End User Tools for Evaluating Scholarly Content

Carol Anne Meyer
CrossRef, cmeyer@crossref.org

Follow this and additional works at: <https://docs.lib.purdue.edu/charleston>

An indexed, print copy of the Proceedings is also available for purchase at:

<http://www.thepress.purdue.edu/series/charleston>.

You may also be interested in the new series, Charleston Insights in Library, Archival, and Information Sciences. Find out more at: <http://www.thepress.purdue.edu/series/charleston-insights-library-archival-and-information-sciences>.

Carol Anne Meyer, "End User Tools for Evaluating Scholarly Content" (2011). *Proceedings of the Charleston Library Conference*.
<http://dx.doi.org/10.5703/1288284314959>

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.

End User Tools for Evaluating Scholarly Content

Carol Anne Meyer, *Business Development and Marketing*, CrossRef

Abstract:

The existence of multiple versions of scholarly content (from author websites, institutional repositories, government archives, subject-specific digital libraries, aggregator collections and publisher websites) make it difficult for users to locate the most recent version of a document or to ascertain if the document has had any updates or even been retracted. This session describes tools for end users to evaluate the content they come across to make sure they are citing the most authoritative version of the content available. The reader will learn about the CrossMark version of record service and the importance of educating users about how to locate current information.

What Happens When Scholarly Content Changes?

Marc Hauser was a well-known primate researcher at Harvard University. He was on the faculty in the Psychology Department, ran the Cognitive Evolution Laboratory, and was a popular teacher and a leader in his field. Not unusually, the lab had its own web site, and publications of affiliate faculty and researchers were listed and copies were hosted on this site. One of the papers on the siteⁱ, a 2002 article co-authored by Hauser, became the focus of a year-long Harvard University ethics investigation. The results of the investigation were unfortunate for Hauser. The paper, originally published in the Elsevier journal *Cognition*, was retracted, the case was made public in the Boston Globe,ⁱⁱ he was suspended, and he ultimately left Harvard.

Published scholarly research is supposed to be self-correcting. But what happens when articles that have been corrected, updated or even retracted, as in the case of this paper, are still available in their original form? I first came across the case of Marc Hauser in December 2010. At that time, the retracted paper was still available on the Cognitive Evolution Labs Web site as a PDF. The article was formatted for publication in *Cognition*, and was clearly labeled "Article in Press." There was no mention of a retraction or any investigation. In preparation for an article I was writing and several subsequent presentations, I checked the lab's site in January, March, June, and November of 2011. It wasn't until November that the PDF was removed from the lab's site. In fact, the entire site had disappeared by that time. At no time was there any indication that the paper had been retracted.

If you follow the CrossRef DOI link to the paper, you find that the publisher has clearly marked it "Re-

tracted" by adding the word to the title. So the researcher who is using the publisher version will pretty clearly be able to learn about the retraction.

If, on the other hand, a researcher starts at Google, and we heard at another session at this conference a study of librarians, faculty and students that reminds us that that is exactly what academics do,ⁱⁱⁱ and he or she types in a search for the article "Rule Learning by Cotton-top Tamarins" the first result is the PDF of this article, on another site (CiteSeer), which does not indicate the article has been retracted. The second result is to PubMed, which does have a way to notify researchers of retractions, but not of other types of corrections.^{iv}

Of course Google results vary. When I retried this search in preparing this article for the proceedings version of this talk, the first result was in a group from Google Scholar, and had a link to the properly identified retracted article at *Cognition*, but one of the other 13 results for the same citation was a still-available PDF of the paper on the co-author's departmental web site at New York University.^v In today's world scholars have a broad array of sources for the research they rely on, from author home pages, to institutional repositories, web aggregators, government repositories, multiple e-book formats and readers, and federated search engines. This array is convenient, but it also presents challenges for notifying readers when something has happened to a document after peer review and publication.

The Concerns of Librarians

About a year ago, CrossRef conducted several focus groups with librarians. We wanted to know whether they had concerns about different versions of arti-

cles proliferating on the web. Librarians told us that they did not believe that users are always clear about which version of a document they are reading. They noted, as we have discussed, that both Google and Google Scholar return search results from multiple versions of documents. They also told us that they do not have time or resources to track post-publication changes at the article level for research they may have helped patrons identify and acquire. They expressed concern that readers might cite “incorrect” versions instead of an official version of record. They easily understood the concept of a version of record, though they may not have heard that exact terminology before. They also noted that researchers complain to them that Google does not return up-to-date versions of the researcher’s own articles. In other words, researchers will search for their own name (sometimes called ego-surfing) and find inaccurate or outdated versions of their own papers. Finally, librarians expressed concern that usage statistics might not accurately represent resource usage if multiple versions are available.

Communicating Corrections

So how do publishers communicate corrections to scholarly research? A colleague of mine, Geoffrey Bilder, noted on a visit to a well-known medical publisher a bulletin board on which a paper correction—a product recall—to an important reference work was tacked up. This is not an effective communication method in a world where information wants to be electronic.

Reputable publishers do notify their users of updates. But they do it in a non-standard way. *Science* links to corrections in red in a left-handed side bar. *BMC Genomics* chooses to place a blue banner in the center pain above the title of an article. In *PLOS Medicine* the user finds the correction in a bar on the right side under a heading called “Related Content” and they have to scroll down the page to find it. And none of these publishers show an indication

on the PDF versions of their papers that an update has been issued. CrossRef content includes documents from 26,000 journals and 247,000 books. It is not hard to imagine the variety of ways publishers may choose to notify their users of important changes. Nor is it hard to imagine that readers might easily miss those notifications.

A lot of attention has been given to retractions recently. They make for interesting and controversial stories in the news. But retractions are not the only types of changes that should be noted by careful researchers. Other updates might include errata, corrigenda, enhancements withdrawals, new editions, protocol updates, and notices of concern, just to name a few. Though certainly most articles do not have serious changes that would effect the interpretation of the research, it happens frequently enough to warrant concern. When content changes, readers need to know about it.

The CrossMark Version Identification Service

CrossRef, a not-for-profit association of 1300 scholarly publishers, is piloting a service to help solve this problem. Several publishers are participating in this pilot.^{vi} The CrossMark service displays a logo on participating publishers’ content (Figure 1). When a scholar sees a CrossMark logo, he or she knows that the publisher has made a commitment to keeping the document up-to-date. In addition, clicking on the logo tells the researcher several things:

- Whether there actually have been any updates, and if so, where they can be found
- If this copy of the paper is being maintained by the publisher
- Where the publisher-maintained copy is located
- Other important non-bibliographic publication record information about the document

Lithuanian Industry Personnel Value Orientation Towards Innovativeness: Analysis, Evaluation and Improvement Possibilities

Augustinas Maceika, Bronius Jančiauskas

Abstract

The authors of this article established the goal to detect and evaluate the wide possibilities for strengthening the value orientation towards innovativeness in the industrial enterprises. For this purpose exploration of the valuation on the innovativeness as a value was made. The authors of this article performed necessary research, analyses,

Figure 1

CrossMark logo displayed on Vilnius Gediminas Technical University journal page <http://dx.doi.org/10.3846/btp.2010.34>

The clickable CrossMark logo works on PDF documents as well as on publisher's HTML pages. This is particularly important for researchers who save PDFs locally on their computers or load them into paper managements systems far in advance of reading them or citing them in their own submissions. If a CrossMark logo has been applied to a PDF, then as long as a researcher is connected to the Internet, they will find updated status and record information about that document, even if an update occurred months after the researcher saved the PDF.

When a user clicks on the CrossMark logo, a pop-up box appears on top of the grayed-out browser window. The box is open to a Status tab, which either has a bright green checkmark with a message that the content is up to date (Figure 2) or a blue exclamation mark that an update is available (Figure 3). In either case, basic bibliographic information appears with the CrossRef DOI link to the original content and, in the case of a status change, the CrossRef DOI link to the update or updates. The Status tab also has a link to the publisher's specific CrossMark policy page, which explains how corrections or updates are managed at that organization.

Figure 2

Up-to-date CrossMark Status tab from The Royal Society <http://dx.doi.org/10.1098/rspb.2011.0465>.

Figure 3

Example of CrossMark Status tab with correction from the International Union of Crystallography <http://dx.doi.org/10.1107/S0907444906030915>.

If a publisher chooses to provide additional publication record information in addition to the status information, that data will be presented on the Record tab of the popup box (Figure 4). The Record tab can provide a consistent place for important publication information, in essence non-bibliographic metadata. This information might include:

- Research funding information
- Conflict of interest statements
- Publication history like submission, revision, and acceptance dates; location of related data deposits or registries

- Description of the peer review process used
- Whether the document has been screened by Crosscheck for plagiarism
- Copyright or license information

This publication record information is optional, and not pre-defined by CrossRef, though we do expect that communities of interest (particularly in related disciplines) may develop to create best practices for ways to introduce consistency.

Figure 4

Sample CrossMark Record tab from the International Union for Crystallography <http://dx.doi.org/10.1107/S0108767310044892>

Applications for CrossMark Beyond the Original Published Content

CrossRef intends to make the CrossMark metadata openly available and discoverable. The implications of this openness is that search and discovery services will be able to display CrossMark logos in their search results to differentiate content that is being maintained by the publisher from other copies. Imagine if in our searches for Hauser's cotton-top tamarin paper, Google or Google Scholar had dis-

played CrossMark logos for those instances of the paper that the publisher maintained. There would have been an easy way to tell which of the copies a researcher could rely on.

The same function could work in records and results from secondary abstracting and indexing databases, search results from library link resolvers, and results lists from federated search tools.

CrossMark Business Model and Rules

Just as CrossRef itself has no say or interest in the business model of publishers who assign DOIs (there are more not-for-profit publishers than commercial publishers among our membership, and CrossRef DOIs link to many open access documents), the CrossMark service can work as well for documents that a publisher deems exclusive to its own web site as for documents from a publisher that has very liberal reposting and republication policies. In fact, one might argue that there is even greater value to displaying a CrossMark logo on an open access document. If the publisher license allows widespread reposting, it can become even more difficult to ascertain whether there has been a correction on any particular document. But if the re-posted documents include the CrossMark logo, clicking on the logo will query the CrossMark metadata at CrossRef and bring up the current information about that document, no matter when it was posted to the third-party site, and no matter if that site is an author home page, an institutional repository, or a collection of papers on specific topics maintained by individuals or departments.

The rules for attaching CrossMark logos to documents are simple. First, anything that has a CrossRef DOI assigned can have a CrossMark logo. The publisher of the content is the organization responsible for assigning the CrossMark logo, and that means that CrossMark logos would only appear on content of CrossRef members. Second, CrossMark logos will not be applied to pre-publication content. If a publisher makes preprints available in an “online ahead

of print” program, CrossMark logos could be attached to them. But if an author submitted a manuscript for publication, it could not have a CrossMark logo displayed before acceptance.

The CrossMark pilot is underway now. The plan is to launch the service (in other words allow all CrossRef publishers to participate) in the spring of 2012. Additional information about CrossMark is available from the CrossRef website^{vii} and from previously published articles^{viii,ix}. The service will be completely funded by participating publishers. Institutions and end users will not incur any charges for the display of the logos or clicking on them to get status and record information.

CrossRef would like to invite the library community to contribute suggestions for the best way to communicate the value of CrossMark to researchers. CrossRef is a not-for-profit organization of publishers. Our affiliate members, including about 1700 libraries, are all organizations. Reaching individual end-users directly has not been a major activity until now. CrossRef will work cooperatively with publishers to explain the service, its benefits, and how it works. CrossRef would also like to encourage collaboration between librarians and faculty who teach research skills to students in order to emphasize the importance of identifying the provenance of research that scholars rely on. The goal of the CrossMark service is to maintain the integrity of the scholarly record—a goal that is consistent with the mission of academic librarians and scholars themselves.

ⁱ Hauser, Marc D., Daniel Weiss, and Gary Marcus. “RETRACTED: Rule learning by cotton-top tamarins.” *Cognition* 86, no. 1 (November 2002): B15-B22. [http://dx.doi.org/10.1016/S0010-0277\(02\)00139-7](http://dx.doi.org/10.1016/S0010-0277(02)00139-7).

ⁱⁱ Johnson, C., “Author on leave after Harvard inquiry. Boston Globe, 10 Aug 2010. Available at http://www.boston.com/news/local/massachusetts/articles/2010/08/10/author_on_leave_after_harvard_inquiry/?s_campaign=8315 (accessed 16 Dec 2011).

ⁱⁱⁱ Anderson, H. and Clark, K., “Finding Trusted Information: How Do Faculty, Students and Librarians Look For Information?” Charleston Library Conference, November 2011.

^{iv} Davis, Philip, “The Persistence of Error: A Study of Retracted Articles on the Internet and in Personal Libraries,” unpublished manuscript, results presented at CrossRef Annual Meeting, November 15, 2011. Slide presentation available at <http://www.slideshare.net/CrossRef/retracted-articles-studyphildavis>. (Accessed December 16, 2011.)

^v Hauser, Marc D., Daniel Weiss, and Gary Marcus. “Rule learning by cotton-top tamarins.” *Cognition* 86, no. 1 (November 2002): B15-B22.

<http://www.psych.nyu.edu/gary/marcusArticles/Hauser%20Marcus%20Weiss%202002.pdf> (Accessed December 16, 2011)

vi The pilot participants include Elsevier, International Union for Crystallography, Oxford University Press, The Royal Society, Wiley-Blackwell, and Vilnius Gediminas Technical University.

vii <http://www.crossref.org/crossmark/index.html> (accessed December 30, 2011).

viii Meyer, C, "Distinguishing published scholarly content with CrossMark," *Learned Publishing*, 24:2, April 2011, pp. 87-93, April 2011, <http://dx.doi.org/10.1087/20110202>.

ix Meyer, C, "What's JAV got to do with it? Indicating versions of record with CrossMark," *Against the Grain*, 23:2, p. 18, April 2011, <http://www.against-the-grain.com/2011/07/v23-2-what%E2%80%99s-jav-got-to-do-with-it-indicating-versions-of-record-with-crossmark/>