

Opening the non-open access medical journals: Internet-based sharing of journal articles on a medical web site

K Masters

Citation

K Masters. *Opening the non-open access medical journals: Internet-based sharing of journal articles on a medical web site.* The Internet Journal of Medical Informatics. 2008 Volume 5 Number 1.

Abstract

Introduction: Open access (AO) journals are freely available, but non-open access (NOA) journals are available only through payment. Similar to the music industry, one might expect a sharing of NOA articles on the Internet. This paper investigates a site facilitating such sharing amongst medical professionals. **Method:** A six-month snap-shot (25 May to 24 November 2008) of activities on the site. **Results:** Total articles requested: 6,587; total found: 5,464 (82.9%). Mean number of views of each article: 4.47. Total estimated saving (or loss): \$1.4M for the year of 2008. Nature articles were the most highly requested, followed by Science, and other major medical journals. **Discussion and Conclusions:** This method of accessing data is highly effective, but issues are raised. Ethical issues and financial implications are the most important. NOA journal publishers should recognise the problem, research its size and implications, but the discussions must occur in the open access area.

There is no theory of protecting content other than keeping secrets – Steve Jobs, Apple. [1]

INTRODUCTION

BACKGROUND

Although there are many definitions of “Open Access” (OA) journals, a useful guide is that offered by the “Budapest Open Access Initiative,” (BOAI) which speaks of “The literature that should be freely accessible online [...] which scholars give to the world without expectation of payment” [2]. This description fits best with Willinsky’s definition of an “unqualified open access journal” [3] which allows its articles to be available “immediately, completely, and exclusively free-to-read” [3].

OA journal publishing has strong international support amongst researchers and librarians [4-6], and even amongst Research Councils and Institutes [7; 8]. At the time of writing this article, some 5,000 individuals and more than 460 organisations had added their signature supporting the BOAI [2].

There is a growing body of evidence to suggest that OA journals are being cited as much as, and even more than, non-open access (NOA) journals [9-11]. These and other issues, such as the principle of free access, faster publishing times, wider readership, and greater number of citations, are

strong motivators to publishing in OA journals [12; 13].

In OA journals, the publishing costs (or a portion of them) are sometimes carried by the authors, or their institutions, or by journal advertising [14; 15]. The author-pay models are not popular amongst researchers [16], and there is some discussion about the possible disadvantages that this model has for researchers in the developing world [17].

The path of OA publishing has not been smooth, and the debates around OA publishing continue. Researchers who do not publish in OA journals have cited unfamiliarity with OA, low prestige and impact, author payment issues, and low readership [12; 13]. The arguing points include copyright and other legal issues, peer-review, the value of OA articles, business models, publishing costs, technology infrastructure, business models, indexing services, standards, rewards for researchers and marketing [4; 5; 9; 18-20].

In the midst of this are the researchers who want access to the information in NOA journals. The term “researchers,” of course, has a very broad application. The Public Library of Sciences (PLoS), now a publisher of the OA journal PLoS Biology, aims at a “much wider audience, including millions of students, teachers, physicians, scientists, and other potential readers, who do not have access to a research library that can afford to pay for journal subscriptions” [21]. While the term “digital divide” is usually used to apply to

the divide between those who have access to the Internet and those who don't, there also appears to be a digital divide between those who have access to NOA journals and those who do not.

To those who do not have access to NOA journals, the information is, as it were, a closed book.

Unless they take matters into their own hands.

The environment that facilitated the growth of OA publishing has been the Internet. Simultaneously, the Internet has allowed for the dissemination and sharing of information on an unprecedented scale. The music industry was a prime example of this sharing, when the sharing (frequently illegally) of digitised music exploded. Given that so many journal articles, both OA and NOA, are now available electronically, it is not difficult to imagine a scenario in which researchers share NOA articles with others, without the permission of the authors, publishers or copyright holders. These researchers would be people who have access (usually through institutions) to journals, and they use that access to share the articles with other researchers who do not have access to those journals.

Indeed, web sites that facilitate this exchange do exist. Question that need to be answered include: How many articles are affected? How many journals are affected? Which journals are most commonly affected? What is the estimated cost to these journals? This paper examines a site that is used for such a purpose, to begin answering these questions.

THE SETTING

The web site to be discussed (referred to hereafter as "the site") is aimed specifically at medical professionals and students. At the time of this study (January 2009), the site had a total of 127,626 registers users, and nearly 300,000 postings in the electronic discussion forums. Although aimed at the medical field, the site did not prevent non-medical persons from joining.

The site had a wide range of electronic discussion forums that one might expect to find on a medical site, including "Medical Education Forums", "Allied Health & Nursing Forums", "Medical Student Forums" and "Physician and Resident Forums." In addition, however, there was a set of forums collectively labelled "Electronic Library." "Electronic Library" contained forums and sub-forums aimed specifically at the sharing of information that one

might typically find in a library. This paper focuses on a sub-forum named "Databases & Journals – Requests and Enquiries."

In "Databases & Journals – Requests and Enquiries", users who do not have access to a particular NOA journal or conference proceedings post a request for a desired article. Other users who have access to those journals then download the requested articles from those journals, and make them available to the requester by posting the articles either into the forum or to a publicly-accessible web site. Although many participants on this site are students, other users have identified themselves as practicing professionals or academic staff.

Strict rules govern the requests. For instance, only three articles may be requested by each user each day, and the hypertext links to the articles must be submitted in the request. There are, however, no rules concerning the type of journals, licences, etc, of the articles requested or posted. Naturally, not only does the requester have access to the downloaded article: anyone reading the forum also has access to the article, and, in most cases, the site tracks the total number of times an article has been accessed or 'viewed.'

Until January 2009, the site was publicly available. Until then, although all the forums could be read by guests, full participation required a free registration, requiring merely a user name and an email address. (The site and forum messages were routinely indexed by search engines such as Google). Since January 2009, the site has been available to a small group only, and has also frequently been offline. This study deals with data that were publicly available during 2008.

The aim of this study was to investigate the interactions in the "Databases & Journals – Requests and Enquiries" forum and to report on the number of requests and postings, the journals from which articles are being accessed, and to comment on the broader implications of the findings.

METHOD

The last posting to the forum was on 24 November 2008. A 6-month snapshot (25 May to 24 November) of all requests for journal articles and conference proceedings was taken. For each article, the bibliographic details, and whether or not the article was successfully found, were captured. In addition, where possible, the number of times the article had

been viewed by participants was also captured. (Occasional requests for e-books were ignored). The data were placed into an Excel spreadsheet. From the data, the following was determined: total number of articles requested, total successfully supplied, total views, journals from which the most articles had been sought.

Permission to search the forum was not required. Using [22] as a guide, it was established that:

- All postings to the site were publicly visible (registration was not required for viewing. Registration was necessary for posting only).
- The site was searched and indexed by general search engines, such as Google.
- The site has a large number of registered users (127,626).
- Nothing in the site's rules restricted access to any particular groups. In fact, access was encouraged.
- No intrusive or identifying information (including direct quotations) about any of the participants was researched.

RESULTS

According to the site's main page, on 10 January 2009, the 127,626 registered users had created 30,558 threads and 298,280 posts. The most users online simultaneously had been 3,212 on 20 November 2008.

ARTICLES AND VIEWS

Within the six-month period, the total number of unique articles requested in the sub-forum Databases & Journals – Requests and Enquiries" was 6,587. Of the 6,587 articles, 6,463 (98.1%) were in English, 68 in German, 42 in French, and 14 in other languages. The total number of unique articles returned (i.e. found by other users) was 5,464, giving a success rate of 82.95%. The oldest article was from Science, dated 1884 [23].

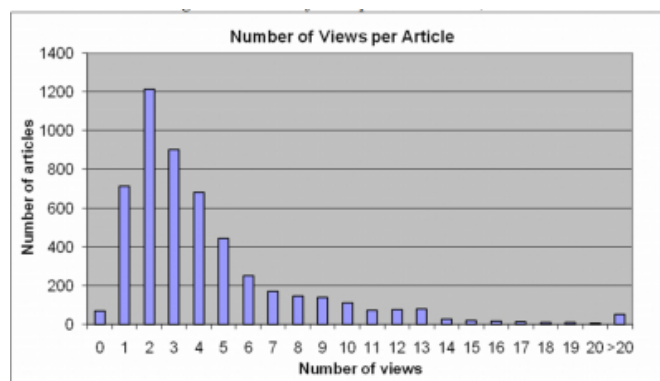
The success rate of 83% would probably have been higher if the requesters had more frequently followed the basic rules of the site, which included requesting a maximum of three articles per day, supplying direct links to the articles, and ensuring that the text was available electronically. Although the rule of three articles per day was usually strictly enforced, there were occasions when people requested and

obtained a greater number of articles. This included requests for entire issues.

Of the 5,464 articles returned, the number of 'views' could be determined on 5,251 articles (96.1%). (Others had been placed in formats in which the number of views could not be determined). From these 5,251 articles, the total number of views was 23,461, the minimum number of views was 0 (for 69 articles), and the maximum was 177 [24]. The mean number of views was 4.47, the median 3, and the mode 2. These last three descriptive statistics, however, do not show the skewness of the views – this is shown in Figure 1 below:

Figure 1

Figure 1: Number of views per article. N=5,251.



It is also obvious that, although many of the articles are viewed by participants who did not originally request the article, generally, the article is viewed by the requester, and only a few others.

A question to consider, then, is, how much money was saved by the people who viewed the articles? (Conversely, how much was lost by the publishers?) The prices of individual articles ranges. Some are available for as little as \$7.00, and others for as much as \$50.00 for academics, and \$100.00 for non-academics (e.g. [25]). At an approximate value of \$30.00 per article, the 23,461 views of the articles amount to \$703,830.00 over the six-month period. Extrapolated across a year, this translates into a little more than \$1.4M.

JOURNALS AND CONFERENCES

Articles from a total of 2,867 journals and conferences were requested. The 22 journals with the highest number of requests are summarised in Table 1 below:

Figure 2

Table 1: Table showing the 22 journals with the highest number of requests, ranked according to number of unique articles requested.

Rank	Journal Name	Articles Requested	Articles Received	Mean no. of views per article
1	Nature ¹	118	118	5.0
2	Science	41	39	4.2
3	Lancet	40	36	3.0
4	Neurology	36	24	3.7
5	Gastrointest Endoscopy Clin N Am ¹	33	13	12.7
6	Journal of Neurology	31	31	2.5
7	Am J Gastroenterology	25	24	6.3
7	J. Agric. Food Chem.	25	24	5.8
7	Spine	25	24	4.0
10	Planta Medica	24	18	4.9
11	J History of the Neurosciences	22	22	3.5
11	Journal of Urology	22	17	3.1
13	British Journal of Dermatology	21	18	2.6
15	Urologic Clinics of North America	19	19	2.9
16	Fuzzy Sets and Systems	19	18	3.3
17	J Geriatr Psychiatry Neurol ¹	18	17	5.6
18	New England Journal of Medicine	17	13	6.4
19	Appl. Biochemistry & Biotechnology	16	14	3.9
19	Diseases of the Colon & Rectum	16	15	3.3
19	Journal of Personality Assessment	16	16	2.4
19	Analytical Biochemistry	16	16	5.9

At the top of the list are two general scientific journals, Nature and Science. Not surprisingly, some of the major medical journals are also found in this list. Other journals that one might expect to find (e.g. BMJ), allow free access to their articles after 6 or 12 months, and so would not necessarily be placed in this list. (As a matter of interest, 11 articles had been requested from BMJ, of which 9 were successfully delivered).

DISCUSSION

The results presented in this paper give an indication of what researchers without access to NOA journals can do, and are doing, to access required articles. The large number of articles requested, the wide range of journals, and the success rate (83%) indicate that this method of access is very effective.

There are, however, issues, of which the most obvious is ethical. It is not the place of this paper to take an ethical stand on the results, but the ethics of the activities should be considered. In the field of medicine, ethics plays a pivotal role, and yet the site displays activities by medical students, teachers and practicing professionals that are ethically dubious.

From the participants' comments made in the forums, however, there does not appear to be any vindictiveness on the part of the participants against the journals or holders of copyright, but a mood of togetherness, of openness and sharing, and communal assistance. Most remarkable, is that

the activity described in this paper did not occur within closed, secure, password- and firewall-protected environments, but within open environments, easily publicly accessible, and easily searchable and referenced by general search engines such as Google.

A second consideration is the financial issue. One might argue that, while \$1.4M is a sizable amount, this is spread across 2,867 journals and conference proceedings, so the loss to each journal is not substantial. In addition, there is, ironically, a benefit to the journals: their articles would possibly otherwise go unseen and uncited.

On the other hand, this article deals with one site only, and more information is needed on the number of such web sites so that an accurate impact on journal revenue can be established. Further, one may also argue that the amount is irrelevant – the revenue has been lost through theft, and that is concern enough. Finally, any possible benefit to the journals is incidental, and, while it might be some small consolation, does not make the practice acceptable.

Is there a solution?

It is obvious that the interests of legitimate enterprises must be protected, but it is equally obvious from these activities that current methods do not work. In this light, Steve Jobs' quotation at the top of this article bears repeating: "There is no theory of protecting content other than keeping secrets" This was written in 2007 in an article that can be seen as a softener for what was to follow in 2009 – relaxing the copy-protection in Apple's music. In his article, Jobs writes about the problems of keeping secrets for the protection of material, and says there "are many smart people in the world, some with a lot of time on their hands, who love to discover such secrets and publish a way for everyone to get free (and stolen) music" [1].

This applies equally to texts, and the NOA publishers need to recognise this. The exact route for a solution is unclear. What is clear, however, is that further research is required on the size of the problem, and the possible solutions. Further, much of this debate cannot occur in areas where these issues are currently being discussed – many of the most illuminating discussions are occurring in journal contributions (e.g. [6; 17]) that are not accessible, except by payment. (Those articles are, of course, freely available through sites such as the one described in this article). The real debate on open access to journals, if it is to be of value, can only occur if the access to the debate is open.

CONCLUSION

This paper has investigated a web site that allows for the sharing of non-open access journal and conference proceeding articles. Although the site covers a range of subjects, the focus is on the medical field. It has found that a large number of such articles, from a very wide range of journals, are shared. Major scientific and medical journal articles are frequently shared. There are ethical and financial issues at stake. While the solution to the problem is unclear, it is certain that the problem requires further research, and further discussion in an open access area.

CONFLICT OF INTEREST

The author is a signatory of the BOAI.

References

1. Jobs S, 2007, Thoughts on Music <<http://www.apple.com/hotnews/thoughtsonmusic/>> (Accessed 09/01/2009).
2. Budapest Open Access Initiative (BOAI), 2002, Budapest Open Access Initiative <<http://www.soros.org/openaccess/read.shtml> Budapest, Hungary> (Accessed 07/01/2009).
3. Willinsky J. The nine flavours of open access scholarly publishing. *J Postgraduate Med* 2003;49:263-7.
4. May C. The Academy's new electronic order? Open source journals and publishing Political Science. *European Political Science* 2005;4(1):14-24.
5. Plutchak TS. Embracing open access. *J Med Libr Assoc* 2004;92(1):1-3.
6. Tamber PS, Godlee F, Newmark P. Open access to peer-reviewed research: Making it happen. *The Lancet* 2003;362:1575-7.
7. BioMed Central, 2009, 11 February, Last UK Research Council Mandates Open Access <<http://www.biomedcentral.com/html/info/about/updatesarchive/update11-02-09.html>> (Accessed 07/03/2009).
8. Canadian Cancer Society Research Institute, 2009, Open Access Policy <http://cancer.ca/research/policies_and_administration/policy/open_access.aspx?sc_lang=en> (Accessed 11/03/2009).
9. Antelman K. Do open access articles have a greater research impact? *College & Research Libraries News* 2004;65(5):372-82.
10. Eysenbach G. The open access advantage. *J Med Internet Res* 2006;8(2):e8.
11. Eysenbach G. Citation advantage of open access articles. *PLoS Biology* 2006;4(5):e157.
12. Swan A, Brown S. Authors and open access publishing. *Learned Publishing* 2004;17:219-24.
13. Swan A, Brown S. Open access self-archiving: An author study. <http://www.jisc.ac.uk/uploaded_documents/Open%20Access%20SelfArchiving-an%20author%20study.pdf> 2005, (Accessed 08/01/2009).
14. Delamothe T, Smith R. Open access publishing takes off. *BMJ* 2004;328:1-3.
15. Eysenbach G. Peer review and publication of research protocols and proposals: A role for open access journals. *J Med Internet Res* 2004;6(3):e37.
16. Schroter S, Tite L, Smith R. Perceptions of open access publishing: Interviews with journal authors. *British Medical Journal* 2005;330(7474):756.
17. Gadagkar R. Open-access more harm than good in developing world. *Nature* 2008;453:450.
18. Björk B-C. Open Access to Scientific Publications - an Analysis of the Barriers to Change? *Information Research* 2004;9(2) <<http://InformationR.net/ir/9-2/paper170.html>> (Accessed 07/01/2009).
19. Butler D. Who will pay for open access? *Nature* 2003;425(6958):554-5.
20. Crawford BD. Open-access publishing: Where is the value? *The Lancet* 2003;362:1578-80.
21. Brown PO, Eisen MB, Varmus HE. Why PLoS Became a Publisher. *PLoS Biol* 2003;1(1):e36.
22. Eysenbach G, Till JE. Ethical issues in qualitative research on internet communities. *BMJ* 2001;323:1103-5.
23. Ryder J. The pedunculated lateral-line organs of *Gastrostomus*. *Science* 1884;3(48):5.
24. Hosemann W, Gross R, Göde U, Kühnel T, Röckelein G. The anterior sphenoid wall: Relative anatomy for sphenoidotomy. *American Journal of Rhinology* 1995;9(3):137-44.
25. Sheng C, Zhang W, Zhang M, Song Y. Homology Modeling of Lanosterol 14 α -Demethylase of *Candida albicans* and *Aspergillus fumigatus* and Insights into the Enzyme-Substrate Interactions (p. 91-100). *Journal of Biomolecular Structure and Dynamics* 2004;4:91-100.

Author Information

Ken Masters, MA, ITHHealthEd