

Llana Martín, Amelia de la; Avila Barredo, Pilar; Barredo Sobrino, M<sup>a</sup> Pilar; Martínez García, Francisco\*

Universidad Autónoma de Madrid. Biblioteca Medicina.

\* EAP; CS Colmenar Viejo Sur; Area V IMSALUD, Madrid

C/Arzobispo Morcillo s/n. 28029 Madrid

91 4972414

[amelia.delallana@uam.es](mailto:amelia.delallana@uam.es)

## LA IMPORTANCIA DE LOS RECURSOS ELECTRÓNICOS EN LA CITACIÓN BIBLIOGRÁFICA

### Resumen

Objetivos: Estudiar la tendencia y la repercusión de la citación de referencias de páginas Web en la bibliografía de los artículos científicos publicados en la revista española con mayor repercusión internacional.

Métodos Se extrajeron todas las referencias bibliográficas electrónicas de los artículos publicados en un año en la revista Medicina Clínica para su análisis: comprobando las normas de citación, así como la mención de la fecha de acceso al documento, fecha de recepción del artículo y fecha de publicación. Se procedió al acceso en Internet de las citadas referencias desde dos IP distintas en, al menos, dos ocasiones con un intervalo de tiempo superior a dos semanas.

Resultados: De un total de 504 artículos publicados en el año 2003 con, al menos, una cita bibliográfica, se obtuvieron 11622 referencias bibliográficas, de las cuales, 164 fueron electrónicas correspondiendo a 73 artículos. Por secciones de la revista, los Artículos Especiales y las Editoriales fueron el tipo de artículo con mayor porcentaje de citas electrónicas: 41 y 29%, respectivamente. El intento de acceso a las referencias dio algún mensaje de error en el 43% de los casos, siendo el error 404 el mensaje más frecuente (77%). La principal causa de dicho error fueron las incorrecciones en la escritura de la dirección URL. Además, el error 404 fue más frecuente a medida que aumentó el tiempo de la consulta en Internet con respecto a la fecha de publicación

Discusión: El aumento de la citación de recursos electrónicos en la bibliografía de los artículos científicos es cada vez más frecuente. No obstante, la desaparición de muchos de estos recursos en Internet aumenta con el paso del tiempo. Si la finalidad de una cita bibliográfica es la mención de las fuentes consultadas por el autor para

elaborar sus propias hipótesis y para que otros autores puedan, a su vez, consultarlas, ¿qué sentido tiene la cita de documentos que con el tiempo, en un promedio dos años, pueden desaparecer? Se necesitan medidas para preservar dichos recursos ya que las iniciativas existentes, no resuelven completamente el problema.

**Palabras clave:** (MeSH)

Internet/utilización; Internet/tendencias; Bibliografía; Almacenamiento y recuperación de la información /normas

(Internet/utilization; Internet/trends; Bibliography; Information Storage and Retrieval/standards)

## Introducción

La rápida difusión de la información a través de Internet, y su obligada mención en los trabajos de investigación, llevó a la necesidad de incorporar la citación de estos recursos en los artículos científicos, incluso con anterioridad a la normalización en la forma de citar dichos recursos. La manera de citar estos recursos en las Revistas Biomédicas data de la 5ª edición del año 1997 de las Instrucciones elaboradas por el Comité Internacional de Editores de Revistas Médicas<sup>1</sup>. Desde entonces, y con la sucesiva incorporación de nuevas revistas científicas a dicho Comité, el uso de referencias bibliográficas de páginas web se ha generalizado. No obstante, el uso de éste material, tiene repercusiones importantes como consecuencia del propio carácter efímero de la información disponible en Internet. Si la red crece a un ritmo vertiginoso, también es cierto que muchos recursos desaparecen de forma rápida: el tiempo medio de *persistencia* de una página web se ha establecido en algo menos de dos años<sup>2-4</sup>, aunque podría ser variable en función del tipo de material<sup>5</sup>. La principal consecuencia de esto es que, no sin cierta frecuencia, al intentar acceder a un recurso electrónico nos aparezca el denominado error 404: no existe o no encontrado<sup>6-10</sup>. Esto plantea una cuestión muy importante: si la finalidad de una referencia bibliográfica en un artículo científico es la mención de las fuentes sobre las que el autor elabora sus propias hipótesis para que otros autores puedan consultarlas, ¿qué sentido tiene la citación de un documento que, con el tiempo (en promedio unos dos años), puede desaparecer? ¿Cuestiona este hecho el uso del material electrónico como fuente bibliográfica citable en un artículo?<sup>11</sup>. Para hacernos una idea de este fenómeno nosotros hemos estudiado la tendencia y la repercusión de la citación de referencias de páginas Web en la bibliografía de los artículos científicos publicados en la revista española con mayor repercusión internacional.

## Material y Métodos

Las referencias bibliográficas de los artículos científicos publicados en la revista Medicina Clínica durante el año 2003 fueron sometidos a un estudio pormenorizado. Se consideraron todos aquellos artículos publicados, de todas las secciones de la revista, que contuvieran al menos una referencia bibliográfica al final del artículo. Las

citas con referencia a un documento electrónico publicado en Internet fueron identificadas y analizadas. Se consideraron sólo aquellas referencias que contuvieran una dirección *URL* (*Uniform Resource Locator*) del tipo *http* o *www* y se descartaron aquellas correspondientes a la prepublicación en la web de artículos de revistas en formato papel así como una referencia a un correo electrónico.

Las referencias electrónicas recuperadas fueron procesadas con el fin de acceder al recurso citado. Se intentó su acceso en dos ocasiones, con dos semanas de diferencia y desde dos direcciones IP distintas. Los documentos con un inicial mensaje de error fueron explorados, a través de sus autores, título, fuentes, etc., en el buscador *Google* (<http://www.google.es>) para su identificación. Se analizó, además, el tiempo transcurrido entre la fecha de recepción del manuscrito y la fecha de publicación del artículo en aquellos casos en que figuraron estos datos (por ejemplo, en las Cartas al Editor no figura la fecha de envío). Además, se procedió a la búsqueda del recurso, aparentemente inexistente, en el *Internet Archive* (<http://www.archive.org>).

## Resultados

De un total de 504 artículos publicados en el año 2003, con al menos una referencia bibliográfica, y un total de 11622 referencias bibliográficas, se obtuvieron 164 referencias electrónicas procedentes de 73 artículos. Los resultados se presentan en la Tabla 1 y se comparan con los de otras revistas internacionales en lengua inglesa de elevado prestigio: *NEJM*, *JAMA* y *Science*<sup>10</sup>. Como se puede apreciar en dicha tabla, los resultados para Medicina Clínica son muy similares a los de la revista *NEJM*. Si bien de forma global el número de referencias electrónicas puede parecer pequeño, el porcentaje de artículos con al menos una referencia electrónica representa ya el 14,5% del total de los artículos publicados en esta revista, y con una tendencia progresiva a ir en aumento. Por secciones de la revista, los Artículos Especiales y las Editoriales son los que con mayor frecuencia incluyen referencias electrónicas: 41% y 29%, respectivamente. En el lado opuesto, las Cartas al Editor son las que menos incluyen este tipo de referencias (6%), cifras muy similares a lo que ocurre en *NEJM*.

Tabla 1: Comparación de nuestros resultados obtenidos de Medicina Clínica (año 2003) con el de otras revistas de elevado factor de impacto.

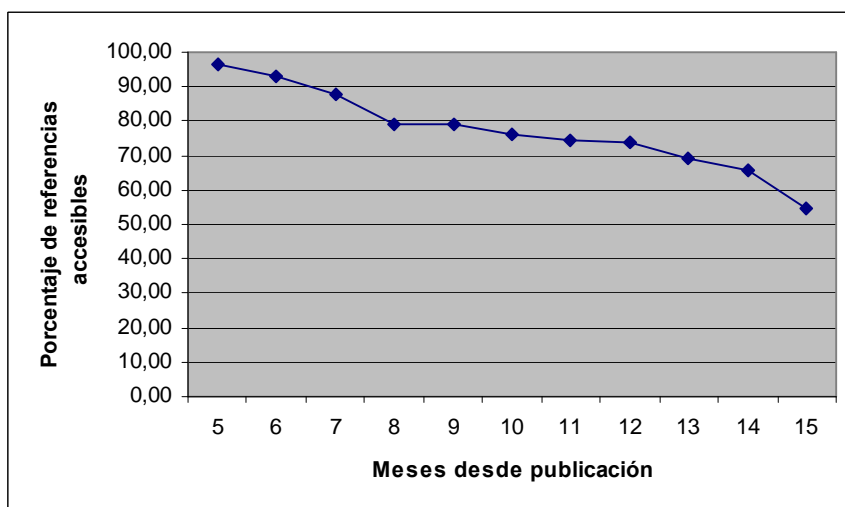
	Med Clin	NEJM*	JAMA*	Science*
Números analizados	40	18	18	18
Artículos	504	290	300	472
Total de Referencias	11622	6704	7045	11799
Referencias por Artículo (%)	23,06	23,12	23,48	25
Referencias electrónicas	164	97	189	386
Referencias electrónicas (%)	1,41	1,45	2,68	3,27
Artículos con Referencias electrónicas	73	39	93	190
Artículos con Referencias electrónicas (%)	14,48	13,45	31	40,25
Referencias electrónicas por artículo	2,25	2,49	2,03	2,03

\*: los datos para las revistas mencionadas han sido obtenidos de la referencia nº 10.

Se comprobó que, el 43% de las referencias daba algún tipo de mensaje de error. El error más frecuente fue el error 404 en el 77% de los casos. Un análisis pormenorizado de las direcciones *URL* detectó un 31% de errores de escritura en las mismas, que fueron el motivo del citado mensaje en el 73% de los casos.

En el 19% de las citas que dieron un mensaje de error, se pudo recuperar el documento citado desde otro dominio distinto al mencionado en la bibliografía, y en un 29% de los casos se constató un cambio importante de la dirección *URL* probablemente debido a actualizaciones internas del propio recurso. El mensaje de error fue más frecuente a medida que aumentó el tiempo desde la fecha de publicación del artículo (Ver Figura 1): para los artículos que contenían al menos una referencia electrónica, el tiempo medio fue de  $174 \pm 90$  días. Pero si se analiza el tiempo transcurrido entre la fecha de acceso al recurso por el/los autor/es, en los casos en que figuraba este dato, y la fecha final de publicación del artículo, este tiempo se incrementó a  $257 \pm 174$  días.

Figura 1. Evolución de la accesibilidad de los recursos electrónicos citados en función del tiempo tras la publicación del artículo fuente.



La búsqueda de las referencias electrónicas de Medicina Clínica que dieron algún mensaje de error, pudieron ser identificadas, bien en *Google*, bien en el *Internet Archive*, salvo cuatro citas.

## Discusión

Hemos podido comprobar que las dificultades para acceder a un recurso electrónico previamente citado por otro autor en un artículo científico, son muy importantes. El principal motivo es el error tipográfico en la dirección *URL* de la cita. En este sentido, habría que insistir en la meticulosidad necesaria por parte de los autores a la hora de transcribir fielmente este tipo de direcciones. Pero, aún así, las probabilidades de no obtener respuesta del servidor a la hora de intentar acceder al recurso, aumentan con el paso del tiempo, como consecuencia de un cambio de la dirección *URL*, de las actualizaciones del servidor o de la desaparición de la red del recurso. ¿Qué posibilidades se plantean para intentar subsanar esta cuestión?

Se ha propuesto el *DOI* (*Digital Object Identifier*) como alternativa a la *URL*<sup>12</sup>, pero esta solución, no parece exenta de problemas propios. Es uno de los sistemas que ha surgido en la red para identificar al propietario de los derechos de autor de las obras que circulan por ella, y también para permitir la compra de obras enteras o de partes de ellas. El *DOI* es un sello de identidad que se puede añadir a cualquier tipo de material que sea objeto de propiedad intelectual. Además, puede incluir otras informaciones como el *ISBN*, o claves para la clasificación temática de la obra. Una ventaja indudable de este sistema es la posibilidad de acceder al documento,

independientemente de su cambio de localización, a través de dicho número de identificación (Cross Ref). No obstante, este sistema no resuelve el problema de los documentos que desaparecen de la red. Otro inconveniente sería el coste que supondría para pequeños editores y autores que de forma individual publiquen sus trabajos en Internet. La ventaja indiscutible de la *URL* sobre el *DOI* sigue siendo que es, precisamente a través de la *URL* como nosotros podemos acceder o seguir su rastro en el *Internet Archive*. Sin embargo, el principal problema radica en el elevado porcentaje de *URL* mal escritas y que, por ello mismo, no pueden ser recuperadas sin una comprobación previa.

Finalmente, algunos autores han sugerido la implicación de los Editores de las revistas en la salvaguarda de los recursos citados<sup>10</sup>. Esta posibilidad, no parece que sea muy bien acogida por dichos editores ya que ello supone dedicación de más recursos a algo que, en definitiva, parece ser responsabilidad de los autores. Sin embargo, la credibilidad de la información por Internet depende de la capacidad de crear recursos que permitan preservar dicha información y asegurar su perdurabilidad. En este sentido, se están llevando a cabo diferentes iniciativas<sup>13-16</sup>. Es probable que, con el tiempo, alguna de ellas se imponga sobre las demás, pero, hoy por hoy, todas tienen ventajas e inconvenientes. Una posibilidad podría ser que los Editores actuaran como repositorios provisionales, solicitando, junto con los manuscritos enviados, los recursos electrónicos consultados en un formato digital. Una vez publicado el artículo, el material custodiado podría enviarse a un servidor creado para tal fin e independiente de las Editoriales, donde el material recibido sería registrado y archivado en una base de datos que podría ser accesible, de forma gratuita, para cualquier autor que lo solicitase, siempre que el documento original no pudiese ser accesible desde la red o produjese un mensaje de error como consecuencia de un cambio de dirección o su desaparición de Internet. De esta manera se evitaría cualquier posible conflicto de intereses con respecto a los derechos de reproducción o copia, puesto que el material sería tratado de forma similar al de los artículos científicos que gestiona cualquier biblioteca cuando le es solicitado por alguno de sus usuarios.

## Bibliografía

1. Comité Internacional de Editores de Revistas Médicas. Requisitos de uniformidad para manuscritos presentados para publicación en revistas biomédicas. *Med Clin (Barc)*. 1997; 109:756-763.
2. Koehler W. Digital libraries and World Wide Web sites and page persistence. *Information Research* 1999; 4(4), [consultado 18-06-2004]: Disponible en: <http://informationr.net/ir/4-4/paper60.html>.
3. Koehler W. Web page change and persistence - a four-year longitudinal study,. *Journal of the American Society for Information Science and Technology*. 2002; 53:162-171.
4. Koehler W. A longitudinal study of Web pages continued: a consideration of document persistence. *Information Research* 2004; 9(2), [consultado 15-06-2004]: Disponible en: <http://informationr.net/ir/9-2/paper174.html>.
5. Nelson M, Allen B. Object persistence and availability in digital libraries. *D-Lib Magazine* 2002; 8, [consultado 15-06-2004]: Disponible en: <http://www.dlib.org/dlib/january02/nelson/01nelson.html>.
6. Benbow SMP. File Not Found: the problem of changing URLs for the World Wide Web. *Internet Research: Network Applications and Policy* 1998; 8, 247-250.
7. Kiley R. Archiving the Web. *Health Information on the Internet*. 2000; 18, [consultado 15-06-2004]: Disponible en: <http://www.rsmppress.co.uk/hii/issue18.pdf>.
8. Schloman BF. Now you see it, now you don't: the ephemeral nature of digital information. *Online J Issues Nurs*. 2003;8:7 [consultado 15-06-2004]: Disponible en: [http://nursingworld.org/ojin/infocol/info\\_11.htm](http://nursingworld.org/ojin/infocol/info_11.htm).
9. Martínez García F, de la Llana Martín A, Gil García MA, Barredo Sobrino P. Citación de páginas Web y error 404. *Med Clin (Barc)* 2003; 121:676-7.
10. Dellavalle RP, Hester EJ, Heilig LF, Drake AL, Kuntzman JW, Graber M, Schilling LM. Information science. Going, going, gone: lost Internet references. *Science*. 2003;302:787-8.
11. Kitchens JD, Mosley PA. Error 404: or, WWWhat is the shelf-life of printed Internet guides? *Library Collections, Acquisitions & Technical Services* 2000; 24: 467-478.
12. The Benefits of DOI Linking vs. URL Linking. *Content Directions*, 2003. [consultado 18-06-2004]: Disponible en: [http://doi.contentdirections.com/edudoc/Benefits\\_Doi\\_vs\\_URL.pdf](http://doi.contentdirections.com/edudoc/Benefits_Doi_vs_URL.pdf) .



13. Germain CA. URLs: uniform resource locators or unreliable reliable resource locators? *College and Research Libraries* 2000; 61: 359-365.
14. Reich V, Rosenthal D. Preserving today's scientific record for tomorrow. *BMJ* 2004; 328:61-62.
15. Lawrence S, Coetzee F, Glover E, Pennock D, Flake G, Nielsen F, Krovetz B, Kruger A, Giles L. Persistence of web references in scientific research. *IEEE Computer* 2001; 34:26-31.
16. Veronin M. 'URL 404 File Not Found': dealing with the transient nature of the Web. *J Audiov Media Med.* 2003; 26:153-5.