

ARTÍCULOS

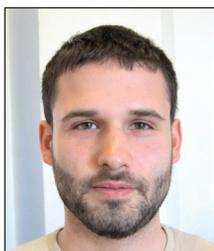


NUEVOS DATOS, NUEVAS POSIBILIDADES: REVELANDO EL INTERIOR DE *ALTMETRIC.COM*



Nicolás Robinson-García, Daniel Torres-Salinas, Zohreh Zahedi y Rodrigo Costas

Note: This article can be read in its original English version on:
<http://www.elprofesionaldelainformacion.com/contenidos/2014/jul/03.pdf>



Nicolás Robinson-García es licenciado en ciencias de la información y máster en información científica por la *Universidad de Granada*. Doctorando y miembro del *Grupo de Investigación EC3 (Evaluación de la Ciencia y de la Comunicación Científica)*. Investiga sobre evaluación de la investigación en el ámbito institucional, y nuevas fuentes de datos para el análisis bibliométrico. Participa en los proyectos *I-UGR Rankings of Spanish universities* y *MetaRanking EC3 de Universidades Españolas*.
<http://orcid.org/0000-0002-0585-7359>

*Evaluación de la Ciencia y de la Comunicación Científica, Departamento de Información y Documentación
Universidad de Granada, EC3 Research Group
Campus Cartuja. 18071 Granada, Spain
elrobin@ugr.es*



Daniel Torres-Salinas, doctorado en documentación científica por la *Universidad de Granada*, trabaja como técnico de gestión en la *Universidad de Navarra*, donde efectúa auditorías del rendimiento científico de la universidad. Investiga en bibliometría como miembro del *Grupo de Investigación EC3*. Desarrolla varias herramientas de evaluación de la investigación, tales como *Cientificacvn*, *I-UGR Rankings of Spanish universities* y *CIRC (Clasificación integrada de las revistas científicas)*. Es co-fundador y CEO de la spin-off consultora científica *EC3metrics*.
<http://orcid.org/0000-0001-8790-3314>

*Universidad de Granada, EC3 Research Group
Campus Cartuja. 18071 Granada, Spain
torressalinas@gmail.com*



Zohreh Zahedi es estudiante de doctorado en el *Centre for Science and Technology Studies (CWTS)* de la *Universidad de Leiden*, Holanda. Su investigación doctoral se centra en el uso de las alométricas como alternativa y nueva fuente de datos para la medición del desempeño de la investigación. Más allá del tema inmediato de su tesis doctoral, también participa en los grupos de trabajo *Career in Policy & Path in Science* y *Societal Impact of Research (SURE)*. En este segundo grupo contribuye en la línea de investigación de alométricas.
<http://orcid.org/0000-0001-5801-1886>

*Centre for Science and Technology Studies, Leiden University
Wassenaarseweg 62, 2333 AK Leiden, The Netherlands
zohrehzahedi@gmail.com*



Rodrigo Costas es investigador senior experto en bibliometría y ciencias de la información. Con un doctorado en biblioteconomía y documentación obtenido en el *CSIC* en España, trabaja en el *Centre for Science and Technology Studies (CWTS)* de la *Universidad de Leiden* desde 2009. Sus líneas de investigación abarcan una amplia gama de temas bibliométricos. Recientemente en el *CWTS* ha comenzado algunas líneas de investigación nuevas como el estudio de las alométricas y de las posibilidades de reconocimiento de la financiación, con el fin de ampliar las posibilidades analíticas de la cienciometría.
<http://orcid.org/0000-0002-7465-6462>

*Centre for Science and Technology Studies, Leiden University
Wassenaarseweg 62, 2333 AK Leiden, The Netherlands
rcostas@cwts.leidenuniv.nl*

Artículo recibido el 11-03-2014
Aceptación definitiva: 03-07-2014

Resumen

Este trabajo analiza *Altmetric.com*, una de las fuentes de datos altmétricos más usadas actualmente. Para ello hemos cruzado un set de publicaciones con doi indexadas en la *Web of Science* para el periodo 2011-2013 con la API de *Altmetric.com*. Sólo el 19% de las publicaciones de nuestro set estaban indexadas en *Altmetric.com*. Este recurso obtiene datos altmétricos de 16 redes sociales distintas. No obstante, cinco de ellas representan el 95,5% del set de datos recuperado. *Twitter* (87,1%) y *Mendeley* (64,8%) cubren un mayor número de publicaciones. Concluimos destacando *Altmetric.com* como una herramienta rica, transparente y precisa en sus datos altmétricos. No obstante, ofrece aún algunas dudas acerca de la exhaustividad de la recuperación así como de la selección de fuentes, que requieren más investigación.

Palabras clave

Altmetric.com, *Twitter*, *Mendeley*, Indicadores altmétricos, Impacto social, Cobertura, Web 2.0, Confiabilidad.

Title: New data, new possibilities: Exploring the insides fo *Altmetric.com*

Abstract

This paper analyzes *Altmetric.com*, one of the most important altmetric data providers currently used. We have analyzed a set of publications with doi number indexed in the *Web of Science* during the period 2011-2013 and collected their data with the *Altmetric* API. 19% of the original set of papers was retrieved from *Altmetric.com* including some altmetric data. We identified 16 different social media sources from which *Altmetric.com* retrieves data. However five of them cover 95.5% of the total set. *Twitter* (87.1%) and *Mendeley* (64.8%) have the highest coverage. We conclude that *Altmetric.com* is a transparent, rich and accurate tool for altmetric data. Nevertheless, there are still potential limitations on its exhaustiveness as well as on the selection of social media sources that need further research.

Palabras clave

Altmetric.com, *Twitter*, *Mendeley*, Altmetrics, Social impact, Coverage, Web 2.0, Reliability.

Robinson-García, Nicolás; Torres-Salinas, Daniel; Zahedi, Zohreh; Costas, Rodrigo (2014). "New data, new possibilities: Exploring the insides of *Altmetric.com*". *El profesional de la información*, July-August, v. 23, n. 4, pp. 359-366.

<http://dx.doi.org/10.3145/epi.2014.jul.03>

Introducción

Tradicionalmente, el análisis de citas ha sido puesto en entredicho con puntos de vista diferentes y opuestos en cuanto a su idoneidad para medir cuantitativamente el "impacto científico" de las publicaciones. Por citar algunos, se le achacan sesgos en la citación, demoras de publicación e incluso limitaciones derivadas del proceso de revisión por pares (Bollen; Van-de-Sompel, 2006). Se han propuesto varias alternativas, especialmente desde los 90s y gracias a la expansión de internet y las ediciones digitales. Entre otros destacamos el uso de reconocimientos o *influmetrics* (Cronin; Weaver, 1995), enlaces web o *webometrics* (Almind; Ingwersen, 1997) y las métricas de uso (Kurtz; Bollen, 2010). Sin embargo, la alternativa más reciente al tradicional análisis de citas se ha convertido en un tema candente en la comunidad bibliométrica: las altmétricas (métricas alternativas) que usan indicadores basados en las redes sociales para cuantificar el impacto de un trabajo académico. Fueron propuestas por primera vez por Priem *et al.* (2010), y desde entonces este tema se ha convertido en un frente de investigación.

Los proponentes de las altmétricas afirman que estos indicadores permiten complementar o mejorar los sistemas de evaluación científica tradicionales (Priem *et al.*, 2010), y que proporcionan una imagen más amplia de la relevancia y el impacto de las contribuciones científicas como "productos resultantes de la investigación" (Piwowar, 2013). Asimismo,

se generan antes que las citas y terminan con el monopolio ejercido por los índices de citas (*WoS*, *Scopus*...), ya que vienen de fuentes abiertas. Es más, su punto más fuerte es que pueden reflejar aspectos del impacto diferentes de los que ofrecen las citas. Sin embargo, la realidad es que están poco desarrolladas y que se necesita mucho más estudio antes de confirmar esos argumentos, que actualmente son cuestionables o sólo simples promesas (Wouters; Costas, 2012).

Los indicadores altmétricos complementan o mejoran los sistemas de evaluación científica tradicionales

Por lo tanto, todavía hay serias dudas sobre el significado de estos indicadores (Torres-Salinas; Cabezas-Clavijo; Jiménez-Contreras, 2013; Torres-Salinas; Cabezas-Clavijo, 2013) y la idoneidad de las fuentes de datos utilizadas (Thelwall *et al.*, 2013). Hasta el momento los estudios han reportado 1) una correlación relativamente débil con las citas (i.e., Thelwall *et al.*, 2013; Costas; Zahedi; Wouters, 2014), 2) que su potencial para ofrecer aspectos complementarios del impacto sigue siendo desconocido; y 3) que tweets, menciones en blogs, lectores en *Mendeley*, recomendaciones en *F1000* y apariciones en los medios parecen ser las fuentes más relevantes (Li; Thelwall, 2012; Li; Thelwall; Giustini,

2012; **Haustein et al.**, 2013; **Costas; Zahedi; Wouters**, 2014; **Zahedi; Costas; Wouters**, en prensa). En los últimos años han aparecido muchos servicios que recogen y proporcionan esas nuevas métricas alternativas. Las principales son *ImpactStory.org*¹, *Plum Analytics*² y *Altmetric.com*³.

Altmetric.com es uno de los proveedores de altmétricas más importantes. Captura información sobre el impacto de un documento de varios medios sociales y elabora una puntuación ponderada. Para ello elimina ambigüedades en los enlaces a los artículos⁴ unificando los enlaces a los registros de *PubMed*, a identificadores del repositorio *arXiv*, números doi y sitios web de los editores. Aunque algunos han advertido contra el uso de puntuaciones altmétricas agregadas (**Davis**, 2013), no ha habido mucho debate sobre la riqueza y la diversidad de los datos. Uno de los principales problemas a los que nos enfrentamos es la diversidad y riqueza de datos, que es realmente difícil de entender. Aunque *Altmetric.com* ofrece amplia información sobre sus contenidos aún hay dificultades para evaluar la exhaustividad de los datos y las posibilidades que esta fuente puede proporcionar. <http://support.altmetric.com>

El objetivo de este trabajo es explorar *Altmetric.com* como fuente de altmétricas. Con el fin de conocer el uso potencial de esta web, en este artículo ofrecemos una visión global y práctica sobre sus contenidos. En concreto, vamos a responder a las siguientes preguntas de investigación:

1. ¿Qué fuentes de datos están incluidas en *Altmetric.com* y cómo están éstos estructurados?
2. ¿Cuál es la cobertura de *Altmetric.com* y qué fuentes de datos cubren mejor el impacto altmétrico de las publicaciones?

Para ello realizamos una extracción de datos de *Altmetric.com* y luego los analizamos pormenorizadamente.

Material y métodos

Seleccionamos todas las publicaciones entre 2011 y 2013⁵ indexadas en la base de datos *Web of Science (WoS)* utilizando la versión de esta bdd existente en el *Centre for Science and Technology Studies (CWTS)* de la *Universidad de Leiden*. De este conjunto de documentos se seleccionaron los 2.792.706 que incluían número de identificación doi. Luego, en enero de 2014 comprobamos los números doi con la API de *Altmetric*.

<https://api.altmetric.com>

Con dicha API recuperamos 516.150 registros. Esto significa que aproximadamente el 19% de todos los artículos con doi durante el período de tiempo del estudio había recibido algún tipo de atención de los medios sociales (figura 2). Sin embargo debemos señalar que hay errores en algunos de los dois presentes en *Altmetric.com*. Además, no todos los artículos de *Altmetric.com* incluyen doi. Para cada registro se obtuvo un archivo en formato *javascript object notation format (JSON)*⁶. Los ficheros JSON incluyen datos en bruto recopilados por *Altmetric.com* para cada artículo. La tabla 1 muestra la estructura de cada fichero indicando el tipo de información suministrada en cada sección.

Como se observa, se identificaron cinco secciones. La primera es un resumen de las puntuaciones globales recuperadas de cada fuente. En la segunda se da una breve descripción del artículo científico incluyendo no sólo la referencia bibliográfica, sino también la fecha en que el documento fue incluido por primera vez en el sistema y enlaces alternativos al mismo. La tercera sección ofrece una evolución temporal de la puntuación agregada de *Altmetric* para varios períodos de tiempo, junto con las comparaciones con las puntuaciones de la revista. En la cuarta se muestra información demográfica, por país y tipo de público, obtenida de la cuenta de *Twitter* de los usuarios que mencionan el artículo. Final-

Tabla 1. Estructura de un registro proporcionado por la API de *Altmetric.com*

Descripción	Ejemplo de campos extraídos
Resumen de métricas según se muestra en el bookmarklet ⁷ de <i>Altmetric.com</i>	"counts":{"readers":{"mendeley":"citeulike","connotea"},"facebook":{"unique_users_count","unique_users":[]},"posts_count"},"blogs":{"unique_users_count","unique_users":[]},"news":{"unique_users_count","unique_users":[]},"pinterest":{"unique_users_count","unique_users":[]},"posts_count"},"reddit":{"unique_users_count","unique_users":[]},"posts_count"},"twitter":{"unique_users_count","unique_users":[]},"posts_count"},"video":{"unique_users_count","unique_users":[]},"posts_count"},"linkedin":{"unique_users_count","unique_users":[]},"posts_count","total":[]}...
Descripción bibliográfica del documento	"citation":{"title","authors":[],"pubdate","volume","issue","startpage","endpage","doi","PMID","arxiv_id","journal","altmetric_jid","links":[],"first_seen_on"}
Comparación y evolución de la puntuación agregada <i>Altmetric</i>	"altmetric_score":{"score","score_history":{"1d","2d","3d","4d","5d","6d","1w","1m","3m","6m","1y","at"},"context_for_score":{"all":{"rank","mean","median","sample_size","sparkline","total_number_of_other_articles","this_scored_higher_than","this_scored_higher_than_pct","percentile","rank_type":"approximate"},"similar_age_3m":{"rank","mean","median","sample_size","sparkline","total_number_of_other_articles","this_scored_higher_than","this_scored_higher_than_pct","percentile","rank_type":"approximate"}},...
Demografía (<i>Twitter</i>): tipo de público y país	"demographics":{"poster_types":{"member_of_the_public","researcher","practitioner","science_communicator"},"geo":{"twitter":{"Country":"*number of users*}}}
Datos de <i>Altmetric</i> desglosados por proveedor	"posts":{"twitter":{"url","posted_on","license","summary","author":{"name","image","id_on_source","followers"},"tweet_id"},"blogs":{"title":{"title","url","posted_on","summary","author":{"name","url","description"}}},"facebook":{"title","url","posted_on","summary","author":{"name","url","facebook_wall_name","image","id_on_source"},"url","posted_on","summary","author":{"name","url","facebook_wall_name","image","id_on_source"},"googleplus":{"title","url","posted_on","summary","author":{"name","url","image","id_on_source"}},...

mente, la última sección incluye información y los campos registrados en el sistema obtenidos de las fuentes de las que *Altmetric.com* recupera los datos.

Descripción de las fuentes recolectadas por *Altmetric.com*

Se identificaron 16 fuentes, que presentamos en la tabla 2 con una breve descripción, el tipo de métrica que miden y los elementos o campos de los datos recuperados. Cada registro mantiene un histórico de todas las métricas registradas desde 2011 o desde la inclusión del artículo en el

sistema. Para capturar esos datos, *Altmetric.com* identifica menciones reconociendo los enlaces a los artículos. La única excepción se hace con los blogs y noticias, donde también se emplea un mecanismo de seguimiento con técnicas de minería de textos para capturar las menciones que no tienen url de enlace a la publicación. Estas técnicas se emplean sólo para fuentes en idioma inglés.

Como se observa, el tipo más común de las métricas recopiladas son conversaciones y menciones (con 4 y 5 fuentes respectivamente), seguidas por el recuento de los lectores (*Mendeley*, *Connotea* y *CiteULike*). Además se pueden ver

Tabla 2. Resumen de los elementos de datos proporcionados por las fuentes de *Altmetric.com*

Fuente	Descripción	Tipo de métrica	Elementos de datos
Blogs	Lista RSS elaborada manualmente	Discusión	Título del blog; título del post; url del post; fecha y hora de publicación; resumen; nombre del autor; url del autor; descripción del autor
Noticias	Lista RSS elaborada manualmente	Discusión	Título de la noticia; url de la noticia; fecha y hora de publicación; licencia; resumen; nombre del medio de comunicación; url del medio; identificador del medio; imagen del medio
<i>Reddit</i>	Proveedor de noticias	Discusión	Título de la noticia; url de reddit; fecha y hora de publicación; nombre del autor; url del autor; identificador del autor; seguidores; subreddit
<i>Facebook</i>	Red social	Menciones	Título de la mención; url de la mención; fecha y hora de publicación; resumen; nombre del autor; url del autor; nombre del muro de <i>Facebook</i> ; imagen del autor; identificador del autor
<i>Google Plus</i>	Red social	Menciones	Título de la mención; url de la mención; fecha y hora de publicación; resumen; nombre del autor; url del autor; imagen del autor; identificador del autor
<i>Pinterest</i>	Red social	Menciones	Url de la mención; imagen de la mención; fecha y hora de publicación; resumen; nombre del autor; <i>pinboard</i>
<i>Twitter</i>	Micro-blogging	Menciones	Url; fecha y hora de publicación; licencia; resumen; nombre del autor; imagen del autor; número de seguidores; identificador del tweet; tipo de público; país
<i>Stack Exchange</i>	Preguntas y respuestas	Discusión	Título de la discusión; url de la discusión; fecha y hora de publicación; resumen; identificador del autor
<i>CiteULike</i>	Marcadores sociales	Lectores	Recuento total de marcadores
<i>Connotea</i>	Marcadores sociales (servicio cerrado)	Lectores	Recuento total de marcadores
<i>Mendeley</i>	Marcadores sociales	Lectores	Recuento total de marcadores
<i>F1000</i>	Servicio de peer review posterior a la publicación	Revisiones	Recomendado en <i>F1000</i> ; fecha de publicación (probablemente de la última actualización); tipo de recomendación
<i>YouTube</i>	Compartición de vídeos	Vídeo	Título del vídeo; url del vídeo; imagen de vídeo; fecha y hora de publicación; licencia; resumen; tipo de inserción; identificación de <i>YouTube</i> ; nombre del autor; identificador del autor
<i>LinkedIn Groups</i>	Red social profesional	Menciones	Total de usuarios únicos; nombre de los usuarios únicos; total de posts; título del post; resumen; fecha y hora de publicación; nombre del autor; descripción del autor; url del post; url del logo del grupo; nombre del grupo; descripción del grupo
<i>Research Highlights</i>	Aspectos destacados por <i>Nature</i>	Citas	url del destacado; fecha añadido a <i>Altmetric.com</i> ; título del destacado; total de destacados; descripción bibliográfica del destacado; fecha visto por primera vez
Misc	Otros	Otros	Datos procedentes de otros medios sociales que se añaden a petición de los autores (Adie , 2014)

otras métricas similares como los vídeos, reseñas o hilos de discusión tipo ‘preguntas y respuestas’. Con la excepción de *Research highlights*, que incluye datos de las citas recuperadas de la sección *Highlights* (aspectos destacados) de la revista *Nature*, todas las fuentes son 2.0. Algunas de estas fuentes están sesgadas hacia ciertos campos: por ejemplo, *F1000* es un servicio de revisión inter pares posterior a la publicación de la investigación biomédica (Waltman; Costas, 2014); *Stack Exchange* lo utilizan sobre todo los investigadores en informática y ciencias naturales.

Con la excepción del campo Misc, que se dedica a otros medios de comunicación no incluidos en el conjunto original de *Altmetric.com*, todos están incluidos en el cálculo de la puntuación agregada de cada artículo. La mayor parte de esta información se puede visualizar a través del bookmarklet *Altmetric.com* (figura 1). Sin embargo, se han observado algunas diferencias entre los registros recuperados de la API *Altmetric* y los mostrados en el bookmarklet *Altmetric*: algunos indicadores y elementos de datos no aparecen en el bookmarklet (por ejemplo, todos los tweets y retweets) y también discrepancias en la información proporcionada de las fuentes (por ejemplo, errores ocasionales en las conversaciones de preguntas y respuestas).

Cobertura de *Altmetric.com* de publicaciones *WoS* con doi en 2011-2013⁵

Como se dijo, del total de publicaciones de la muestra original obtenida de la *WoS*, sólo el 19% estaba incluido en *Altmetric.com* con algún tipo de impacto altmétrico (figura 2). *Twitter* es la fuente que proporciona más datos altmétricos (87,1%) seguido de *Mendeley* (64,8%). Ninguno de los otros medios de comunicación social alcanza valores superiores al 20% de la cuota total de artículos que disponen de indicadores altmétricos, aunque *Facebook* se acerca con un 19,9% de los documentos incluidos en *Altmetric.com*.

En la tabla 3 se incluye más información sobre el número de artículos que disponen de métricas, recuento total de cada métrica y usuarios únicos para las cinco fuentes principales (*Twitter*, *Mendeley*, *Facebook*, *CiteULike* y blogs). Estas fuentes están presentes en el 95,5% del

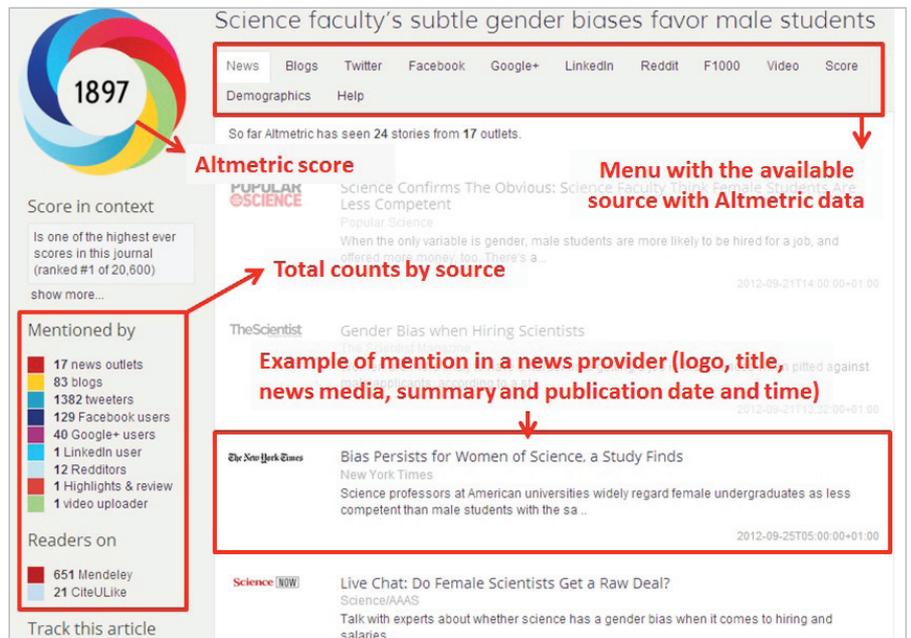


Figura 1. Ejemplo de los datos proporcionados por el bookmarklet *Altmetric.com*

total de documentos recuperados de *Altmetric.com*. Aunque *Twitter* es el medio social con más menciones, *Mendeley* incluye un mayor número de usuarios que señalan trabajos científicos. Estas dos fuentes de datos son las más extendidas entre todas las fuentes altmétricas analizadas. De hecho, la presencia de menciones de artículos científicos de los medios sociales como *Facebook*, *CiteULike* o incluso blogs, nunca llega a 5% de los trabajos totales con doi indexados en la *Web of Science* en el período de tiempo estudiado.

Discusión y observaciones finales

En este trabajo se ha analizado *Altmetric.com* como proveedor de datos altmétricos para medir el impacto de las publicaciones científicas. El principal problema de este tipo de servicios es la dificultad en identificar las menciones a los trabajos científicos, de manera similar a las dificultades observadas cuando se utilizan técnicas webmétricas (Thelwall, 2011). Aunque *Altmetric.com* afirma hacer serios esfuerzos para desambiguar los enlaces, aún hay que investigar mu-

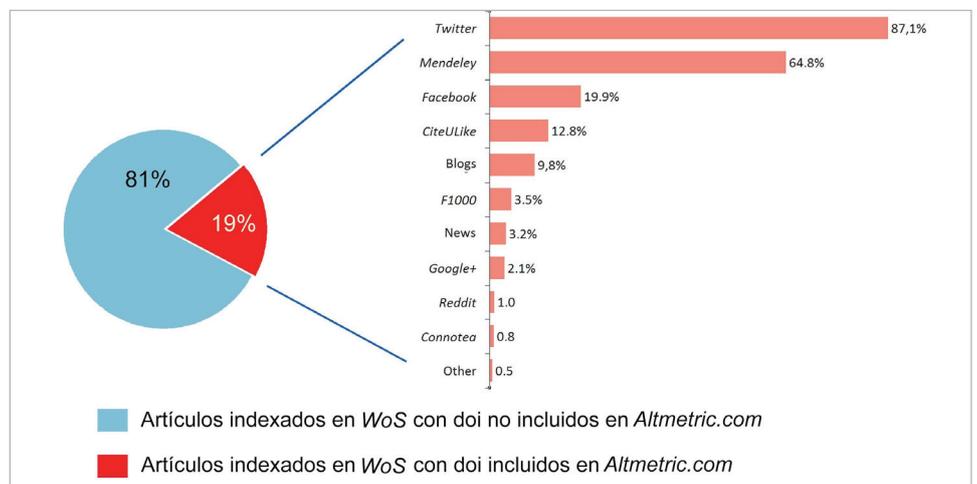


Figura 2. Cobertura por los medios de comunicación social de los artículos indexados en *WoS* según *Altmetric.com* en el período 2011-2013⁵.

Tabla 3. Cobertura de *Altmetric.com* de redes sociales que mencionan artículos indexados en la *Web of Science* en el período 2011-2013⁵

Medio social	Artículos	Recuentos	Usuarios únicos	% de artículos en WoS
Twitter	449,493	1,819,194	1,621,396	16.1
Mendeley	334,616	2,631,396	2,631,396	12.0
Facebook	102,923	197,449	182,422	3.7
CiteULike	65,799	130,756	130,756	2.4
Blogs	50,529	84,927	75,946	1.8

cho sobre la exhaustividad, precisión y exactitud de la información obtenida por esos sistemas (por ejemplo, ¿cuántas menciones de las fuentes cubiertas faltan en *Altmetric.com*?). Esto es especialmente relevante cuando se analiza el método de identificar menciones a artículos científicos en las fuentes más problemáticas como son los blogs o los medios de comunicación.

<http://support.altmetric.com>

En este último caso aplican un mecanismo de seguimiento basado en minería de textos como un complemento al reconocimiento del enlace. Sin embargo, se aplica a una lista de recursos seleccionados manualmente, no siendo evidentes los criterios seguidos.

<http://www.altmetric.com/sources-blogs.php>

Además, esta técnica se aplica sólo a las fuentes en idioma inglés; para las fuentes que no están en inglés sólo se consideran los enlaces directos a las publicaciones, lo que introduce un sesgo de idioma importante que debe tenerse en cuenta.

<http://www.altmetric.com/sources-news.php>

Una limitación muy seria se debe a las fuentes usadas por *Altmetric.com*. Sería importante conocer las razones de por qué esas y no otras, particularmente en un entorno en el que cada vez hay más medios sociales. Esta deficiencia se aplica también a los demás proveedores de altmétricas, ya que no siempre justifican empírica o conceptualmente las fuentes seleccionadas. Se podría argumentar que si está cubierto *Twitter*, por qué no *Tumblr*; o por qué no está el servicio español *Menéame* junto a *Reddit*? En la misma línea, y en relación con la investigación científica, es importante mencionar la omisión de las redes sociales científicas como *Academia.edu* o *ResearchGate* que parecen ser utilizadas por muchos investigadores (Mas-Bleda; Thelwall; Kousha; Aguillo, 2014). En este sentido se han producido algunas mejoras: el 7 de abril de 2014 *Altmetric.com* informó de la inclusión de *Weibo*, un servicio chino híbrido entre *Twitter* y *Facebook*, como nueva fuente (Adie, 2014).

Probablemente, la razón de la selección de las fuentes actuales se debió más a una cuestión práctica que conceptual (estas fuentes son muy populares, tienen APIs públicas, son internacionales, etc.) y, aunque con limitaciones, encontrar y escanear menciones a artículos de investigación a través de ellas es relativamente fácil. Sin embargo, los problemas técnicos no deben evitar una discusión más conceptual y teórica sobre lo que debe estar cubierto y las posibles limi-

taciones o sesgos de las fuentes actuales, de forma similar a los análisis sobre la cobertura y las limitaciones de otras bases de datos bibliométricas como la *Web of Science*, *Scopus* o *Google Scholar* (ver por ejemplo Jacsó, 2009).

Nuestros resultados muestran que de las 16 fuentes cubiertas por *Altmetric.com* sólo 5 representan el 95,5% del total de artículos que tienen altmétricas. Esto abre la cuestión de la pertinencia de las fuentes y si los más pequeños realmente puede proporcionar resultados significativos de impacto. La concentración en un pequeño número de medios de comunicación social ya se ha estudiado (Priem *et al.*, 2012; Cabezas-Clavijo; Torres-Salinas, 2010). Las fuentes más importantes son *Twitter* y *Mendeley* (figura 2) y son los que parecen ser más prometedoras para determinar los impactos altmétricos, ya que tienen una mayor densidad de menciones y las métricas extraídas de ellos son más fiables. Como se observa en nuestros resultados, mientras que *Twitter* ofrece datos de un mayor número de publicaciones, *Mendeley* muestra las cifras más altas (tabla 2), incluyendo un mayor número de recuentos y los usuarios. *Mendeley* parece haberse extendido mucho entre la comunidad científica (Haustein *et al.*, 2014). Sorprendentemente, *Altmetric.com* no recopila datos de lectores de *Mendeley*, a menos que se recojan otros indicadores bibliométricos (Costas; Zahedi; Wouters, 2014).

“*Altmetric.com* es un proveedor abierto y transparente, que ofrece datos de calidad sobre menciones a publicaciones científicas en medios sociales”

Como resumen podemos decir que *Altmetric.com* es un proveedor abierto y transparente, que ofrece datos de calidad sobre menciones a publicaciones científicas en medios sociales. La reciente alianza establecida entre *ImpactStory* (otro importante proveedor altmétrico) y *Altmetric.com* (Piwowar, 2014) es un claro reconocimiento del valor de esta herramienta. Nuestro estudio pone de manifiesto la riqueza de los datos recogidos, que no se limitan a los recuentos y menciones, sino también a información sobre los usuarios, su origen o de las fechas de la mención, por ejemplo. Tiene dos características positivas importantes: 1) los datos se almacenan y se registran de forma permanente, lo que permite la reproducibilidad de los resultados y el análisis retrospectivo, dando así una solución al problema de la volatilidad de los datos altmétricos (Wouters; Costas, 2012); 2) la riqueza de los datos registrados abre las posibilidades a realizar análisis más allá del simple recuento de las menciones. Por ejemplo, es posible analizar los tipos de público, los intereses de las audiencias, sus relaciones, etc., nuevas posibilidades aún no exploradas.

Por último, nuestro estudio muestra que aún hay temas importantes que necesitan resolverse para comprender completamente los datos altmétricos. Se necesita más investigación para hallar métodos de recuperación de datos altmétricos válidos y fiables. En la misma línea, la selección de fuentes, de medios sociales, debe ser rigurosa y crítica,

atendiendo a su uso dentro de las diferentes comunidades y audiencias y evitar posibles sesgos de disciplina o de idioma.

Agradecimientos

Los autores desean agradecer a Erik van Wijk, del *CWTS*, su ayuda en la recuperación de los datos. Euan Adie, de *Altmetric.com*, aclaró algunas de nuestras dudas sobre los datos. Stefanie Haustein contribuyó con sus comentarios a mejorar la versión final del manuscrito.

Nicolás Robinson-García tiene una beca FPU del *Ministerio de Economía y Competitividad* de España.

Notas

1. <http://impactstory.org>

Fundado por **Jason Priem** y **Heather Piwowar** en 2011, originalmente fue llamado *Total-Impact*.

2. <http://www.plumanalytics.com>

Fundado a finales de 2011 por **Andrea Michalek** y **Mike Buschman**, recientemente ha sido adquirido por *Ebsco Publishing*.

3. <http://www.altmetric.com>

Fundado por **Euan Adie** en 2011, se ha convertido en uno de los principales proveedores de datos altmétricos.

4. Para simplificar hablamos de “artículos”, pero en *Altmetric.com* también se incluyen ponencias de congresos, o sea, se trata de *papers* en general.

5. El año de publicación 2013 no está completo: en ese momento sólo estaba cargado un tercio de las publicaciones. En cualquier caso esto no es un problema para nuestro trabajo, ya que sólo estamos haciendo un análisis descriptivo.

6. Para obtener más información sobre el formato JSON se remite al lector a <http://en.wikipedia.org/wiki/JSON>

7. Un bookmarklet es un marcador almacenado en un navegador web que contiene órdenes javascript para extender la funcionalidad del navegador. Por ejemplo, un bookmarklet podría permitir al usuario seleccionar un término en una página, hacer clic en el bookmarklet, y presentársele una página de resultados de un buscador con el término de búsqueda.

Bibliografía

Adie, Euan (2014). “Announcing Sina Weibo support”.

<http://www.altmetric.com/blog/announcing-sina-weibo-support>

Adie, Euan (2014). *Personal communication*.

Almind, Tomas C.; Ingwersen, Peter (1997). “Informetric analyses on the world wide web: Methodological approaches to ‘webometrics’”. *Journal of documentation*, v. 53, n. 4, pp. 404-426.

<http://dx.doi.org/10.1108/EUM0000000007205>

Altmetric.com. *Knowledge base*.

<http://support.altmetric.com/knowledgebase>

Bollen, Johan; Van-de-Sompel, Herbert (2006). “Mapping the structure of science through usage”. *Scientometrics*, v. 69, n. 2, pp. 227-258.

<http://dx.doi.org/10.1007/s11192-006-0151-8>

Cabezas-Clavijo, Álvaro; Torres-Salinas, Daniel (2010). “Los investigadores en la ciencia 2.0: El caso de *PLoS one*”. *El profesional de la información*, v. 19, n. 4, pp. 431-434.

<http://dx.doi.org/10.3145/epi.2010.jul.14>

Costas, Rodrigo; Zahedi, Zohreh; Wouters, Paul (2014). “Do ‘altmetrics’ correlate with citations? Extensive comparison of altmetric indicators with citations from a multidisciplinary perspective”.

<http://arxiv.org/abs/1401.4321>

Cronin, Blaise; Weaver-Wozniak, Sherrill (1995). “The praxis of acknowledgement: From bibliometrics to influmetrics”. *Revista española de documentación científica*, v. 18, n. 2, pp. 172-177.

<http://dx.doi.org/10.3989/redc.1995.v18.i2.654>

Davis, Phil (2013). “Visualizing article performance - Altmetric searches for appropriate display”.

The scholarly kitchen. <http://scholarlykitchen.sspnet.org/2013/09/30/visualizing-article-performance-altmetrics-searches-for-appropriate-display>

Haustein, Stefanie; Peters, Isabella; Bar-Ilan, Judit; Priem, Jason; Shema, Hadas; Terliesner, Jens (2014). “Coverage and adoption of altmetrics in the bibliometric community”. *Scientometrics*.

<http://dx.doi.org/10.1007/s11192-013-1221-3>

Haustein, Stefanie; Peters, Isabella; Sugimoto, Cassidy R.; Thelwall, Mike; Larivière, Vincent (2013). “Tweeting biomedicine: An analysis of tweets and citations in the biomedical literature”.

Journal of the American Society for Information Science and Technology, v. 65, n. 4, pp. 656-669.

<http://dx.doi.org/10.1002/asi.23101>

Jacsó, Péter (2009). “Testing the calculation of a realistic h-index in Google Scholar, Scopus, and Web of Science for F.W. Lancaster”.

Library trends, v. 56, n. 4, pp. 784-815.

<http://dx.doi.org/10.1353/lib.0.0011>

Kurtz, Michael J.; Bollen, Johan (2010). “Usage bibliometrics”. *Annual review of information science and technology*, v. 44, pp. 1-64.

<http://dx.doi.org/10.1002/aris.2010.1440440108>

Li, Xuemei; Thelwall, Mike (2012). “F1000, Mendeley and traditional bibliometric indicators”. In *17th Intl conf on science and technology indicators*, v. 3, pp. 1-11.

http://2012.sticonference.org/Proceedings/vol2/Li_F1000_541.pdf

Li, Xuemei; Thelwall, Mike; Giustini, Dean (2012). “Validating online reference managers for scholarly impact measurement”. *Scientometrics*, v. 91, n. 2, pp. 461-471.

<http://dx.doi.org/10.1007/s11192-011-0580-x>

Mas-Bleda, Amalia; Thelwall, Mike; Kousha, Kayvan; Aguillo, Isidro F. (2014). “Successful researchers publicizing research online: An outlink analysis of European highly cited scientists’ personal websites”. *Journal of documentation*, v. 70, n. 1, pp. 148-172.

<http://dx.doi.org/10.1108/JD-12-2012-0156>

Piwowar, Heather A. (2013). “Altmetrics: Value all research products”. *Nature*, v. 493, n. 159, January 10.

<http://dx.doi.org/10.1038/493159a>

Piwowar, Heather A. (2014). "Impactstory partners with Altmetric.com". *ImpactStory blog*.

http://blog.impactstory.org/2014/01/28/altmetric_com

Priem, Jason; Piwowar, Heather A.; Hemminger, Bradley M. (2012). "Altmetrics in the wild: Using social media to explore scholarly impact".

<http://arxiv.org/html/1203.4745>

Priem, Jason; Taraborelli, Dario; Groth, Paul; Neylon, Cameron (2010). "Altmetrics: A manifesto-altmetrics.org".

<http://altmetrics.org/manifesto>

Thelwall, Mike (2011). "A comparison of link and URL citation counting". *Aslib proceedings*, v. 63, n. 4, pp. 419-425.

<http://dx.doi.org/10.1108/00012531111148985>

Thelwall, Mike; Haustein, Stefanie; Larivière, Vincent; Sugimoto, Cassidy R. (2013). "Do altmetrics work? Twitter and ten other social web services". *PLoS one*, v. 8, n. 5, e64841.

<http://dx.doi.org/10.1371/journal.pone.0064841>

Torres-Salinas, Daniel; Cabezas-Clavijo, Álvaro (2013). "Altmetrics: no todo lo que se puede contar, cuenta". *Anuario*

ThinkEPI, v. 7, pp. 114-117.

Torres-Salinas, Daniel; Cabezas-Clavijo, Álvaro; Jiménez-Contreras, Evaristo (2013). "Altmetrics: New indicators for scientific communication in web 2.0". *Comunicar*, v. 21, n. 41, pp. 53-60.

<http://dx.doi.org/10.3916/C41-2013-05>

Waltman, Ludo; Costas, Rodrigo (2014). "F1000 Recommendations as a potential new data source for research evaluation: A comparison with citations". *Journal of the Association for Information Science and Technology*, v. 65, n. 3, pp. 433-445.

Wouters, Paul; Costas, Rodrigo (2012). "Users, narcissism and control - Tracking the impact of scholarly publications in the 21st Century". In: *Procs of 17th Intl conf on science and technology indicators*, v. 2, pp. 847-857.

http://2012.sticonference.org/Proceedings/vol2/Wouters_Users_847.pdf

Zahedi, Zohreh; Costas, Rodrigo; Wouters, Paul (en prensa). "How well developed are altmetrics? Cross disciplinary analysis of the presence of 'alternative metrics' in scientific publications". *Scientometrics*.

<http://dx.doi.org/10.1007/s11192-014-1264-0>



5^a Conferencia internacional sobre calidad de revistas de ciencias sociales y humanidades

CRECS 2015
7-8 de mayo

Universidad de Murcia. Facultad de Comunicación y Documentación

<http://www.thinkepi.net/crecs2015>



UNIVERSIDAD DE MURCIA

CENTUM
CIEN AÑOS DE LA UNIVERSIDAD DE MURCIA
1915 | 2015

El profesional de la información

ANUARIO
Think
EPI