) fue puesta en producción en 2021 [4], [5]. Al momento de escribir este informe, una [segunda versión de HERA](#) se encuentra en BETA, la cual contempla diferentes mejoras y extensiones a la versión original. Considerando que saldrá a producción antes de fin de año, el informe se realiza considerando esta nueva versión.

## Aspectos de Funcionamiento Interno

### Arquitectura

HERA consiste de una aplicación web, que está dividida en una aplicación *frontend*<sup>2</sup>, desarrollada con React+Bootstrap, y en una aplicación *backend*, desarrollada con NodeJS, como se muestra en la Fig. 1. La aplicación backend tiene como objetivo servir de puente entre la aplicación frontend y los diversos servidores que proveen las métricas e indicadores de los recursos académicos, concentrando de esta forma la mayor carga de trabajo en sí misma. La aplicación frontend realiza peticiones a la de backend mediante una API<sup>3</sup> REST, recopilando los resultados, y es la encargada de generar las visualizaciones de los mismos en un portal web.

---

<sup>2</sup> Desde el punto de vista del diseño de aplicaciones informáticas, el concepto de frontend se asocia a los módulos y/o capas de la aplicación dedicados a presentar información al usuario, y a brindar algún tipo de interactividad (enlaces, menús, formularios, etc). En el diseño de aplicaciones multi-capas, el frontend carece de lógica de negocios (código dedicado a resolver los problemas para los que se construyó la aplicación); dicha lógica se concentra en la capa de *backend*, y esto suele incluir aspectos vinculados a la seguridad, acceso a datos, procesamiento intensivo, cumplimiento de reglas específicas del contexto, etc.

<sup>3</sup> Pieza de código que permite a diferentes aplicaciones comunicarse entre sí y compartir información y funcionalidades. Una API es un intermediario entre dos sistemas, que permite que una aplicación se comunique con otra y pida datos o acciones específicas

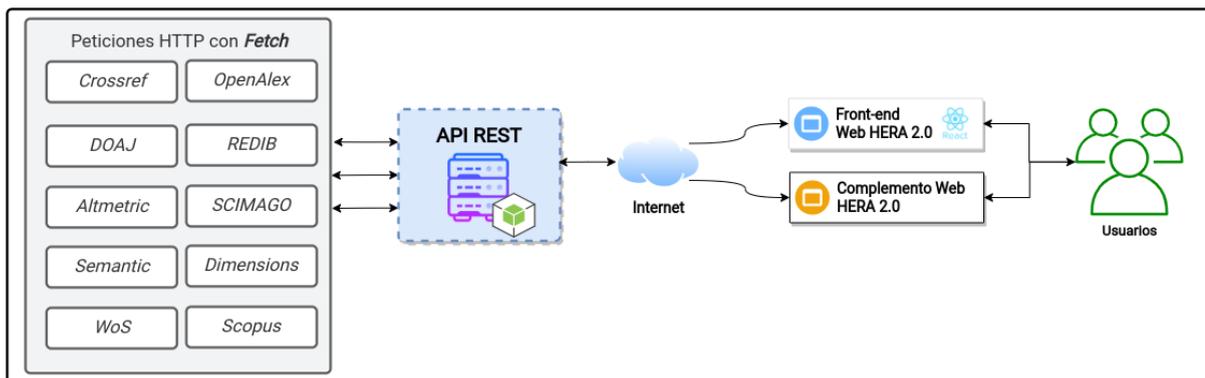


Figura 1: Esquema de comunicación de la API REST . El backend se compone tanto de la API REST como de los módulos encargados de recuperar datos desde las distintas fuentes, y se comunica con el frontend y con el complemento web a través de Internet.

## Medios de acceso a las bases de datos académicas

Al momento de consultar las bases de datos, se cuenta con 3 opciones según las posibilidades que estas provean:

- API: afortunadamente la mayoría de las bases de datos cuenta con una API disponible, que permite obtener información de un recurso mediante su identificador. Este proceso resulta efectivo, además de ser relativamente sencillo de usar.
- Web scrapping<sup>4</sup>: para aquellos sitios que no ofrecen una API pero resultan de interés para extraer datos, se opta por utilizar la técnica de *web scrapping*. Este proceso no es viable a gran escala ya que implica desarrollar un algoritmo específico para cada base de datos, que puede requerir ser actualizado ante cambios en su código HTML.
- Botón para búsqueda externa: finalmente, para los casos en que no disponen de una API pero tampoco resulta posible emplear técnicas de web scrapping, se opta por incluir un botón que permita la búsqueda del recurso en un sitio externo. Si bien no se muestra la información directamente en HERA, se simplifica y agiliza el acceso a esta al sólo tener que hacer clic.

## Bases de datos académicas incluidas

La Tabla 1 lista las bases de datos que consulta HERA 2.0, así como el medio acceso y las métricas que se obtienen en cada caso. En general, se han considerado las de mayor popularidad entre las multidisciplinarias y regionales.

Base de Datos	Medio de acceso	Principales métricas e indicadores	Recurso
Crossref	API	<ul style="list-style-type: none"> <li>● Número de citas</li> <li>● Estadísticas de papers por año y autores con ORCID</li> <li>● Información bibliográfica</li> </ul>	Artículo; Revista

<sup>4</sup> Refiere al proceso de recolectar datos desde la web usando software, ya sea de forma manual o automática.

<b>OpenAlex</b>	API	<ul style="list-style-type: none"> <li>• Número de citas</li> <li>• Estadísticas de <i>papers</i> por año y autores con ORCID+ROR</li> <li>• Información bibliográfica</li> <li>• Evento/Venue</li> <li>• Sello de DOAJ</li> </ul>	Artículo; Revista
<b>DOAJ</b>	API	<ul style="list-style-type: none"> <li>• Licencias</li> <li>• Sello de DOAJ</li> </ul>	Artículo; Revista
<b>Red Iberoamericana (REDIB)*</b>	Web scrapping	<ul style="list-style-type: none"> <li>• Posición de ranking</li> <li>• Información bibliográfica</li> <li>• Indicadores de calidad editorial y acreditaciones</li> </ul>	Revista
<b>Web of Science</b>	API	<ul style="list-style-type: none"> <li>• Índice de WoS</li> <li>• Información bibliográfica</li> </ul>	Revista
<b>Scopus</b>	API + Botón para búsqueda externa	<ul style="list-style-type: none"> <li>• Número de citas</li> <li>• Predicción de citas</li> <li>• Factores de impacto</li> </ul>	Revista; Artículo
<b>Scimago Journal &amp; Country Rank</b>	Web scrapping	<ul style="list-style-type: none"> <li>• Posición en ranking global</li> <li>• Cuartil</li> <li>• H-index del journal</li> <li>• Información bibliográfica</li> </ul>	Revista
<b>Semantic Scholar</b>	API	<ul style="list-style-type: none"> <li>• Número de citas</li> <li>• Frecuencia de citas</li> <li>• Número de citas influyentes</li> </ul>	Artículo
<b>Google Scholar</b>	Botón para búsqueda externa	<ul style="list-style-type: none"> <li>• Citas</li> </ul>	Artículo; Revista
<b>Dimensions</b>	API	<ul style="list-style-type: none"> <li>• Número de citas</li> <li>• Proporción de citas por campo y relativa</li> </ul>	Artículo
<b>Altmetric</b>	API	<ul style="list-style-type: none"> <li>• Puntuación de atención basada en menciones ponderadas según el sitio</li> <li>• Presencia de menciones en redes sociales</li> </ul>	Artículo

\*REDIB: Temporalmente sin servicio.

Tabla 1: Resumen de bases de datos consultadas en HERA 2.0.

## Gestor de fuentes de información

Para administrar las diferentes fuentes de información (bases de datos académicas), HERA dispone de un gestor específico, que le permite adaptarse de manera ágil y flexible ante cambios en las mismas. El gestor consiste en un archivo de configuración que almacena información relevante sobre cada fuente de datos, incluyendo su URL, su grado de prioridad, y si está habilitada o no. Gracias a esto, en caso que una fuente de datos deje de estar disponible (por ejemplo, como ocurrió con Microsoft Academic en 2021), los cambios necesarios se podrían hacer directamente en el archivo de configuración sin necesidad de modificar el código de la aplicación. De la misma forma, en caso de querer incorporar una nueva fuente (por ejemplo, OpenAlex), sólo debe incluir la información que el gestor requiere.

## API REST

La incorporación de una API REST en HERA 2.0 constituye una actualización estratégicamente significativa, dado que contribuye a una mayor integración con aplicaciones y servicios externos, a la optimización del rendimiento en las solicitudes y al aumento de la escalabilidad de la plataforma.

## Aspectos de Funcionamiento Externo

### Interfaz Gráfica de Usuario (GUI)

En la aplicación, se presenta una interfaz de usuario única que se caracteriza por la presencia de una destacada barra de búsqueda. En esta barra, los usuarios deben ingresar el identificador del recurso que desean buscar, tal como se ilustra en la Figura 2.



Figura 2: Interfaz de búsqueda de la herramienta con input de búsqueda y selectores de tipo de identificador (DOI e ISSN)

Dentro de la interfaz, los usuarios deben elegir el tipo de búsqueda que desean realizar, ya sea mediante el uso del DOI o del ISSN. Una vez seleccionado el tipo de búsqueda, deberán ingresar el identificador correspondiente en la barra designada. Más allá de la nueva interfaz, el flujo de búsqueda es el mismo que en el caso de HERA 1.0 en general. El único cambio consiste en la posibilidad de introducir múltiples identificadores en la barra de búsqueda (se deben separar por comas) para permitir su búsqueda simultánea.

### Búsqueda por ISSN

Al introducir un ISSN en la barra de búsqueda y presionar el botón Buscar, se realiza la búsqueda de ese recurso. De este se obtiene primero algunos metadatos para luego presentar la vista del resumen de métricas recopiladas de los diferentes sitios, visualizado en la Fig. 3. Adicionalmente, se cuenta con un botón para expandir la información y así visualizar la totalidad de su contenido, como se muestra en las Fig.4 y 5.

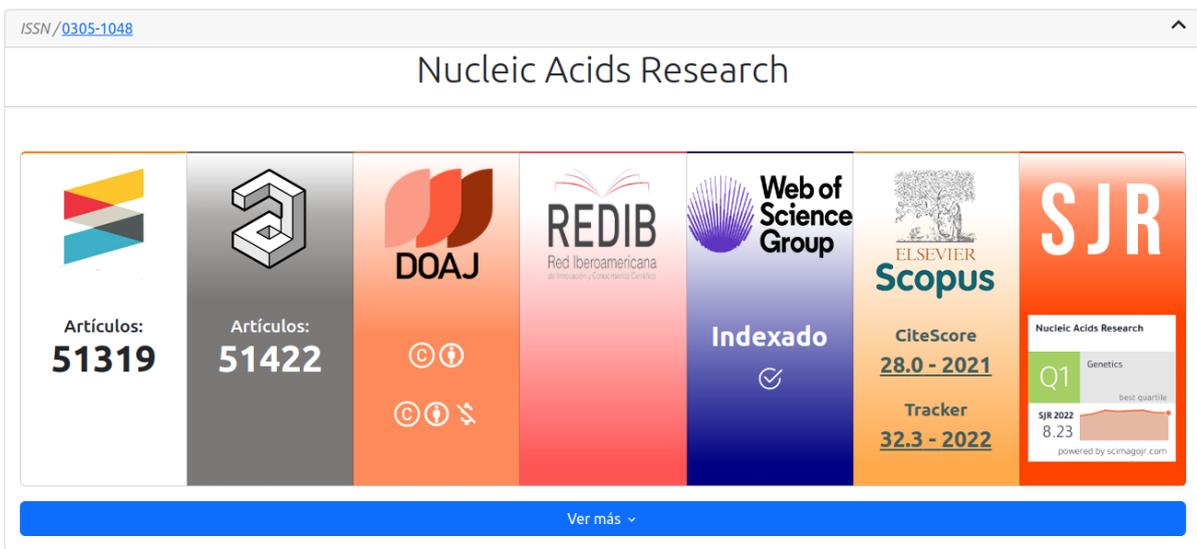


Figura 3. Resultados generales de la búsqueda de un journal (vista resumida)

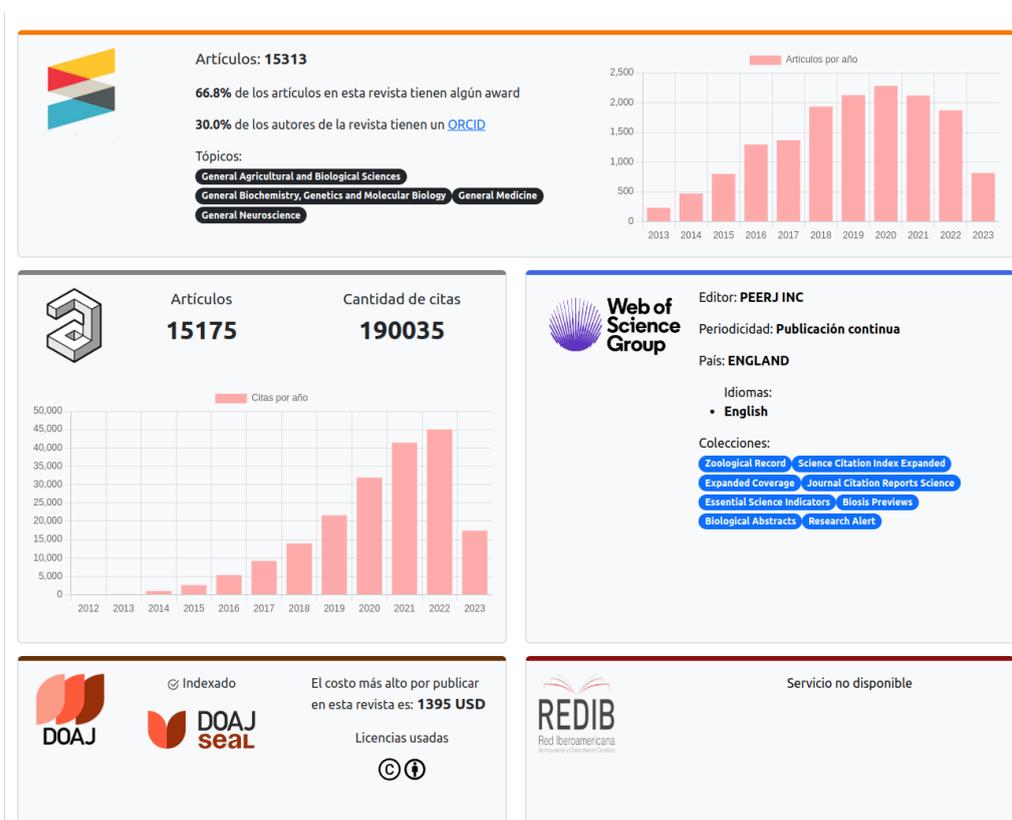


Figura 4. Vista expandida de métricas de un journal (parte 1)



Figura 5. Vista expandida de métricas de un *journal* (parte 2)

## Búsqueda por DOI

Después de ingresar un DOI en la barra de búsqueda y hacer clic en el botón de búsqueda, las métricas se recopilan de diversas bases de datos y se muestran en la pantalla, tal como se ilustra en la Figura 6. Entre los resultados, se pueden observar varios metadatos del recurso, como el título, el tipo de recurso, el resumen, entre otros. Además, HERA consulta las fuentes de datos ORCID y ROR, para integrar información de autores e instituciones, con los enlaces correspondientes. Finalmente, se muestra el nombre de la revista en la que se publicó el artículo, junto con un enlace que dirige a la página original del mismo. En caso de ser posible, la aplicación también intentará recopilar métricas del *journal* en el que se encuentra el artículo, y las mostrará en conjunto, como se muestra en la Figura 7. Para esto, se extrae el ISSN del *journal* de los metadatos recuperados y se realiza automáticamente una búsqueda por ISSN en segundo plano. De esta manera, se amplía el contexto de evaluación de un artículo, permitiendo ver sus métricas de forma aislada y también acceder a las métricas del *journal*, lo que ayuda a comprender qué influencia puede tener en las métricas del artículo. Al igual que en el caso de ISSN, al hacer clic en el botón "Ver más" que se muestra en esta última figura, se genera una vista expandida de las métricas, donde se puede obtener información más detallada (en los casos en que las bases de datos proporcionen datos adicionales), como se ilustra en la Figura 8.

10.7717/peerj.4375 DOI ISSN Buscar  
 Sólo exportar ?

Exportar Exportar

DOI/10.7717/peerj.4375

# The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles

**Autores:**

- **Heather A. Piwowar**  
Institución: [Impact Technology Development \(United States\)](#)
- **Jason Priem**  
Institución: [Impact Technology Development \(United States\)](#)
- **Vincent Larivière**  
Institución: [Université de Montréal](#)
- **Juan Pablo Alperin**   
Institución: [Simon Fraser University](#)
- **Lisa Matthias**   
Institución: [Simon Fraser University](#)
- **Bree Norlander**   
Institución: [University of Washington](#)
- **Ashley Farley**  
Institución: [University of Washington](#)
- **Jevin D. West**   
Institución: [University of Washington](#)
- **Stefanie Haustein**   
Institución: [University of Quebec at Montreal](#)

**Año de publicación:** 2018

**Título de la revista:** PeerJ

**Editorial:** PeerJ, Inc.

**Tipo de recurso:** journal-article

**Resumen:**

Despite growing interest in Open Access (OA) to scholarly literature, there is an unmet need for large-scale, up-to-date, and reproducible studies assessing the prevalence and characteristics of OA. We address this need using oaDOI, an open online service that determines OA status for 67 million articles. We use three samples, each of 100,000 articles, to investigate OA in three populations: (1) all journal articles assigned a Crossref DOI, (2) recent journal articles indexed in Web of Science, and (3) articles viewed by users of Unpaywall, an open-source browser extension that lets users find OA articles using oaDOI. We estimate that at least 28% of the scholarly literature is OA (19M in total) and that this proportion is growing, driven particularly by growth in Gold and Hybrid. The most recent year analyzed (2015) also has the highest percentage of OA (45%). Because of this growth, and the fact that readers disproportionately access newer articles, we find that Unpaywall users encounter OA quite frequently: 47% of articles they view are OA. Notably, the most common mechanism for OA is not Gold, Green, or Hybrid OA, but rather an under-discussed category we dub Bronze: articles made free-to-read on the publisher website, without an explicit Open license. We also examine the citation impact of OA articles, corroborating the so-called open-access citation advantage: accounting for age and discipline, OA articles receive 18% more citations than average, an effect driven primarily by Green and Hybrid OA. We encourage further research using the free oaDOI service as a way to inform OA policy and practice.

Figura 6. Resultados generales de la búsqueda de un *paper* (vista resumida)



Figura 7. Métricas correspondientes al *paper* y *journal* donde se encuentra el *paper* (vista resumida en tarjetas)

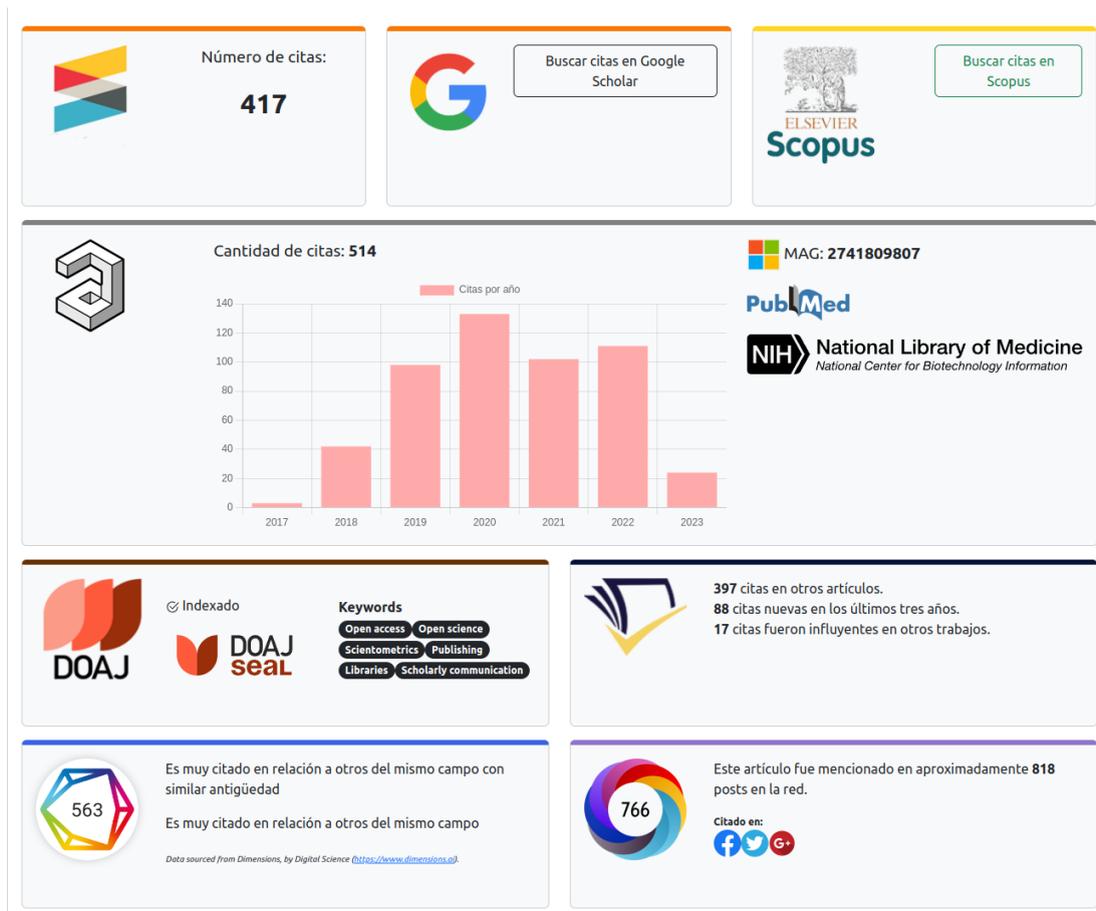


Figura 8. Vista expandida de métricas de un paper

## Exportación a formatos descargables

Esta funcionalidad proporciona una considerable ventaja para los usuarios que necesiten realizar análisis de datos, investigaciones académicas, auditorías bibliográficas, entre otros fines. En particular, se consideraron dos formatos que son ampliamente usados en la actualidad, como CSV y JSON.

La exportación de datos permite a los usuarios obtener los principales indicadores y métricas para un recurso académico (tanto para DOI como para ISSN) de forma conjunta y estructurada. Cuando se realizan búsquedas simultáneas, la información recolectada de cada DOI o ISSN se presenta en filas separadas, facilitando su lectura y posterior procesamiento. Debajo de la barra de búsqueda, los usuarios pueden encontrar una casilla de verificación etiquetada como "Sólo exportar". Esta opción, como se indica en el mensaje de ayuda emergente que aparece al pasar el cursor sobre el icono de ayuda contiguo, permite a los usuarios obtener los datos de exportación sin tener que visualizarlos primero en la interfaz web. Al seleccionar esta opción, HERA genera una lista de los DOI o ISSN solicitados, y presenta los botones para exportar los datos en formato CSV o JSON, reduciendo así el tiempo de procesamiento general. La Fig. 9 ilustra cómo aparece esta nueva característica en la interfaz de usuario.

10.3389/fenvs.2020.581591,10.1109/LCOMM.2019.2898944,10.1016/

DOI

ISSN

Buscar

Sólo exportar ?

Exportar información

Criterio: DOI

Exportar datos de:

- 10.3389/fenvs.2020.581591
- 10.1109/LCOMM.2019.2898944
- 10.1016/j.telpol.2021.102261
- 10.24215/26838559e28
- 10.1007/978-3-031-07802-6\_9
- 10.1109/CLEI53233.2021.9640225
- 10.1007/978-3-030-61702-8\_25
- 10.7717/peerj.4375
- 10.24215/18522971e087
- 10.5710/PEAPA.21.05.2020.298

Exportar  Exportar 

Figura 9. Función de “Sólo exportar” en la web de HERA 2.0

## Extensión web para navegadores

La [extensión web de HERA](#) ofrece a los usuarios la capacidad de acceder rápidamente a métricas e información relevante sobre documentos académicos a través de su navegador. Su principal función es reconocer los identificadores DOI presentes en la página web en la que el usuario se encuentra y, mediante una consulta a la API de HERA 2.0, recopila y presenta la información asociada a estos documentos. Aunque aun se encuentra en estado de prototipo, su estado actual ya es suficiente para considerarlo una prueba de concepto.

La información proporcionada por el complemento se adapta al contexto de la página web en la que se encuentra el usuario. En particular, el complemento de HERA 2.0 está orientado a reconocer dos contextos diferentes, los cuales se describen a continuación.

### Página específica de una publicación

Si la página es específica de un documento (es decir, se refiere a un único DOI), el complemento ofrece información detallada de dicho artículo. En particular, se presentan métricas e indicadores tanto del artículo como de la revista en la que se publicó (ver Figura 10), y se incluye un enlace a los resultados de buscar el artículo en cuestión en HERA 2.0.

The screenshot shows the PeerJ website interface. On the left, the article title is "The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles" by Heather Piwowar, Jason Priem, Vincent Larivière, Juan Pablo Alperin, Lisa Matthias, Bree Norlander, Ashley Farley, Jevin West, and Stefanie Haustein. The article is published on February 13, 2018. The abstract discusses the need for large-scale, up-to-date, and reproducible studies assessing the prevalence and characteristics of OA. It mentions using oaDOI, an open online service that determines OA status for 67 million articles. The abstract also notes that at least 28% of the scholarly literature is OA (19M in total) and that this proportion is growing, driven particularly by growth in Gold and Hybrid. The most recent year analyzed (2015) also has the highest percentage of OA (45%). Because of this growth, and the fact that readers disproportionately access newer articles, we find that Unpaywall users encounter OA quite frequently; 47% of articles they view are OA. Notably, the most common mechanism for OA is not Gold, Green, or Hybrid OA, but rather an unclassified category we call Bronze articles made available through Unpaywall.

On the right, the HERA sidebar displays the following information:

- Total citations: 588
- Recent citations: 197
- Field Citation Ratio: 486
- Relative Citation Ratio: 9.87
- DOI: 10.7717/peerj.4375
- Citas: 541 (for DOAJ)
- Citas: 434 (for DOAJ Seal)
- Citas: 411 (for DOAJ Seal)
- Citas: 588 (for Google Scholar)
- Menciones: 817 (for Scopus)
- Publicado en: PeerJ (2167-8359), PeerJ, Inc.
- Ver en HERA button

Figura 10. Complemento de HERA en funcionamiento en la web de PeerJ: The state of OA

## Página que contiene varias publicaciones

Por otro lado, si el usuario se encuentra en una página que presenta varios documentos, como podría ser una página de resultados en un motor de búsqueda académico (por ejemplo, Google Scholar) o un agregador de artículos académicos (por ejemplo, DOAJ), el complemento identifica todos los DOI presentes y proporciona una visión simplificada para cada uno de ellos. En este escenario, la información proporcionada se centra en el conteo de citas. Esta decisión se justifica principalmente por dos razones: por un lado, desde un punto de vista visual, mostrar una cantidad limitada de información permite una lectura rápida y clara para el usuario; por otro lado, existen limitaciones inherentes a los servicios que proveen los datos, lo que restringe la cantidad de información que puede ser obtenida y mostrada en un contexto de múltiples documentos.

Este enfoque ofrece una visión rápida del impacto de cada documento sin abrumar al usuario con una cantidad excesiva de datos, dándole la posibilidad de acceder a la información completa haciendo clic en el botón "Ver en HERA". La Fig. 11 ilustra un ejemplo de dicho funcionamiento.

The image shows a Semantic Scholar search interface for the term 'clustalw'. The search results are displayed in a list format, showing the title, authors, publication date, and a brief description (TLDR) for each paper. The HERA sidebar on the right provides a quick access to these results, with buttons labeled 'Ver en HERA' and citation counts (Citas (OpenAlex)).

Figura 11. Búsqueda de 'clustalw' en Semantic Scholar y funcionamiento del complemento de HERA

En resumen, el objetivo principal del complemento web de HERA 2.0 es proporcionar a los usuarios un acceso rápido y sencillo a la información relevante sobre documentos académicos, directamente desde el sitio que están navegando, evitando el proceso de búsqueda manual en la página web de HERA.

## Oportunidades actuales y a futuro de HERA

### Incorporación de nuevas bases de datos

La incorporación de nuevas bases de datos en HERA 2.0 es una tarea relativamente sencilla gracias a la implementación del gestor de fuentes de información descrito anteriormente. Simplemente se debe ingresar una nueva entrada en el archivo de configuración completando los datos que el gestor requiere. Es importante aclarar que incorporar una nueva fuente al gestor, no implica necesariamente su visualización en la web de HERA. De hecho, el uso del gestor permite configurar qué información es la que se muestra.

El acceso a la información de la base de datos a incorporar dependerá en gran medida de los medios que ésta brinde. El caso más sencillo y práctico es si dispone de una API, a la cual consultar (p.e., el caso de OpenAlex). En caso negativo, se debería analizar la posibilidad de usar alguna técnica de web scrapping (p.e. el caso de SJR). Por último, si las dos opciones anteriores no fueron viables, se podría incluir un botón para la búsqueda en un sitio externo (p.e. el caso de Google Scholar). Asimismo, el diseño actual de HERA permite considerar e incorporar nuevos mecanismos para acceder a información remota, como por ejemplo consultas GraphQL, uso de API *ad-hoc*, webservices SOAP o incluso protocolos de interoperabilidad como SRU y SRW [6, 7].

## Incorporación de nuevos identificadores

Si bien HERA permite la búsqueda por DOI o Handle, fue diseñado originalmente para soportar cualquier tipo de identificador persistente. Para que esto sea posible, todos los identificadores utilizados deben o bien integrarse con las bases de datos consultadas por HERA como un metadato más, o bien los recursos deberán contar con dicho identificador junto a otro que soporte interoperabilidad. Por ejemplo, es posible extender HERA para dar soporte a identificadores como Microsoft Academic Graph ID (mag), OpenAlex ID, Pubmed Identifier (PMID) y Pubmed Central Identifier (PMCID), ya que todos estos son soportados por la API REST de OpenAlex, lo que permitiría a HERA obtener metadatos de calidad de los recursos desde OpenAlex y utilizar esta información para obtener más información desde otras fuentes. Por otro lado, en el futuro sería posible integrar otros identificadores como ARK o HANDLE.NET a medida que estos sistemas incorporan servicios de interoperabilidad y recuperación de metadatos.

## Aplicaciones externas que se nutran de HERA

La implementación de una API REST en HERA 2.0 permite una mayor interoperabilidad con aplicaciones externas que puedan beneficiarse de sus servicios. La provisión de una interfaz única y estandarizada simplifica la incorporación de la funcionalidad, ya que posibilita que las diferentes bases de datos incluidas en HERA puedan ser consultadas a través de un único medio y de una misma manera. De hecho, la extensión web de HERA es justamente un ejemplo de esta clase de aplicación.

A continuación, se plantean algunas cuestiones vinculadas a la implementación de una plataforma agregadora de revistas a partir de HERA, junto a los posibles beneficios y limitaciones asociados.

### Plataforma *agregadora* de revistas científicas

Una plataforma agregadora de revistas es un espacio web que ofrece acceso a un conjunto de revistas científicas y que brinda servicios de valor agregado sobre las mismas. Si bien la construcción de una plataforma de estas características excede los objetivos planteados en este informe, es posible imaginar un conjunto de desarrollos que recuperen o *cosechen* metadatos de revistas científicas, seleccionadas previamente a partir de determinados criterios (por ej. por la pertenencia a una organización o a un país), utilizando un protocolo de interoperabilidad adecuado para cosechas masivas de metadatos normalizados (por ej. OAI PMH); toda la información recuperada podría entonces ser organizada, clasificada e incorporada en una base de datos centralizada para luego brindar servicios como exploración, búsqueda, visualización o acceso a datos crudos. Una plataforma de estas características puede hacer uso de la API REST de HERA para *enriquecer* los metadatos recuperados de la revista, y esto podría hacerse a través de diferentes estrategias:

- Al visualizar una revista desde la plataforma, se podría enviar una consulta REST a HERA utilizando el ISSN de la revista, y así recuperar información sobre bases de datos donde se encuentra integrada dicha revista e indicadores de visibilidad e impacto que aquellas bases de datos que los ofrecen (por ej. Scopus o WoS)

- Al visualizar un artículo de una revista desde la plataforma, se podría enviar una consulta REST a HERA utilizando el DOI del artículo (en el futuro podrían sumarse otros identificadores), y así recuperar indicadores de visibilidad e impacto de dicho artículo, junto a información sobre la revista en la que fue publicado
- Si la plataforma fuera capaz de generar listados de artículos a partir de determinados criterios (artículos de una revista, artículos de un número particular, artículos de determinado autor o artículos de determinada temática), se podrían recolectarse los identificadores de dichos artículos y enviar a la API de HERA un pedido de generación de un dataset. HERA retornaría entonces un archivo (por ej. CSV o JSON) en el que se integre, para cada artículo, tanto sus metadatos como indicadores de visibilidad e impacto recolectados por HERA. De esta manera, la combinación de la plataforma de revistas con HERA permitiría ofrecer una herramienta de gran utilidad para la elaboración de estudios y análisis sobre cantidad e impacto de la producción científica, medición de citas, evaluación de áreas del conocimiento, transparencia de instituciones académicas, entre otros.

Estos procesos de integración de datos provenientes de HERA junto a datos de la plataforma de revistas podría hacerse en vivo: al momento de realizar una consulta en la plataforma, ésta consulta a HERA e integra los datos a medida que los recibe. Esta solución tiene como principal ventaja que el usuario siempre recibirá datos actualizados, ya que HERA consultará sus fuentes en el mismo momento que el usuario realizó su consulta. Sin embargo, también presenta algunas desventajas que deben ser tenidas en cuenta. Una de las principales desventajas será un mayor tiempo de respuesta para el usuario final, ya que HERA deberá esperar que todas sus fuentes envíen datos para generar una respuesta a la plataforma de revista. Otra desventaja está asociada a la dependencia de HERA con fuentes externas, ya que en caso que una fuente no responda (por ejemplo porque se alcanzó un límite de consultas, porque está saturada de peticiones o porque se encuentra fuera de servicio), HERA retornará datos incompletos para el recurso solicitado.

Una alternativa a las consultas en vivo podría ser la recolección de datos de manera diferida: la plataforma de revistas puede alojar en su propia base de datos la información que genera HERA, y actualizar esta información de manera periódica (por ejemplo semanalmente). Esto requerirá establecer un esquema de actualizaciones que permita mantener los recursos actualizados sin saturar el servicio que brinda HERA ni los sistemas que HERA consulta; por ejemplo, podría considerarse un modelo escalonado, en el que periódicamente (por ej. todos los días) se selecciona un conjunto diferente de recursos alojados en la plataforma de revistas, que serán enviados a HERA y actualizados con la información recibida. De este modo, la plataforma de revistas sería capaz de generar respuestas de manera mucho más ágil (ya que tanto los metadatos como las métricas estarán alojados localmente) y podrá por lo tanto brindar servicio a un mayor número de usuarios; sin embargo, es importante observar que esta aproximación podría generar información con datos desactualizados, y el grado de actualización de estos datos estará supeditado al período entre actualizaciones de cada recurso.

## Consideraciones finales

El diseño de HERA 2.0 fue pensado de manera modular sobre una arquitectura en capas (*layers*), lo que brinda la flexibilidad necesaria para considerar a futuro muchas más funciones o integraciones de las mencionadas aquí. Sin embargo, es necesario remarcar que actualmente HERA 2.0 no corre sobre una infraestructura adecuada para soportar un alto tráfico, y carece de personal dedicado al mantenimiento y soporte. Para poder brindar un servicio de manera permanente y estable será necesario incorporar recursos humanos que sean capaces de velar por el buen funcionamiento del servicio, realizar ajustes cuando los sistemas externos así lo requieran, incorporar nuevos sistemas de consulta, y desarrollar nuevas funcionalidades y servicios sobre HERA. Asimismo, para permitir un uso intensivo de HERA como el propuesto aquí mediante la integración con una plataforma de revistas, será necesario ampliar la infraestructura tecnológica para soportar una mayor carga. Finalmente, deberá considerarse si los límites impuestos por algunos sistemas de consulta resultan suficientes para brindar un servicio a gran escala, o si sería conveniente la contratación de servicios premium (cuando sea posible) que permitan sortear dichos límites.

## Referencias

- [1] C. Rozemblum, C. Unzurrunzaga, G. Banzato, y C. Pucacco, «Calidad editorial y calidad científica en los parámetros para inclusión de revistas científicas en bases de datos en Acceso Abierto y comerciales», *Palabra Clave*, vol. 4, n.º 2, pp. 64-80, abr. 2015.
- [2] R. Johnson, A. Watkinson, y M. Mabe, «The STM Report. An overview of scientific and scholarly publishing», The International Association of Scientific, Technical and Medical Publishers, 2018.
- [3] R. Repiso, «Cómo identificar una revista de calidad», *Cardiocre*, vol. 50, n.º 2, pp. 46-48, 2015, doi: 10.1016/j.carcor.2014.12.002.
- [4] J. F. Porto, E. Rucci, y G. L. Villarreal, «HERA-Herramienta para Enriquecimiento de Recursos Académicos», en *Actas de la XI Conferencia Internacional de Bibliotecas y Repositorios Digitales (BIREDIAL-ISTEC)*, oct. 2022, pp. 251-266. [En línea]. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/148922>
- [5] J. F. Porto, E. Rucci, y G. L. Villarreal, «HERA, una Herramienta para la Evaluación de Recursos Académicos», en *Actas del XXVIII Congreso Argentino de Ciencias de la Computación (CACIC 2022)*, 2022, pp. 546-557. Accedido: 3 de marzo de 2023. [En línea]. Disponible en: <http://sedici.unlp.edu.ar/handle/10915/149623>
- [6] De Giusti, Marisa R. ; Villarreal, Gonzalo L.; Sobrado, Ariel; Lira, Ariel J.; Vila, María M. «SeDiCI | Servicio de Difusión de la Creación Intelectual», *D-Lib Magazine*, vol 14. n. 9/10, 2008. [En línea]. Disponible en <https://doi.org/10.1045/september2008-villarreal>
- [7] Oviedo, Nestor F.; Lira, Ariel J.; De Giusti, Marisa R.; «Service Cloud for information retrieval from multiple origins», en *International Conference on Engineering Education (ICEE) (Gliwice, Polonia, 2010)*. [En línea]. Disponible en <http://sedici.unlp.edu.ar/handle/10915/5561>