

# Indicadores de medición del acceso abierto: fuentes y herramientas

## Open access indicators: sources and tools

Ángel Borrego

**Borrego, Ángel** (2022). "Indicadores de medición del acceso abierto: fuentes y herramientas". *Anuario ThinkEPI*, v. 16, e16a18.

<https://doi.org/10.3145/thinkepi.2022.e16a18>

Publicado en *IweTel* el 7 de junio de 2022

### Ángel Borrego

<https://orcid.org/0000-0002-6462-3966>

Universitat de Barcelona

Facultat d'Informació i Mitjans Audiovisuals

Centre de Recerca en Inf., Com. i Cultura (CRICC)

Melcior de Palau, 140. 08014, Barcelona (España)

[borrego@ub.edu](mailto:borrego@ub.edu)



**Resumen:** La determinación del porcentaje de publicaciones disponibles en acceso abierto es el indicador más pragmático para medir la concreción de las políticas de ciencia abierta y está presente en la mayoría de los planes en este ámbito. Esta nota ofrece un repaso de la bibliografía publicada en la última década sobre la evolución del acceso abierto a la bibliografía científica y el análisis de algunos indicadores de acceso abierto desarrollados en diferentes países. Los resultados ponen de manifiesto que, al crear un indicador de acceso abierto, es necesario tomar decisiones relativas a la colección documental sobre la cual se realizará el cálculo, la herramienta empleada para determinar la disponibilidad en acceso abierto, la definición y tipologías de acceso abierto consideradas, la posible utilización de fuentes de información adicionales, los aspectos temporales y la publicación de la documentación necesaria para garantizar la reproducibilidad del indicador.

**Palabras clave:** Acceso abierto; *Crossref*; Indicadores; Monitorización; Repositorios; *Unpaywall*.

**Abstract:** Determining the percentage of open access publications available is the most pragmatic indicator for measuring the concreteness of open science policies and can be found in most plans in this field. This note offers a review of the literature published in the last decade about the evolution of open access to the scientific literature and the analysis of open access indicators developed in different countries. The results show that, when creating an open access indicator, it is necessary to make decisions regarding the collection of documents on which the calculation will be carried out, the tool that will be used to determine the availability of open access documents, the definition and typologies of open access being studied, the possible inclusion of additional sources of information, aspects related to timing, and the publication of the documentation necessary to guarantee the reproducibility of the indicator.

**Keywords:** Open access; *Crossref*; Indicators; Monitoring; Repositories; *Unpaywall*.

**Nota:** Esta nota ThinkEPI se basa en el informe *Creació d'un indicador d'accés obert a la producció científica de Catalunya* realizado por encargo del Àrea de Ciència Oberta del Consorci de Serveis Universitaris de Catalunya (CSUC). El informe está disponible en <http://hdl.handle.net/2072/520604>

## 1. Introducción

El acceso abierto es un conjunto de principios y prácticas que han modificado el sistema de comunicación científica en las dos últimas décadas. En respuesta al incremento continuo de los precios de suscripción a las revistas científicas por encima del coste de la vida y las consiguientes dificultades para acceder a los contenidos publicados, el movimiento por el acceso abierto defiende que las publicaciones científicas se tienen que distribuir en línea, sin barreras económicas ni limitaciones en los derechos de copia o reutilización.

Hay dos estrategias para facilitar el acceso abierto a la bibliografía científica:

- el depósito de los textos en repositorios institucionales o temáticos (ruta verde);
- la publicación en revistas de acceso abierto (ruta dorada).

Entre estas últimas, se acostumbra a distinguir las revistas que, para ser viables, cobran a los autores unas tasas de edición (*article processing charges*, APC), de las financiadas por instituciones y sociedades científicas (revistas diamante o platino). Por otro lado, las revistas híbridas son revistas de suscripción que permiten a los autores publicar en acceso abierto sus artículos mediante el pago de unas APC, mientras el acceso al resto de contenidos continúa siendo restringido. Finalmente, el acceso abierto bronce hace referencia a artículos que se pueden consultar gratuitamente en la web pero que no disponen de una licencia que garantice el reaprovechamiento de sus contenidos, de forma que no está claro si estos artículos están disponibles en abierto para su lectura de forma temporal o permanente.

Actualmente, la determinación del porcentaje de publicaciones disponibles en acceso abierto es el indicador más pragmático para medir la concreción de las políticas de ciencia abierta y está presente en la mayoría de los planes en este ámbito, que acostumbran a marcarse como objetivo llegar a un 100% de acceso abierto en un determinado horizonte temporal. Algunos rankings universitarios, como el *CWTS Leiden Ranking*, han introducido indicadores de acceso abierto en paralelo a los de producción, impacto o colaboración para clasificar ordenadamente los centros de enseñanza superior.

<https://www.leidenranking.com>

De entrada, parece que medir la disponibilidad en acceso abierto de un conjunto de publicaciones debería ser un proceso relativamente sencillo. Sin embargo, no siempre es así. Recientemente, el *National Open Research Forum*, creado en Irlanda en 2017 para orientar la agenda nacional hacia la ciencia abierta, alertaba de la carencia de datos para medir la disponibilidad de publicaciones en acceso abierto:

“No disponemos de ninguna fuente de datos autorizada sobre el nivel de acceso abierto a nivel nacional. Especialmente problemática es la medición del OA en las Humanidades y las Ciencias Sociales. [...] Se necesitan mejores datos y mucho más análisis para abordar el progreso en relación con esta y otras recomendaciones del *NORF* y para poder proyectar la trayectoria del crecimiento del OA con el fin de establecer un objetivo apropiado para el 100% del OA en todas las disciplinas” (*NORF*, 2021, p. 16-17).

Al comparar los datos de acceso abierto suministrados por diferentes fuentes, es fácil sorprenderse por la diversidad de cifras aportadas. Recientemente, revisando la evolución del acceso abierto en España, **Martínez-Galindo, Rubio y Fernández-Burguete** (2022, p. 3) constataban “las disparidades de los resultados según quien recopila los datos”.

Estas inconsistencias son atribuibles a varios factores:

- colección documental sobre la cual se realiza el cálculo;
- herramienta utilizada para determinar la disponibilidad en acceso abierto;
- definición de acceso abierto empleada; y
- momento de recogida de los datos.

Esta nota describe las metodologías empleadas para realizar estudios y construir indicadores que midan la disponibilidad en acceso abierto de publicaciones científicas. Se pretende poner de manifiesto cómo las diferencias en los métodos empleados pueden dar lugar a discrepancias en los resultados obtenidos. Para ello se revisa en primer lugar la bibliografía publicada en la última década sobre la evolución del acceso abierto a la bibliografía científica. El análisis se centra en los estudios a gran escala que muestran las principales tendencias y las perspectivas de futuro. A continuación, se describen las dos herramientas empleadas en los estudios identificados en el apartado anterior: motores de búsqueda de bibliografía científica y *Unpaywall*, principal instrumento utilizado en trabajos de este tipo. Finalmente, se analizan las características de una serie de indicadores de acceso abierto elaborados en la Unión Europea, Alemania, Dinamarca, España, Francia y Países Bajos. La nota finaliza con algunas recomendaciones.

## 2. Acceso abierto a la bibliografía científica: evolución cronológica y situación actual

Durante la última década, diversos estudios han medido el grado de disponibilidad en acceso abierto de la bibliografía científica, especialmente aquella en forma de artículo de revista. En este apartado revisaremos la metodología y las conclusiones de algunos trabajos que adoptan una perspectiva global, obviando estudios limitados a una disciplina, país o tipología de acceso abierto.

El primer análisis a gran escala fue llevado a cabo por **Archambault et al.** (2014) que, en el marco de un contrato con la *Unión Europea*, crearon un motor de búsqueda para rastrear la disponibilidad gratuita de una muestra de documentos publicados entre 1996 y 2013 indizados en *Scopus*. Los autores midieron lo que **Suber** (2008) bautizó como *gratis open access* —la posibilidad de leer los artículos—, en contraposición al *libre open access* —que incluye la posibilidad de reutilizarlos. Considerando todo tipo de disponibilidad —es decir, incluyendo copias de artículos en redes sociales o en páginas web personales— encontraron un nivel global de acceso abierto del 47%.

**Piwowar, Priem y Orr** (2018) utilizaron *Unpaywall* para medir la disponibilidad en acceso abierto de tres muestras de 100.000 artículos —artículos con DOI registrados en *Crossref*; artículos “recientes” indizados en *Web of Science*; y artículos visualizados por usuarios de *Unpaywall*— y estimaron que al menos el 28% de la bibliografía científica estaba disponible en acceso abierto, si bien esta cifra se elevaba hasta el 45% para los artículos publicados en 2015, último año analizado. En este trabajo se introdujo el concepto de acceso abierto “bronce” para hacer referencia a los artículos “que son de libre lectura en el sitio web del editor, sin una licencia explícitamente abierta”.

Poco después, **Bosman y Kramer** (2018) usaron los datos suministrados por *Unpaywall* a través de *Web of Science* —que, al igual que *Dimensions* o *Scopus*, incorpora un filtro de documentos en acceso abierto basado en *Unpaywall* (**Else**, 2018)— para analizar la disponibilidad de 12,3 millones de artículos y revisiones publicados entre 2010 y 2017. La disponibilidad en abierto llegaba a casi el 30% de los documentos publicados en 2015 y 2016, si bien las cifras de 2017, último año analizado, se reducían. Las razones de esta caída para el período más reciente tenían que ver con el hecho de que las editoriales que ponen contenidos en abierto por la vía bronce acostumbran a hacerlo al menos uno o dos años después de la publicación, un fenómeno similar al que se puede producir con el depósito en repositorios institucionales y temáticos.

Otro estudio de *Science-Metrix* (2018) midió la disponibilidad en acceso abierto de documentos indizados en *Web of Science* y *Scopus* usando la base de datos *1science*. El equipo de trabajo era el del estudio de **Archambault et al.** de 2014 y usó la misma definición de *gratis open access*, es decir, la disponibilidad de una copia gratuita de un artículo científico en la web que se pudiera consultar sin necesidad de autenticación, independientemente de la existencia o no de una licencia de reutilización. Los resultados mostraban que, en el tercer trimestre de 2016, el 50% de los artículos publicados en 2006 indizados en *Web of Science* estaban disponibles en acceso abierto. El porcentaje se incrementaba hasta el 60% para documentos publicados en 2011 y, a partir de entonces, decaía. Los resultados mostraban diferencias por países y disciplinas y, dentro de estas, en las modalidades de acceso abierto más habituales.

Usando una aproximación diferente, **Martín-Martín et al.** (2018) emplearon *Google Scholar* para localizar “copias gratuitas” de casi 2,3 millones de artículos y revisiones con DOI indizados entre 2009 y 2014 en *Web of Science*. Los datos se recogieron el verano de 2016 y mostraban una tasa de disponibilidad del 54,6%. Los documentos estaban disponibles en las webs de las editoriales (23,1% que combinaba acceso abierto dorado, híbrido y bronce), en repositorios (17,6%) y en otras fuentes como *ResearchGate* (40,6%), con solapamientos entre las diversas fuentes.

También en 2018, **Severin et al.** llevaron a cabo un metanálisis que sintetizaba los resultados de once estudios sobre la disponibilidad de la bibliografía científica en acceso abierto. Los resultados ponían de manifiesto la diversidad de cifras obtenidas en estos trabajos. El artículo concluía (p. 17) que “los niveles estimados de OA para los años de publicación posteriores a 2010 variaron entre el 29,4% [cifra ofrecida en el estudio de **Bosman y Kramer** para el año 2016] y el 66% [cifra aportada por **Larivière y Sugimoto** (2018) en un trabajo sobre la disponibilidad en acceso abierto de las publicaciones financiadas entre 2009 y 2017 por doce agencias con mandatos de publicación en acceso abierto], y la mayoría de los estudios informaron de que los niveles de OA se encontraban entre el 50% y el 60%”. Estas discrepancias obedecerían al hecho de que algunos de los estudios que mostraban las tasas más elevadas de disponibilidad en acceso abierto incluían copias disponibles en redes sociales y webs personales bajo la categoría de acceso abierto verde.

Un estudio posterior (**Piwowar; Priem; Orr**, 2019) combinó datos sobre la disponibilidad en acceso abierto de 70 millones de documentos publicados entre 1950 y 2019 con datos de 2,8 millones de con-

sultas de la extensión de navegador de *Unpaywall* durante el mes de julio de 2019. En lugar de medir la disponibilidad en acceso abierto de los documentos en función de su fecha de publicación —lo que hace la mayor parte de la bibliografía— los autores de este trabajo optaban por intentar determinar la disponibilidad en acceso abierto de la bibliografía en una fecha de observación. Los resultados mostraban que en 2019, el 31% de la bibliografía científica analizada estaba disponible en acceso abierto. Estos artículos atraían el 52% de las visualizaciones. A partir de un análisis de tendencias, los autores estimaban —en un modelo que calificaban de conservador— que, en 2025, el 44% de los artículos estará disponible en acceso abierto, acaparando el 70% de las visualizaciones.

**Robinson-Garcia, Costas y Van-Leeuwen (2020)** analizaron la disponibilidad en acceso abierto de 4,6 millones de documentos indizados en *Web of Science* procedentes de 963 universidades. Usando *Unpaywall*, observaron que el 77% de los documentos estaban disponibles en acceso abierto verde, el 33% en dorado, el 20% en bronce y el 16% en híbrido, de nuevo con un elevado grado de solapamiento entre la primera ruta y las otras tres.

Más recientemente, **Basson et al. (2022)** estudiaron las diferencias entre medir la disponibilidad de la bibliografía en acceso abierto a partir de *Dimensions* o de *Web of Science*. A pesar de que ambas bases de datos utilizan *Unpaywall* para determinar si un documento está en abierto o no, los resultados muestran que la proporción de documentos en abierto es mayor en *Dimensions*. Esta discrepancia afecta especialmente a las revistas editadas fuera de Europa y de Norteamérica, que acostumbran a estar recogidas en *Dimensions* (que utiliza *Crossref* como fuente), pero no en *Web of Science*.

En conclusión, tal como anunciaba la revisión sistemática de **Severin et al. (2018)**, se observa una elevada disparidad en las cifras de prevalencia del acceso abierto ofrecidas por diferentes estudios. Estas oscilaciones pueden atribuirse a diferentes factores. En primer lugar, en cuanto a las colecciones documentales analizadas, se observa una elevada coincidencia en la utilización de *Scopus* y *Web of Science* para identificar la producción. No obstante, a menudo hay diferencias en la ventana temporal entre la fecha de publicación de los documentos y el momento de recogida de los datos sobre disponibilidad. La mayor parte de los estudios incluyen un análisis longitudinal para determinar el grado de consolidación del acceso abierto. A pesar de que la tendencia es al incremento progresivo de la disponibilidad, para períodos muy recientes el porcentaje de documentos en acceso abierto se reduce. La razón es que hay artículos que se publican en cerrado pero que, al superar un período de embargo, son liberados por las editoriales o depositados en repositorios. Otro elemento que varía de un estudio a otro es la herramienta utilizada para determinar la disponibilidad en acceso abierto. En líneas generales, se observan dos estrategias que describiremos con más detalle en el siguiente apartado:

- utilizar un motor de búsqueda para localizar copias de los artículos;
- recurrir a *Unpaywall*.

Esta elección incide sobre un tercer elemento: la definición de acceso abierto empleada en los diversos estudios. Los motores de búsqueda localizan copias gratuitas disponibles en páginas web personales o redes sociales. A menudo, estas copias no incluyen una licencia de reutilización, por lo que no cumplen los requisitos de una definición estricta de acceso abierto como la que puede usar *Unpaywall*, que no cosecha información de copias en páginas personales o redes sociales.

### 3. Herramientas para determinar la disponibilidad en acceso abierto

Como hemos visto en el apartado anterior, la mayor parte de los trabajos que miden la prevalencia del acceso abierto usan, o bien un motor de búsqueda de bibliografía científica, o bien *Unpaywall*.

#### 3.1. Motores de búsqueda

Algunos estudios usan un motor diseñado *ex profeso* (**Archambault et al., 2014; Science-Metrix, 2018**), mientras que otros utilizan *Google Scholar* (**Martín-Martín et al., 2018**). En el primer caso, la herramienta empleada en esos trabajos dio lugar a la creación del motor de búsqueda *1findr*, que fue adquirido por *Elsevier* y finalmente abandonado.

[https://www.1science.com/1findr\\_public\\_launch](https://www.1science.com/1findr_public_launch)

Por su parte, *Google Scholar* es un motor que indiza la bibliografía científica que rastrea en la Web, incluyendo contenidos que requieren una suscripción, publicaciones en acceso abierto y copias en repositorios, páginas personales o redes sociales. Al margen de la dificultad de clasificar los documentos recuperados a través de *Google Scholar* como de acceso abierto según la licencia de que disponen, el principal obstáculo para su utilización en estudios de este tipo viene dado por la imposibilidad de replicar los resultados obtenidos debido a la laboriosidad del proceso. Tal como afirman **Martín-Martín et al.:**

“Quizás una de las limitaciones más importantes de este estudio es que no es fácilmente replicable debido a las limitaciones en la extracción de datos impuestas por *Google Scholar*. Extraer una



gran cantidad de datos de esta fuente sólo es posible si se está dispuesto a dedicar una cantidad de tiempo desmesurada a la tarea” (Martín-Martín et al., 2018, p. 835).

### 3.2. Unpaywall

*Unpaywall* es una herramienta creada por la organización sin ánimo de lucro *Our Research*, previamente denominada *ImpactStory*. Indiza contenidos publicados en revistas de acceso abierto y disponibles en repositorios. Hay que remarcar que únicamente recopila documentos que tengan un DOI y que estén disponibles en fuentes “legítimas”, obviando copias en *ResearchGate*. Según la información proporcionada en su web, en mayo de 2022 tenía indizados cerca de 32 millones de documentos en acceso abierto. El sistema se puede implementar como una extensión del navegador que indica al usuario si hay disponible una copia en acceso abierto del documento que está consultando en pantalla. Asimismo, dispone de una API pública.

*Unpaywall* clasifica los artículos en alguna de las siguientes cinco categorías: “cerrado”, “verde”, “dorado”, “híbrido” o “bronce”. La distinción entre “híbrido” y “bronce” radica en que los artículos clasificados en la primera categoría cuentan con una “licencia abierta” (*open license*) que los creadores de *Unpaywall* definen de forma laxa. Como la herramienta informa sobre el tipo de licencia de cada artículo, si el usuario lo desea puede establecer criterios de clasificación más estrictos:

“tendemos a errar en la dirección de la mayor inclusión - en otras palabras, nuestro límite para “es esta licencia abierta” es muy bajo. Incluimos incluso licencias relativamente restrictivas (por ejemplo, CC-BY-NC-ND), y dejamos que los usuarios posteriores las filtren si lo desean.” (Orr, 2021).

Dado que un documento puede estar disponible en acceso abierto a través de diferentes vías, *Unpaywall* identifica una de ellas como la “mejor” (*best\_oa\_location*). Por ejemplo, la versión disponible en la web de la editorial es preferida a la depositada en un repositorio.

Algunos trabajos han ofrecido datos sobre la cobertura y fiabilidad de *Unpaywall*. Piwowar et al. (2018) seleccionaron una muestra de 500 artículos y los buscaron manualmente en Internet. Los resultados mostraban una precisión de la herramienta del 96,6%. Es decir, prácticamente siempre que *Unpaywall* determinaba que un artículo estaba en acceso abierto, efectivamente lo estaba. La exhaustividad era inferior: *Unpaywall* era capaz de encontrar el 77% de los artículos que estaban disponibles en acceso abierto.

## 4. Indicadores de acceso abierto

En este apartado se analizan, desde un punto de vista eminentemente metodológico, las características de varios indicadores nacionales sobre la disponibilidad de bibliografía científica en acceso abierto. Empezaremos con la descripción de dos instrumentos con cobertura internacional y, a continuación, analizaremos los indicadores disponibles en Alemania, Dinamarca, España, Francia y Países Bajos.

### 4.1. Iniciativas internacionales

Unión Europea, *Open Science Observatory*  
<https://osobservatory.openaire.eu>

A nivel europeo, el *Open Science Observatory* hace un seguimiento de la disponibilidad en acceso abierto de publicaciones, *datasets*, software y otros resultados de la investigación de autores afiliados a instituciones de los países miembros. La metodología está descrita en un informe sobre el seguimiento del acceso abierto en el programa *Horizon 2020* (*European Commission*, 2021). Para la compilación de las publicaciones se usan múltiples bases de datos, principalmente abiertas:

“Por primera vez, se consideraron las fuentes de datos abiertos como fuentes primarias para dicho seguimiento (*OpenAire*, *Unpaywall*, *CrossRef*, *OpenAPC*, *DataCite*, *Orcid*, *DOAJ*, *re3data*, por citar algunas). A continuación, se validaron con bases de datos propietarias (*Scopus* y *WoS*) como fuentes secundarias, cuando fue necesario”.

Australia, *Curtin Open Knowledge Initiative*  
<https://open.coki.ac>

La *Curtin Open Knowledge Initiative* (COKI) ofrece un *Open Access Dashboard* que reporta datos de acceso abierto por países e instituciones. Según la metodología, los datos provienen de *Crossref*, *Microsoft Academic Graph* —si bien este proyecto dejó de estar operativo al acabar 2021—, *Unpaywall*, *Research Organization Registry* (ROR) y *Open Citations*.

## 4.2. Alemania

*Open Access Monitor*

<https://open-access-monitor.de>

En Alemania, *Open Access Monitor* es un proyecto desarrollado por el centro de investigación *Forschungszentrum Jülich*. Las publicaciones se identifican mediante búsquedas en *Dimensions*, *Scopus* y *Web of Science*. La determinación de la disponibilidad en acceso abierto se realiza a través de *Unpaywall*, aunque cabe señalar que los documentos en acceso abierto “bronce” se clasifican como “cerrados”. El portal incluye datos sobre gasto en APCs y de citación de los documentos. Los datos se actualizan semanalmente.

## 4.3. Dinamarca

*Open Access Indicator*

<https://www.oaindikator.dk/en>

El *Danish Open Access Indicator* tiene por finalidad hacer el seguimiento del grado de cumplimiento de la *Danish Open Access Strategy 2018-2025*, que se marca como objetivo llegar al 100% de acceso abierto en 2025 con un máximo de doce meses de embargo. El indicador toma como punto de partida las publicaciones incluidas en los sistemas de información de las ocho universidades danesas. Estos datos se complementan con otras fuentes de información:

- *DOAJ*, para conocer las características de las revistas de acceso abierto;
- *Sherpa Romeo*, para averiguar las políticas editoriales;
- sendos listados de repositorios validados y revistas que imponen un período de embargo superior a los doce meses que marca la estrategia danesa.

En cuanto a las rutas de acceso abierto, el indicador contempla el acceso abierto “dorado” (en revistas de acceso abierto) y “verde” (en repositorios, diferenciando los repositorios de las universidades danesas de otros repositorios externos validados), pero no considera el acceso abierto “híbrido”, a no ser que los artículos en cuestión estén también en un repositorio.

El indicador dispone de una detallada descripción metodológica. Destacaremos tres aspectos:

- el indicador se publica anualmente, en el mes de abril, permitiendo el embargo de doce meses previsto en la estrategia danesa. Existe un calendario que detalla las etapas de construcción del indicador: recogida de datos, audiencia a las universidades, ajustes, etc.;
- el indicador no solo mide el acceso abierto logrado, sino también el acceso abierto “potencial”, es decir, el porcentaje de documentos que podrían estar disponibles en abierto a través de repositorios cumpliendo las políticas editoriales de las revistas en las cuales se han publicado;
- para garantizar la transparencia y la reproductibilidad del indicador, es posible descargar los *datasets* con los datos individuales de todos los artículos considerados para el cálculo.

Fruto de este minucioso procedimiento metodológico, el indicador ha sido señalado como un ejemplo de buena práctica en el informe de *NORF* (2021).

## 4.4. España

En España no existe un indicador de acceso abierto similar a los descritos para otros países. No obstante, *Rebiun* elaboró unas recomendaciones sobre la monitorización del acceso abierto (*Rebiun*, 2019a) a partir de las cuales preparó un informe sobre la disponibilidad en acceso abierto de la producción de las universidades españolas y del *CSIC* (*Rebiun*, 2019b). El estudio se basa en la identificación de las publicaciones con DOI indizadas en *Scopus* y *Web of Science* para comprobar, a través de *Unpaywall*, su disponibilidad en abierto.

En Cataluña coexisten dos indicadores:

- El *Observatorio del Acceso Abierto*, alojado en la web de la *Universitat Politècnica de Catalunya*, tiene por objetivo “hacer el seguimiento del estado del acceso abierto en las universidades catalanas”.

<https://bibliotecnica.upc.edu/observatori>

El fondo documental sobre el cual se lleva a cabo este seguimiento resulta de la combinación de los registros con DOI extraídos de *Scopus* y *Web of Science*. Para determinar la disponibilidad en acceso abierto de las publicaciones se utiliza *Unpaywall* (**Rovira; Labastida**, 2019) considerando cuatro categorías: dorado, híbrido, bronce y verde. Los resultados se actualizan dos veces al año, los meses de abril y octubre, y se anuncia que, en el futuro, el Observatorio “presentará los costes de los investigadores y de las instituciones a la hora de publicar en abierto”.

- El *Portal de la Recerca de Catalunya* tiene como finalidad visualizar y difundir la actividad investigadora llevada a cabo en Cataluña. Para ello, se aprovecha el trabajo llevado a cabo por las universidades y centros de investigación en sus sistemas de gestión de la investigación (CRIS) de los cuales se nutre el *Portal*. En su apartado de estadísticas, el *Portal* dispone de una sección sobre acceso abierto, con datos globales y por institución.

<https://portalrecerca.csuc.cat/estadistiques/publicacions/accesobert>

Para determinar el porcentaje de publicaciones en acceso abierto, se extrae la URL de las publicaciones que disponen de ella. Si la URL es un DOI, se ejecuta en *Unpaywall*. En el resto de los casos, se determina si se trata de un dominio validado como de acceso abierto por el *Grupo de Trabajo de Apoyo a la Investigación del CSUC*.

#### 4.5. Francia

*Baromètre français de la Science Ouverte*

<https://barometredelascienceouverte.esr.gouv.fr>

A pesar de que su nombre sugiere un abanico más amplio de indicadores, el *Baromètre français de la Science Ouverte* presenta primordialmente información sobre la disponibilidad de las publicaciones en acceso abierto. A partir de la recopilación de publicaciones de diferentes fuentes abiertas, se verifica su disponibilidad en acceso abierto mediante *Unpaywall*.

#### 4.6. Países Bajos

*Monitor*

<https://www.openaccess.nl/en/in-the-netherlands/monitor>

Entre los años 2016 y 2017, la asociación de universidades neerlandesas (VSNU) llevó a cabo un estudio sobre la disponibilidad en acceso abierto de la producción científica del país. Estos datos han sido actualizados en sucesivas ocasiones para ofrecer algunas estadísticas, no muy detalladas, especialmente de acceso abierto híbrido.

También en los Países Bajos, el portal *Narcis* ofrece estadísticas sobre el porcentaje de documentos en abierto, cerrados, de acceso restringido (*restricted*) y embargados (*embargoed*) disponibles desde 2000. Desgraciadamente no hay mucha información sobre el alcance de cada categoría ni sobre su cálculo (por ejemplo, no queda clara la distinción entre *restricted* y *embargoed*).

### 5. Conclusiones

Los apartados anteriores han puesto de manifiesto el creciente interés de agencias de financiación y de instituciones de investigación por medir la evolución de la disponibilidad de las publicaciones en acceso abierto. Fruto de la necesidad de homogeneizar estos procesos para evitar discrepancias como las que venimos de describir, *Science Europe*, una asociación que reúne a una treintena de entidades que financian o llevan a cabo investigación científica, publicó en mayo de 2021 un conjunto de recomendaciones sobre la medición de la disponibilidad de publicaciones en acceso abierto (Philipp et al., 2021). Al hacer un ejercicio de monitorización habría que diferenciar tres etapas:

- definir el propósito del indicador —¿por qué se quiere medir el nivel de acceso abierto?—;
- identificar el conjunto de publicaciones objeto de análisis;
- fijar el proceso de recogida, interpretación y presentación de los datos.

A pesar de que los objetivos de instituciones diversas puedan diferir, habitualmente hay un interés por comparar entidades que requiere de cierta colaboración y estandarización:

“Aunque el propósito, el calendario y otros parámetros pueden diferir entre las organizaciones interesadas en monitorizar el acceso abierto, una implementación y colaboración alineadas pueden mejorar el impacto que tendrán estos esfuerzos. Esto no sólo ahorraría recursos, sino que también contribuiría a las prácticas sostenibles en general, y facilitaría las inevitables comparaciones entre los resultados de los ejercicios de monitorización del acceso abierto” (Philipp et al., 2021).

La revisión de las iniciativas internacionales de monitorización del acceso abierto muestra que la creación de un indicador de acceso abierto a la producción científica es una decisión oportuna y útil. Se trata de una iniciativa viable, que no requiere de grandes recursos, y que está en sintonía con lo que se está haciendo en la *Unión Europea* para medir la concreción de las políticas de ciencia abierta.

El análisis de diferentes modelos pone de manifiesto que, al crear un indicador de acceso abierto, es necesario tomar decisiones relativas a la colección documental sobre la cual se realizará el cálculo, la herramienta empleada para determinar la disponibilidad en acceso abierto, la definición y tipologías

de acceso abierto consideradas, la posible consideración de fuentes de información adicionales, los aspectos temporales y la publicación de la documentación necesaria para garantizar la reproductibilidad del indicador.

## 6. Referencias

**Archambault, Éric; Amyot, Didier; Deschamps, Philippe; Nicol, Aurore; Provencher, Françoise; Rebut, Lise; Roberge, Guillaume** (2014). *Proportion of open access papers published in peer-reviewed journals at the European and world levels—1996–2013*. European Commission.

[https://science-metrix.com/sites/default/files/science-metrix/publications/ld\\_1.8\\_sm\\_ec\\_dg-rtd\\_proportion\\_oa\\_1996-2013\\_v11p.pdf](https://science-metrix.com/sites/default/files/science-metrix/publications/ld_1.8_sm_ec_dg-rtd_proportion_oa_1996-2013_v11p.pdf)

**Basson, Isabel; Simard, Marc-André; Ouangré, Zoé-Aubierge; Sugimoto, Cassidy R.; Larivière, Vincent** (2022). "The effect of data sources on the measurement of open access: A comparison of Dimensions and the *Web of Science*". *PLoS one*, v. 17, n. 3, e0265545.

<https://doi.org/10.1371/journal.pone.0265545>

**Bosman, Jeroen; Kramer, Bianca** (2018). "Open access levels: a quantitative exploration using *Web of Science* and oaDOI data". *PeerJ preprints*, v. 6, e3520v1.

<https://doi.org/10.7287/peerj.preprints.3520v1>

**Else, Holly** (2018). "How *Unpaywall* is transforming open science". *Nature*, v. 560, pp. 291-292.

<https://doi.org/10.1038/d41586-018-05968-3>

European Commission (2021). *Monitoring the open access policy of Horizon 2020*.

<https://doi.org/10.2777/268348>

**Larivière, Vincent; Sugimoto Cassidy R.** (2018). "Do authors comply when funders enforce open access to research?". *Nature*, v. 562, n. 7728, pp. 483-486.

<https://doi.org/10.1038/d41586-018-07101-w>

**Martín-Martín, Alberto; Costas, Rodrigo; Van-Leeuwen, Thed; Delgado-López-Cózar, Emilio** (2018). "Evidence of open access of scientific publications in *Google Scholar*: A large-scale analysis". *Journal of informetrics*, v. 12, n. 3, pp. 819-841.

<https://doi.org/10.1016/j.joi.2018.06.012>

**Martínez-Galindo, Francisco-Jesús; Rubio, Francisco; Fernández-Burquete, Sergio** (2022). "Monitorización de los mandatos de acceso abierto a través de repositorios institucionales". *Profesional de la información*, v. 31, n. 2.

<https://doi.org/10.3145/epi.2022.mar.04>

NORF (2021). *National open research landscape report*.

<http://bit.ly/NORFLandscape>

**Orr, Richard** (2021). *What is an OA license?*

<https://support.unpaywall.org/support/solutions/articles/44002063718-what-is-an-oa-license->

**Philipp, Tobias; Botz, Georg; Kita, Jean-Claude; Richards, Paul; Sängler, Astrid; Siegert, Olaf; Reumaux, Mathilde** (2021). *Open access monitoring: Guidelines and recommendations for research organisations and funders*.

<https://doi.org/10.5281/zenodo.4905553>

**Piowar, Heather; Priem, Jason; Larivière, Vincent; Alperin, Juan-Pablo; Matthias, Lisa; Norlander, Bree; Farley, Ashley; West, Jevin; Haustein, Stefanie** (2018). "The state of OA: a large-scale analysis of the prevalence and impact of Open Access articles". *PeerJ*, v. 6, e4375.

<https://doi.org/10.7717/peerj.4375>

**Piowar, Heather; Priem, Jason; Orr, Richard** (2019). "The future of OA: A large-scale analysis projecting open access publication and readership". *BioRxiv*, 795310.

<https://doi.org/10.1101/795310>

Rebiun (2019a). *Recomendaciones sobre la monitorización del acceso abierto*. Rebiun Línea 3. Grupo de repositorios.

<http://hdl.handle.net/20.500.11967/245>

Rebiun (2019b). *Medición del acceso abierto en las universidades españolas y el CSIC (2014-2018)*. Rebiun Línea 2. Grupo de acceso abierto.

<http://hdl.handle.net/20.500.11967/354>

**Robinson-García, Nicolas; Costas, Rodrigo; Van-Leeuwen, Thed N.** (2020). "Open Access uptake by universities worldwide". *PeerJ*, v. 8, e9410.

<http://doi.org/10.7717/peerj.9410>

**Rovira, Anna; Labastida, Ignasi** (2019). "The open access observatory". *Liber quarterly*, v. 29, n. 1.

<http://doi.org/10.18352/lq.10295>



Science-Metrix (2018). *Analytical support for bibliometric indicators: Open access availability of scientific publications*.  
[https://www.science-metrix.com/sites/default/files/science-metrix/publications/science-metrix\\_open\\_access\\_availability\\_scientific\\_publications\\_report.pdf](https://www.science-metrix.com/sites/default/files/science-metrix/publications/science-metrix_open_access_availability_scientific_publications_report.pdf)

Severin, Anna; Egger, Matthias; Eve, Martin-Paul; Hürlimann, Daniel (2018). "Discipline-specific open access publishing practices and barriers to change: an evidence-based review". *F1000Research*, v. 7.  
<https://doi.org/10.12688/f1000research.17328.2>

Suber, Peter (2008). "Gratis and libre open access". *SPARC Open access newsletter*, n. 124.  
[https://dash.harvard.edu/bitstream/handle/114322580/suber\\_oagratias.html](https://dash.harvard.edu/bitstream/handle/114322580/suber_oagratias.html)

Ángel Borrego  
 Universitat de Barcelona  
 borrego@ub.edu



## Barreras en la información sobre acceso abierto: el caso de las tesis

Francisco Martínez-Galindo



Agradezco a Ángel Borrego el ThinkEPI dedicado a "Indicadores de medición del acceso abierto: fuentes y herramientas". Me disculpo por tardar tanto en comentarlo. De hecho el borrador con novedades sobre monitorización, que nunca envié, lo he desechado tras ver el tweet de Ángel Delgado con la novedad sobre el acceso a la base de datos Teseo:  
<https://twitter.com/amdelvaz/status/1545109806230474752>

La "mejora" de consulta de Teseo consiste en incluir un *captcha* para identificar que quien consulta es una persona y no un robot. En una base de datos cuya usabilidad es bastante mejorable, ¿de verdad eso es lo prioritario en lugar de implementar un sistema que permita el acceso y la reutilización de la información pública sin

obstáculos? Justo todo lo contrario de las tendencias que, por ejemplo, obligan a que los repositorios que quieren cumplir con el *Plan S* deben poner sus metadatos bajo dominio público, al igual que solicita *Europeana* con los metadatos de los objetos que se envían para su difusión.

Es verdad que vivimos en burbujas, y en mi burbuja nos dedicamos a, entre otras cosas, apoyar el acceso abierto, ayudando a monitorizar los cumplimientos de los mandatos como la *Ley de la ciencia*, por ejemplo, con la implementación de acceso a APIs en nuestros sistemas (repositorios y CRIS) como la de *Recolecta*, que permitan cumplir con las "Instrucciones para la comunicación y publicidad de las ayudas a la I+D+I concedidas por la Agencia Estatal de Investigación. 4 de julio de 2022" (AGI, 2022):

"En el caso de que existan limitaciones de espacio, particularmente en publicaciones en revistas, se mencionará: Ayuda Referencia del/de la proyecto/ayuda financiado/a por MCIN/AEI/10.13039/501100011033 y según proceda, por "Feder. Una manera de hacer Europa", por la "Unión Europea" o por la "Unión Europea NextGenerationEU/PRTR" (AGI, 2022, p. 4).

educacion.gob.es

Está usted en: Portada > Universidades > Educación Superior Universitaria

Tesis doctorales: TESEO

Consulta de la Base de datos de Tesis Doctorales (TESEO)

Rellene los campos del siguiente formulario con las palabras o frases que conozca de los documentos que desea localizar.

Búsqueda Avanzada
  Entrar como usuario registrado
  Registrarme como Doctorando

Buscar en TESEO

Título:

Autor:

Universidad:

Curso académico: Desde:

Atención a usuarios:  
 Correo electrónico: [teseo@minuniversidades.com](mailto:teseo@minuniversidades.com)

Versión 4.4  
 © Ministerio de Universidades  
 Aviso legal | Accesibilidad

Recomiendo una lectura interesante de @JaimeObregon para quien no lo siga en Twitter. Se trata de su artículo “Del limitado acceso a las resoluciones judiciales” (**Gómez-Obregón, 2022**).

¿Y que tiene todo esto que ver con el texto de Ángel Borrego? Tiene relación con el hecho de que cuando hablamos de monitorización del acceso abierto casi siempre se nos olvidan las tesis. El Real Decreto 99/2011, por el que se regulan las enseñanzas oficiales de doctorado, indica en su artículo 14.5:

“Una vez aprobada la tesis doctoral, la universidad se ocupará de su archivo en formato electrónico abierto en un repositorio institucional” (*España, 2011*).

Las tesis de Teseo cuentan con un enlace al repositorio, por lo menos así funciona en nuestra institución. Pero consultar Teseo para comprobar el cumplimiento de este *Real Decreto* no era fácil, y ahora lo será menos. Si alguien está interesado en el tema el único intento que conozco es la presentación: **Barrueco-Cruz, García-Testal y Llamas-Fernández (2019)**.

Aprovecho para recomendar, para quien no los haya visto, los seminarios del Aula Jordi Rubió 2021-2022, y en especial el 3º: *Repositoris: integració en l'ecosistema de ciència oberta*, donde hablamos, entre otras muchas cosas interesantes, sobre la necesidad de hacer seguimiento de cumplimiento de los proyectos.

<https://www.youtube.com/watch?v=T3z52wWJvqY&t=6706s>

Sólo me queda agradecer de nuevo a Ángel Borrego por sacar el tema de la monitorización en un *ThinkEPI*. El borrador original de este texto pretendía comentar el monitor holandés de acceso abierto y cómo la ley de derechos de autor holandesa ofrece una posibilidad adicional de publicar obras breves de ciencia en acceso abierto en los repositorios universitarios, seis meses después de la primera publicación en línea mediante el proyecto “¡Tú compartes, nosotros nos encargamos!”. Al respecto cabe recomendar la lectura del artículo de Sondervan et al. (2021).

## Referencias

AGI (2022). *Instrucciones para la comunicación y publicidad de las ayudas a la I+D+I concedidas por la Agencia Estatal de Investigación. 4 de julio de 2022*. Agencia Estatal de Investigación.

[https://www.aei.gob.es/sites/default/files/page/field\\_file/2022-07/20220704\\_Instrucciones-comunicacion-publicidad-ayudas-concedidas.pdf](https://www.aei.gob.es/sites/default/files/page/field_file/2022-07/20220704_Instrucciones-comunicacion-publicidad-ayudas-concedidas.pdf)

**Barrueco-Cruz, José-Manuel; García-Testal, Cristina; Llamas-Fernández, Emilia** (2019). “Análisis del grado de cumplimiento del mandato de depósito de tesis doctorales en universidades públicas españolas”. En: *18º Workshop de Rebiun de proyectos digitales y 8as Jornadas de OS repositorios: ciencia abierta - ecos, retos y oportunidades de los PlaneS*.

<http://hdl.handle.net/20.500.11967/408>

*España* (2011). “Real Decreto 99/2011, de 28 de enero, por el que se regulan las enseñanzas oficiales de doctorado”. *BOE*, n. 35, 10 febrero.

<https://www.boe.es/buscar/act.php?id=BOE-A-2011-2541>

**Gómez-Obregón, Jaime** (2022). “Del limitado acceso a las resoluciones judiciales”. *Hay derecho*, 7 julio.

<https://www.hayderecho.com/2022/07/07/del-limitado-acceso-a-las-resoluciones-judiciales/>

**Sondervan, Jeroen; Schalken, Arjan; De-Boer, Jan; Woutersen-Windhouwer, Saskia** (2021). “Sharing published short academic works in institutional repositories after six months: The implementation of the article 25fa (Taverne Amendment) in the Dutch Copyright Act”. *Liber quarterly: The Journal of the Association of European Research Libraries*, v. 31, n. 1.

<https://doi.org/10.53377/lq.10915>

**Francisco Martínez-Galindo**  
Universitat Politècnica de València  
[pacomar@upv.es](mailto:pacomar@upv.es)



## La marginación de tesis y literatura gris en la monitorización sobre acceso abierto

Ángel Borrego

Muchas gracias, Paco, por tu comentario. Efectivamente, las tesis doctorales —y, en general, la literatura gris— son las grandes olvidadas en los ejercicios de monitorización del acceso abierto. Hay un trabajo de **Schöpfel** y **Prost** (2019) que explica algunas razones de la exclusión de la literatura gris del *Open Science Monitor* de la UE:

- diversidad de formatos documentales;
- pluralidad de lenguas;
- ausencia de una fuente de información más o menos universal como puede ser *Scopus* (que es la fuente que utiliza la UE) para los artículos de revista.

Otro problema es que la literatura gris no acostumbra a tener DOI, mientras que la mayor parte de indicadores de monitorización se basan en los datos de *Unpaywall* que solo considera los documentos con DOI.

Para ampliar las miras en la monitorización del acceso abierto, en vez de índices de citas como *Web of Science* o *Scopus*, se podrían utilizar como punto de partida los CRIS, que deberían ser exhaustivos en la recolección de la producción bibliográfica de la institución y estar conectados con los repositorios para su utilización en la monitorización, tal como proponéis en vuestro artículo en *EPI* (**Martínez-Galindo; Rubio; Fernández-Burguete**, 2022).

### Referencias

**Martínez-Galindo, Francisco-Jesús; Rubio, Francisco; Fernández-Burguete, Sergio** (2022). "Monitorización de los mandatos de acceso abierto a través de repositorios institucionales". *Profesional de la información*, v. 31, n. 2, e310204.

<https://doi.org/10.3145/epi.2022.mar.04>

**Schöpfel, Joachim; Prost, Hélène** (2019). "The scope of open science monitoring and grey literature". In: *12<sup>th</sup> Conference on grey literature and repositories*, 17 October 2019, Prague, Czech Republic.

<https://hal.archives-ouvertes.fr/hal-02300020>

Ángel Borrego  
Universitat de Barcelona  
borrego@ub.edu



## Algunas iniciativas para una mejor monitorización del acceso abierto a las tesis

Pablo de Castro



Gracias Ángel y Paco por 'revivir' la discusión en torno a este *ThinkEPI* sobre monitorización del acceso abierto. Al hilo de los comentarios sobre las tesis y la posible homogeneización de la recopilación de metadatos sobre la producción académica, en la conferencia CRIS2022 del pasado mayo en Dubrovnik se presentó el portal de investigación danés *Danmarks Forskningsportal*, que tiene un enlace al indicador de acceso abierto que se menciona en la nota *ThinkEPI* original.

<https://forskningsportal.dk>

<https://oaindicator.dk/en/>

Lo novedoso de este portal nacional es que además de las fuentes bibliográficas tradicionales (*Scopus*, *WoS* y *Dimensions*) tiene una opción adicional "*Data from local systems*" que ofrece la opción de comparar la cobertura de estas diferentes fuentes. Esta opción extra de 'datos locales' es factible porque el panorama CRIS en el país es muy homogéneo, con *Pure* (danés en su origen) implantado como CRIS institucional en prácticamente todas las universidades e instituciones de investigación del país. Los repositorios, con su interoperabilidad basada en OAI-PMH, son una herramienta muy valiosa para la agregación de estos metadatos, pero los formatos de metadatos *DublinCore* y familia son un tanto limitados a nivel descriptivo. *Cerif* es mejor opción, aunque la integración es mucho más difícil de lograr en particular en panoramas CRIS fragmentados.

<https://eurocris.org/services/main-features-cerif>

Hay diversos proyectos adicionales de monitorización en Suiza (NOAM, <https://consortium.ch/monitoring/>) y el Reino Unido (*Jisc Monitor*) actualmente en desarrollo, que ofrecen una interesante oportunidad para que la academia retome el control de unos metadatos de publicaciones que son suyos en origen por mucho que las circunstancias lleven a que se recompen a iniciativas comerciales (con cobertura muy variable por ende) a precios no particularmente asequibles. Dado el enorme esfuerzo que las instituciones invierten en recopilar bibliografías de su producción científica y los archivos de acceso abierto asociados, esta externalización sistemática de las fuentes de datos resulta un tanto inexplicable (como tantas otras cosas en el *scholarly comms landscape*). Desconozco si el proyecto *Hércules* de la CRUE y la *Universidad de Murcia* puede tener previsto implantar este nivel de interoperabilidad para fuentes locales, pero a priori podría ser una opción.  
<https://tic.crue.org/hercules/>

La simplificación a la hora de producir DOIs que ha supuesto la infraestructura de *DataCite* hace que muchas instituciones hayan asignado DOIs retrospectivos a sus tesis, lo que debería también facilitar su (eventual) detección via *Unpaywall*. Todo esto apunta a que las condiciones para poder monitorizar el nivel de acceso abierto a tesis doctorales a nivel nacional e incluso europeo están mejorando y pronto podrían ser un ámbito más de los indicadores.

**Pablo de Castro**  
University of Strathclyde  
[pcastromartin@gmail.com](mailto:pcastromartin@gmail.com)

*Te esperamos en*

**www.sedic.es**

c/Rodríguez San Pedro 2,  
oficina 606. 28015 Madrid

Tfno: +34 915 934 059  
[secretaria@sedic.es](mailto:secretaria@sedic.es)

**Sociedad Española de Documentación e Información Científica**

<https://twitter.com/SEDIC20>

<https://www.facebook.com/AsociacionSEDIC>

<https://www.linkedin.com/groups?home=&gid=5060038>