



TENDENCIAS EN MAPAS DE LA CIENCIA: CO-USO DE INFORMACIÓN CIENTÍFICA COMO REFLEJO DE LOS INTERESES DE LOS INVESTIGADORES



Daniel Torres-Salinas, Evaristo Jiménez-Contreras y Nicolás Robinson-García



Daniel Torres-Salinas es doctor en documentación científica por la *Universidad de Granada*. Trabaja como técnico de gestión en la *Universidad de Navarra (UN)* donde realiza auditorías sobre el rendimiento científico y es investigador en el ámbito de la bibliometría y miembro del *Grupo EC3 (Evaluación de la Ciencia y la Comunicación Científica de la Universidad de Granada)*. Es co-creador de diferentes herramientas de evaluación de la investigación como *Cientificacvn*, los *Rankings I-UGR de Universidades* o la *Clasificación integrada de revistas científicas (CIRC)*. Es cofundador y CEO de la spin-off de asesoría científica *EC3metrics*. Habitualmente imparte cursos sobre comunicación científica, web 2.0 y ciencia o servicios de investigación en bibliotecas.

<http://orcid.org/0000-0001-8790-3314>

torressalinas@gmail.com



Evaristo Jiménez-Contreras es catedrático de bibliometría en la *Facultad de Comunicación y Documentación* de la *Universidad de Granada*, y director del grupo de investigación *EC3 (Evaluación de la Ciencia y de la Comunicación Científica)*. Sus líneas de investigación se centran en la evaluación de la actividad científica con metodología bibliométrica. Es uno de los promotores de los índices de impacto de las revistas españolas de ciencias sociales, jurídicas y humanas (respectivamente *In-Recs, In-Recj, In-Rech*), *CIRC (Clasificación integrada de revistas científicas)*, *Metaranking EC3*, y *Científica*.

<http://orcid.org/0000-0001-5668-7057>

evaristo@ugr.es



Nicolás Robinson-García es licenciado en documentación y máster en información científica por la *Universidad de Granada*. Es miembro del grupo *EC3 (Evaluación de la Ciencia y la Comunicación Científica)* de la misma universidad, donde se encuentra haciendo el doctorado. En la actualidad sus intereses se centran en la investigación sobre bibliometría y la evaluación de la calidad investigadora.

<http://orcid.org/0000-0002-0585-7359>

elrobin@ugr.es

*Universidad de Granada, EC3 Research Group
Campus Cartuja. 18071 Granada, España*

Resumen

Se explora la posibilidad de elaborar mapas de la ciencia basados en el co-uso de la información científica que hacen los usuarios de una institución académica. Se define co-uso como la co-ocurrencia de consultas de información científica por parte de pares de usuarios en una plataforma de revistas científicas. Sobre la base de los datos de consulta de la *Universidad de Navarra* en 2012 a la plataforma de revistas científicas *ScienceDirect* se analiza el potencial de dicha aproximación metodológica. Se destaca la viabilidad de dicha metodología para ahondar en los intereses investigadores de una institución académica y las relaciones que se producen entre distintas disciplinas.

Palabras clave

Mapas de la ciencia, Visualización de la información, Co-uso, Consumo de información científica, Bibliografía científica, Metodologías, Recuperación de información, Intereses de investigación, *Universidad de Navarra*.

Title: Trends in science mapping: Co-use of scientific literature as evidence of researchers' interests

Abstract

The possibility of constructing science maps based on co-use of scientific literature by academic users is explored. We define co-use as the co-occurrence of scientific information requests by pairs of users recorded in platforms of scientific journals. We used search data from the *University of Navarra* to the *ScienceDirect* platform in 2012 in order to test the validity and to analyze the possibilities of this methodological approach. We conclude by emphasizing the viability of this methodology when exploring the research interests of an academic institution along with the relations among different disciplines.

Artículo recibido el 30-01-2014

Aceptación definitiva: 21-05-2014

Keywords

Science maps, Information visualization, Co-use, Consumption of scientific literature, Methodologies, Information retrieval, Research interests, *University of Navarre*.

Torres-Salinas, Daniel; Jiménez-Contreras, Evaristo; Robinson-García, Nicolás (2014). "Tendencias en mapas de la ciencia: co-uso de información científica como reflejo de los intereses de los investigadores". *El profesional de la información*, mayo-junio, v. 23, n. 3, pp. 253-258.

<http://dx.doi.org/10.3145/epi.2014.may.05>

Introducción

El formato electrónico ha transformado la forma de consumo de la información científica. En 2005 la quinta parte de los artículos que consultaban los investigadores eran en formato electrónico; actualmente son más de la mitad (**Tenopir et al.**, en prensa). Además los hábitos de lectura han cambiado: los investigadores leen un mayor número de trabajos, la lectura se suele hacer de manera transversal dedicando menos tiempo por trabajo y se tiende a prestar una mayor atención a los trabajos más breves (**Van-Noorden**, 2014).

Este escenario plantea nuevas oportunidades y retos para bibliotecarios y bibliómetras, ya que resulta más sencillo monitorizar la actividad de los usuarios y analizar el consumo que hacen de la información científica en el medio digital. Se comienza a profundizar en el análisis de nuevos indicadores que puedan complementar o sustituir a los tradicionales (préstamos o lecturas en sala en el caso de los bibliotecarios, y citas en el de los bibliómetras). Surgen ambiciosos proyectos como *Counter*, dirigido a bibliotecarios y editores o *Mesur*, dirigido a la comunidad científica.

<http://www.projectcounter.org>

<http://mesur.informatics.indiana.edu>

De manera paralela los sectores profesional y académico emprenden la búsqueda de nuevos indicadores (**Borrego; Urbano**, 2005; **Torres-Salinas; Cabezas-Clavijo**, 2013), como: marcadores sociales (**Haustein; Siebenlist**, 2011), consultas en bibliotecas digitales (**Bollen et al.**, 2005), hiperenlaces (**Aguillo et al.**, 2006) o menciones en redes sociales (**Torres-Salinas; Cabezas-Clavijo; Jiménez-Contreras**, 2013). Se observan principalmente tres frentes de investigación que analizan el uso de la literatura científica utilizando indicadores alternativos a los tradicionales:

1) La webmetría (**Almind; Ingwersen**, 1997): considera internet como el mayor repositorio mundial de documentos y analiza relaciones entre ellos mediante hiperenlaces del mismo modo que lo hacen las citas (**Theilwall**, 2008). Así resulta posible analizar por ejemplo la visibilidad de las instituciones académicas en la Web (**Aguillo et al.**, 2006).

2) La bibliometría basada en el uso o *usage bibliometrics* (**Kurz; Bollen**, 2010): nació a mediados de los noventa y se deriva del análisis de las colecciones electrónicas para determinar las políticas de suscripción de las bibliotecas a paquetes de revistas y asesorar a los editores en la mejora de sus servicios (**Nicholas; Huntington; Watkinson**, 2005; **Rodríguez-Bravo et al.**, 2012). Posteriormente, **Bollen** planteó su uso como herramienta de evaluación científica (**Bollen et al.**, 2005).

3) Las *altmetrics* o métricas alternativas (**Priem et al.**, 2010): la más reciente de estas propuestas. Propone utilizar indicadores de las redes sociales y demás servicios de la web 2.0 como medidas de la atención social que reciben los trabajos científicos.

En todos estos casos se plantean dos tipos de estudios: comparar la relación de los mismos con las citas o adaptar las técnicas bibliométricas a estos nuevos indicadores. En este proceso de extrapolación metodológica uno de los aspectos al que se ha prestado menos atención es la visualización y el mapeo de la ciencia. Se trata de técnicas que se han visto revitalizadas desde los años noventa con el avance de la tecnología y su implementación como herramientas de asesoramiento de la política científica (**Noyons**, 2004). Si bien los mapas de la ciencia están basados en indicadores relacionales como la citación (**Small**, 1973) o la coautoría (**Luukkonen; Persson; Sivertsen**, 1992), en los mapas de co-uso el elemento relacional se establecería cuando dos trabajos han sido consultados en una misma sesión o por un mismo usuario. Esto permite ofrecer una visión más amplia del consumo que se hace de la bibliografía científica al plantearla desde la perspectiva del lector y no de los autores (**Rowlands; Nicholas**, 2007).

“ El objetivo de este trabajo es explorar la posibilidad de realizar mapas de la ciencia de categorías temáticas basados en co-uso de información ”

Este trabajo explora cómo trasladar las técnicas de mapeo de co-citación al entorno de las plataformas online de revistas científicas con el objetivo de generar mapas de la ciencia de co-uso. Con esta perspectiva el objetivo esencial es explorar la posibilidad de realizar mapas de la ciencia de categorías temáticas basados en co-uso de información. Se explora asimismo que estos mapas ayuden a identificar los intereses científicos de una determinada comunidad, en concreto de una universidad.

Material y métodos

Como estudio de caso se ha seleccionado la *Universidad de Navarra* con el fin de elaborar un mapa de la ciencia basado en las consultas de artículos, principalmente descargas y visualización de ficheros html, que hizo su personal en 2012. Los datos de las consultas provienen de *ScienceDirect* y fueron cedidos a través del *Elsevier bibliometric research program*¹. La plataforma *ScienceDirect* contiene más de 2.500

revistas, que cubren las grandes áreas del conocimiento científico. Para cada consulta realizada por un usuario se obtuvo un registro que incluía la IP desde la que se realizó la consulta, el identificador de sesión, la fecha y diferentes datos identificados de la revista. Para asignar una categoría temática a cada revista los datos se cruzaron con el *Scimago journal rank* (SJR). En la tabla 1 se muestra la información básica del dataset.

A continuación se establecieron las relaciones entre categorías temáticas a partir del co-uso de los artículos. Las sesiones comienzan cada vez que un usuario accede a *ScienceDirect* y finalizan cuando éste abandona la plataforma o bien ha pasado cierto tiempo de inactividad. Cada sesión corresponde con una única demanda de información, ya que no puede haber más de un usuario en una misma sesión. Establecemos así un posible vínculo temático o de contenido entre los artículos consultados en un modo similar al que ocurre con la co-citación, donde se asume que dos trabajos citados conjuntamente están relacionados. El término co-uso hace mención por tanto a las relaciones que se establecen entre documentos consultados en una única sesión. Si dos trabajos de categorías temáticas diferentes han sido consultados, co-descargados o co-visualizados en la misma sesión, entre dichas categorías se genera el mencionado vínculo temático. Partiendo de esta idea se ha creado una matriz de distancias de co-uso entre categorías SJR normalizadas según el *índice de Jaccard* (Leydesdorff, 2008). Finalmente se representó la matriz a través de *Pajek*² manteniendo sólo las relaciones más significativas y denotando el número de consultas a través del tamaño de los nodos.

Índice de Jaccard

También conocido como coeficiente de similitud fue definido por Paul Jaccard (1868-1944), es una medida de la semejanza entre dos particiones o grupos de objetos (por ejemplo, resultantes de aplicar dos tipos de categorización al conjunto).

Sea un conjunto de documentos, de los cuales se forman dos grupos:

M: los que tienen el descriptor m

N: tienen el descriptor n

Definimos:

r = número de parejas de documentos que se encuentran ambas tanto en M como en N (o sea, están simultáneamente en M y en N)

s = número de parejas de documentos que se encuentran en M pero no en N

t = número de parejas de documentos que se encuentran en N pero no en M

Índice de Jaccard = $r / (r+s+t)$

Expresado de otra manera, es la intersección de 2 conjuntos $M \cap N$

N dividida por su reunión $M \cup N$

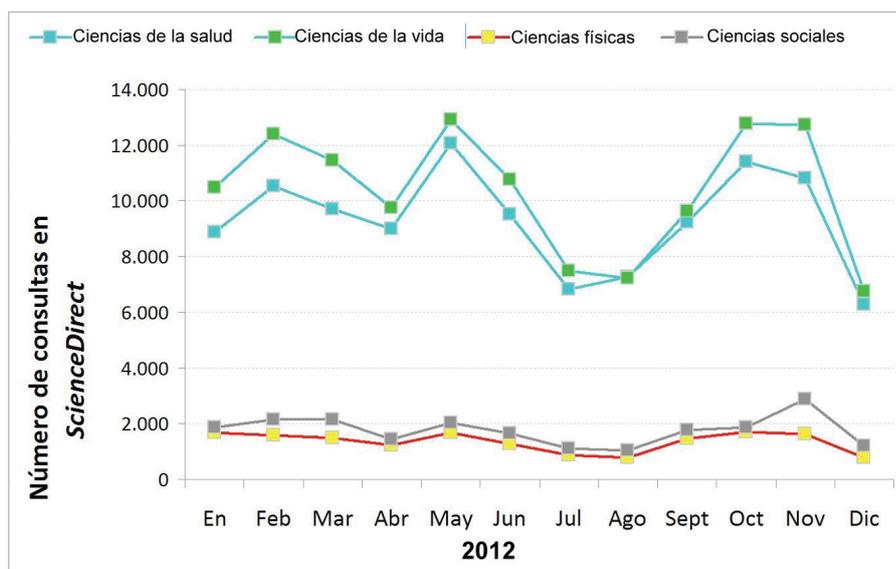


Figura 1. Distribución temporal por grandes áreas de las consultas realizadas por la Universidad de Navarra a ScienceDirect en 2012

Resultados: mapas de la ciencia de co-uso

El 75,98% de las consultas realizadas desde la Universidad de Navarra en 2012 fueron sobre trabajos pertenecientes a las áreas de ciencias de la vida y ciencias de la salud. Como se observa en la figura 1, los trabajos publicados en revistas pertenecientes al área de ciencias de la vida fueron los más consultados seguidos por ciencias de la salud. Por su parte ciencias sociales y humanidades representan poco más del 8% de las consultas recibidas mientras que ciencias físicas apenas superan el 6%.

Esta sobrerrepresentación de ciencias de la vida y de la salud se constata en el mapa de co-uso que se muestra en la figura 2. Vemos un mapa caracterizado por el predominio de las categorías de estas dos grandes áreas que quedan representadas en la parte superior. La parte superior izquierda del mapa está poblada por categorías pertenecientes a ciencias de la vida que conectan a través de farmacia y bioquímica con medicina (parte superior derecha). En el margen inferior derecho están las categorías relacionadas con economía, negocios y ciencias sociales unidas a medicina a través de enfermería. Ingenierías y ciencias de la computación conectan ciencias sociales con ciencias medioambientales (margen inferior izquierda) y ciencias físicas, que se sitúan en el centro del mapa conectando física con química y ésta con bioquímica, que cierra el círculo conectando nuevamente con ciencias de la salud.

Tabla 1. Descripción general de las consultas realizadas desde la Universidad de Navarra a ScienceDirect en 2012

Indicador	Resultado
Consultas	259.547
Sesiones	79.693
Números IP	1.845
Consultas por sesión	3,26
Sesiones por IP	43,21

rida por otros autores (Bollen *et al.*, 2005; Brody; Harnad; Carr, 2006), siendo este trabajo el primero que aplica dicho análisis a nivel institucional para representar los intereses de su comunidad de usuarios. Los mapas basados en el uso enriquecen mucho más el análisis al incluir a otros actores pasivos en el consumo de la literatura científica como son estudiantes, profesionales o docentes.

En líneas generales se observan patrones similares de producción y de consumo de información científica (figura 4), especialmente en el área de las ciencias de la salud. La plataforma de revistas electrónicas *ScienceDirect* es ampliamente utilizada tanto en ésta área como en las ciencias experimentales (Rodríguez-Bravo *et al.*, 2012), lo que hace que se adapte perfectamente a la universidad empleada como caso de estudio al tratarse de una universidad especializada precisamente en el ámbito biomédico. Por otra parte en este artículo empleamos una definición de uso de la información científica muy amplio, considerando como tal la consulta de bibliografía científica. En un futuro sería de interés analizar y comparar estos resultados con definiciones de uso más restrictivas como descargas, visualizaciones del texto completo, etc.

La metodología de visualización mediante co-uso puede fácilmente extrapolarse a otras plataformas y contextos de la web social académica

Del mismo modo, aunque los resultados preliminares son prometedores, todavía sigue siendo necesario ahondar más en el significado de este tipo de mapas y su posible utilidad como complemento a los tradicionales mapas de la ciencia. Su carácter más inclusivo, al contemplar otras dimensiones distintas de la investigación propiamente, consulta profesional o docentes, apunta a tres posibles usuarios:

- El equipo de dirección de bibliotecas: los mapas permiten visualizar de forma rápida el consumo de *ScienceDirect* y el uso real que hacen de ella sus usuarios así como la interconexión que existe entre los intereses de su comunidad de usuarios.
- El equipo de gestión científica: la monitorización en tiempo real de la actividad intelectual de sus investigadores puede resultar de gran utilidad de cara a la toma de decisiones en materia de política científica.
- Editoriales científicas: conocer los hábitos de lectura de sus usuarios les permitirá mejorar la calidad de sus productos, analizando las preferencias de lectura de los investigadores (por ejemplo pdf frente a html).

Finalmente la metodología de visualización mediante co-uso puede fácilmente extrapolarse a otras plataformas y contextos del mundo de la web social académica. La gran mayoría de las plataformas que permiten el cálculo de indicadores alométricos pueden ser susceptibles de ofrecer mapas de co-uso. Este es el caso de *Mendeley*, donde pueden elaborarse matrices de similaridad a partir de los artículos leídos por un mismo lector. La expansión de la red social así como el trasvase al medio digital muestran un nuevo panorama para la redefinición del concepto de uti-

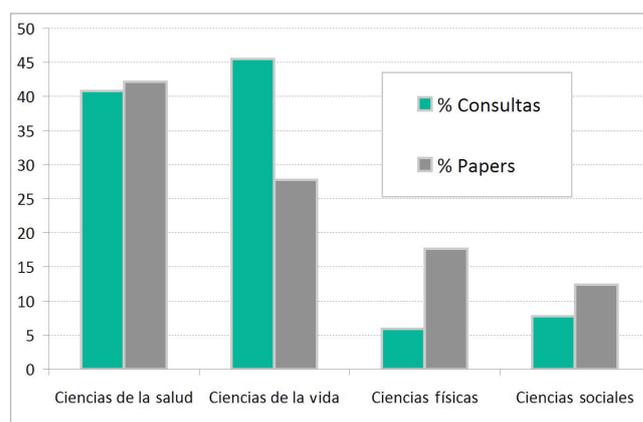


Figura 4. Comparativa entre el porcentaje de consultas realizadas a *ScienceDirect* y la producción según *Scopus* de la Universidad de Navarra en 2012

lidad de cara a la visualización de la ciencia. La utilización de nuevos indicadores ofrece nuevas vías para explorar las relaciones entre disciplinas y entre los intereses de los investigadores.

Notas

1. La información de *Elsevier bibliometric research program* está disponible en:
<http://ebrp.elsevier.com>

El proyecto en el que está basado este trabajo es *Viability of co-downloading data analysis form mapping interdisciplinary research at institutional level* cuya descripción se puede consultar en la página web anterior.

2. *Pajek* (araña, en esloveno) es un software para el análisis y visualización de redes sociales, desarrollado en la Universidad de Ljubljana, Eslovenia, por Vladimir Batagelj, Andrej Mrvar y la contribución de Matjaž Zaveršnik. Es un software libre para uso no comercial y puede descargarse de:
<http://vlado.fmf.uni-lj.si/pub/netwoks/pajek>

Los mapas basados en el uso enriquecen el análisis al incluir a otros actores pasivos en el consumo de la literatura científica como son estudiantes, profesionales o docentes

Agradecimientos

Los autores quieren agradecer a Gali Halevi y Henk Moed la cesión de los datos de uso y a Víctor Sanz Santacruz, director de la Biblioteca de la Universidad de Navarra por permitir el acceso a los mismos.

Bibliografía

Aguillo, Isidro F.; Granadino, Begoña; Ortega, José-Luis; Prieto, José-Antonio (2006). "Scientific research activity measured with cybermetrics indicators". *Journal of the American Society for Information Science and Technology*, v. 57, n. 10, pp. 1296-1306.

http://departments.icmab.es/utc/images/SEO/pdf/2006_JASIST_Webometrics.pdf

<http://dx.doi.org/10.1002/asi.20433>

Almind, Thomas; 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>

Bollen, Johan; Van-de-Sompel, Hebert; Smith, Joan A.; Luce, Rick (2005). "Toward alternative metrics of journal impact: a comparison of download and citation data". *Information processing & management*, v. 41, n. 6, pp. 1419-1440. <http://public.lanl.gov/herbertv/papers/ipm05jb-final.pdf>
<http://dx.doi.org/10.1016/j.ipm.2005.03.024>

Borrego, Ángel; Urbano, Cristóbal (2005). "Estadísticas e indicadores de rendimiento de colecciones y servicios bibliotecarios de carácter electrónico: estudio de caso de las revistas electrónicas". *El profesional de la información*, v. 14, n. 1, pp. 30-38. <http://www.elprofesionaldeinformacion.com/contenidos/2005/enero/4.pdf>

Brody, Tim; Harnad, Stevan; Carr, Leslie (2006). "Earlier web usage statistics as predictors of later citation impact". *Journal of the American Society for Information Science and Technology*, v. 57, n. 8, pp. 1060-1072. <http://dx.doi.org/10.1002/asi.20373>

Haustein, Stefanie; Siebenlist, Tobias (2011). "Applying social bookmarking data to evaluate journal usage". *Journal of informetrics*, v. 5, n. 3, pp. 446-457. <http://dx.doi.org/10.1016/j.joi.2011.04.002>

Kurz, 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>

Leydesdorff, Loet (2008). "On the normalization and visualization of author co-citation data: Salton's cosine versus the Jaccard index". *Journal of the American Society for Information Science and Technology*, v. 59, n. 1, pp. 77-85. <http://arxiv.org/abs/0911.1447>
<http://dx.doi.org/10.1002/asi.20732>

Luukkonen, Terttu; Persson, Olle; Sivertsen, Gunnar (1992). "Understanding patterns of international scientific collaboration". *Science, technology & human values*, v. 17, n. 1, pp. 101-126. <http://dx.doi.org/10.1177/016224399201700106>

Nicholas, David; Huntington, Paul; Watkinson, Anthony (2005). "Scholarly journal usage: the results of deep log analysis". *Journal of documentation*, v. 61, n. 2, pp. 248-280. <http://dx.doi.org/10.1108/00220410510585214>

Noyons, Christiaan M. (2004). "Science maps within a science policy context". En: Moed, Henk F.; Glänzel, Wolfgang;

Schmoch, Ulrich (eds). *Handbook of quantitative science and technology research*. Dordrecht: Kluwer Academic Publishers, pp. 237-255. ISBN: 978 1402027550

Priem, Jason; Taraborelli, Dario; Groth, Paul; Neylon, Cameron (2010). "Altmetrics: a manifesto". *Altmetrics.org*. <http://altmetrics.org/manifesto>

Rodríguez-Bravo, Blanca; Alvite-Díez, María-Luisa; Morán-Suárez, María-Antonia; Marraud, Gerardo (2012). "Impacto de la contratación de colecciones de revistas electrónicas en la productividad de la Universidad de Vigo". *El profesional de la información*, v. 21, n. 6, pp. 585-594. <http://dx.doi.org/10.3145/epi.2012.nov.05>

Rowlands, Ian; Nicholas, David (2007). "The missing link: journal usage metrics". *Aslib proceedings*, v. 59, n. 3, pp. 222-228. <http://dx.doi.org/10.1108/00012530710752025>

Small, Henry (1973). "Co-citation in the scientific literature: A new measure of the relationship between two documents". *Journal of the American Society for Information Science and Technology*, v. 24, n. 4, pp. 265-269. http://polaris.gseis.ucla.edu/gleazer/296_readings/small.pdf
<http://dx.doi.org/10.1002/asi.4630240406>

Tenopir, Carol; King, Donald W.; Christian, Lisa; Volentine, Rachel (en prensa). "Scholarly article seeking, reading and use: A continuing evolution from print to electronic". *Learned publishing*.

Thelwall, Mike (2008). "Bibliometrics to webometrics". *Journal of information science*, v. 34, n. 4, pp. 605-621. <http://www.scit.wlv.ac.uk/~cm1993/papers/JIS-0642-v4-Bibliometrics-to-Webometrics.pdf>
<http://dx.doi.org/10.1177/0165551507087238>

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: nuevos indicadores para la comunicación científica en la web 2.0". *Comunicar*, v. 41, pp. 53-60. <http://dx.doi.org/10.3916/C41-2013-05>

Torres-Salinas, Daniel; Delgado-López-Cózar, Emilio; García-Moreno-Torres, José; Herrera, Francisco (2011). "Rankings ISI de las universidades españolas según campos científicos: descripción y resultados". *El profesional de la información*, v. 20, n.1, pp. 111-118. <http://eprints.rclis.org/15325>
<http://dx.doi.org/10.3145/epi.2011.ene.14>

Van-Noorden, Richard (2014). "Scientists may be reaching a peak in reading habits". *Nature*. <http://dx.doi.org/10.1038/nature.2014.14658>