To Boldly Go Beyond Downloads: How Are Journal Articles Shared and Used?

Carol Tenopir
School of Information Sciences, University of Tennessee, Knoxville, ctenopir@utk.edu

Gabriel Hughes
Elsevier, g.hughes@elsevier.com

Lisa Christian
Center for Information and Communication Studies, University of Tennessee, Knoxville, lchrist2@utk.edu

Suzie Allard
School of Information Sciences, University of Tennessee, Knoxville, sallard@utk.edu

Dave Nicholas
CIBER, Dave.Nicholas@ciber-research.eu

See next page for additional authors

Presenter Information

Carol Tenopir, Gabriel Hughes, Lisa Christian, Suzie Allard, Dave Nicholas, Anthony Watkinson, Hazel Woodward, Peter Shepherd, and Robert Anderson

To Boldly Go Beyond Downloads: How Are Journal Articles Shared and Used?

Carol Tenopir, PI, Chancellor's Professor, School of Information Sciences, University of Tennessee, Knoxville

Gabriel Hughes, Senior Analyst and Program Director, Elsevier

Lisa Christian, Research Associate, Center for Information and Communication Studies, University of Tennessee, Knoxville

Suzie Allard, CO-PI, Professor, School of Information Sciences; Director of Center for Information and Communication Studies, University of Tennessee, Knoxville

David Nicholas, CO-PI, Director of CIBER; Adjunct Professor, School of Information Sciences, University of Tennessee, Knoxville

Anthony Watkinson, Principal Consultant, CIBER

Hazel Woodward, Board of Directors for Project COUNTER; Director of Information Power; and retired Librarian of Cranfield University

Peter Shepherd, Executive Director of Project COUNTER

Robert Anderson, Graduate Research Assistant, University of Tennessee, Knoxville

Abstract

With more scholarly journals being distributed electronically rather than in print form, we know that researchers download many articles. What is less well known is how journal articles are used after they are initially downloaded. To what extent are they saved, uploaded, tweeted, or otherwise shared? How does this reuse increase their total use and value to research and how does it influence library usage figures? University of Tennessee Chancellor's Professor Carol Tenopir, Professor Suzie Allard, and Adjunct Professor David Nicholas are leading a team of international researchers on a the project, "Beyond Downloads," funded by a grant from Elsevier. The project will look at how and why scholarly electronic articles are downloaded, saved, and shared by researchers. Sharing in today's digital environment may include links posted on social media, like Twitter, and in blogs or via e-mail. Having a realistic estimate of this secondary use will help provide a more accurate picture of the total use of scholarly articles.

The speakers will present the objectives of the study, share the approach and avenues of exploration, and report on some preliminary findings. Furthermore, the speakers will discuss how the potential learnings could yield benefits to the library community.

Introduction

Through the efforts of publishers working with Project COUNTER, measurements for downloads of articles have been standardized. The COUNTER reports are now widely used by publishers and by libraries to monitor how many articles from specific journals are downloaded and to compare download amounts across platforms, titles, and time (Shepherd, 2004). Project COUNTER standards have given libraries and publishers a proxy to measure usage, however, tying usage to downloads in order to, in turn, measure value derived through reading misses an important aspect of usage and value of scholarly material. Articles are also shared without downloading by sending links or author's copies following the first instance of downloading (Interviews and Focus Groups Report, 2014). Sharing digital content by email, internal networks, cloud services, or social networks is now widespread (Harley, Acord, Earl-Novell, Lawrence, & King, 2010) (Cheng, Ho, & Lau, 2009) (Interviews and Focus Groups Report, 2014).

This secondary type of usage may be reducing the accuracy of existing value and usage measures, as these existing measures fail to capture secondary

sharing. For instance, widespread sharing may lead to a decrease in repeat downloads or, if only links are shared, an increase in downloads. The extent of this problem is unknown and both downloading and sharing varies by stakeholder group (faculty, postgraduate students, undergraduate students, and nonacademic researchers). What are download counts missing? What is the true or complete value of articles and their access and how might that be measured?

Beyond Downloads: Background to the Research

Gabriel Hughes

Institutions, academics, and publishers all have a common interest in better understanding secondary usage via sharing and related behaviors. Elsevier in particular has been investing in technology which facilitates secondary usage, notably their 2013 acquisition of the Mendeley reference management and collaboration platform. Elsevier is therefore keen to support independent public debate and high quality research about the measurement problems presented by these developments in usage behavior, and is sponsoring this research project conducted by the University of Tennessee.

The measurement problem is that usage data is taken as a proxy measure for actual readership behavior by researchers and yet this relationship is influenced by changes in article sharing behavior. It seems reasonable to assume that downloading of full text articles correlates with subsequent reading of those articles by the person who has downloaded them. It also seems reasonable to believe that sharing behavior would amplify this relationship, as any downloaded article which is shared can lead to more reading events, for example, reading by colleagues of researchers who have downloaded articles and shared them. We strongly suspect that this amplification effect due to sharing is changing over time, and that the assumed proxy relationship between usage measured as downloads and actual readership is unstable and breaking down.

The reason for hypothesizing that sharing behavior is growing is simply the wider context of enormous worldwide growth and transformation in network, social, and cloud technologies, in all areas of human activity. It is now easier and much more common for anyone to share digital content with anyone else via web, mobile, and network platforms. There is already evidence that the research community is subject to these same trends.

In early 2014 Elsevier conducted its own private survey of researchers in which they found that 65% agreed with the statement that they "access or share articles from a shared folder or platform" an increase of 6% against the same response to a survey in early 2013 (the 2014 base was a representative sample of 611 academic researchers, and in 2013 a sample of 1,468). High quality public research is needed to better validate and explore sharing behavior and explore what this means for usage data and the analysis of download and reading behavior.

Beyond Downloads: The Study

Carol Tenopir

There are many methods to share articles. For the purposes of our study, we have divided these methods into two basic categories: formal methods and informal methods. Formal methods are platforms specifically designed to share academic material (such as journal articles) and citations within the researcher's existing research activity and community. Some of the formal methods identified so far include Blackboard, CiteULike, EndNote, Dropbox, Google Docs, Zotero, ResearchGate, Academia.edu, Mendeley, Wizfolio, and RefWorks and newer and better systems appear all the time. These methods fit into a scholar's worklife, while making the process of saving, citing, and sharing easier. They are a positive evolution and aspect of scholarship.

Informal methods, on the other hand, are those methods not specifically developed for article management, citation, and sharing. As we discovered in interviews and focus groups, they are widely used for these (and other) purposes. Some of the most frequent informal methods

include: Twitter, blogs, email, Facebook, and LinkedIn.

Methodology

Measuring, or at least estimating, the amount of use and value from both formal and informal methods is not an easy task. To start this process, a two-prong approach has been used. The first prong included interviews in the United Kingdom and focus groups in the United States. Focus group and interview analysis helped inform the development of the second prong of our study: an international survey. This survey is currently deployed and as of December 3, 2014, we had 985 responses.

The aim of the international survey is to gain further insights into the sharing of a scholar's own work and the works by others. We also want to explore if there is a way to estimate or calculate the average amounts of sharing by various methods and to look at differences by discipline. Our survey questions examine: download and postdownload behavior, different methods of sharing, perceptions of sharing behavior in regards to technology, embargo periods, and copyright, as well as differences based on disciplines, years in academia/research, level of education, and geographic location.

Results from Interviews and Focus Groups

Since the interviews and focus groups are completed, this presentation will mostly focus on results from the first prong of the study. Interviewees and focus group participants ranged from senior researchers and academics to doctoral students. We had 15 interviewees and 14 participants in four focus groups for a total of 29 participants. Twelve participants were from the sciences and 17 participants from the social sciences. Most participants (20 of 29) held a PhD in their field and the average years of experience of working in their field was 11.07 years. In terms of age, most participants were in their thirties (10), while only two participants were in their twenties. Six participants were in their forties, five in their fifties, and six at were more than 59 years old. The mean year of birth was 1968.

Types and Methods of Sharing

Two main types of sharing emerged among participants. The first is sharing just a citation or link to an article. This may actually be the most common way to share. Fifteen participants reported sharing a citation or a link to an article. This presents an interesting situation since if the citation or link goes back to the publisher's site, their download stats will capture it, but the library from which the sending scholar originated will not. The link that is shared may not be to the author's own work; often is to another's work.

The second means of sharing is to send the full text, most often as a PDF. This is more likely to be the sender's own work. In the focus groups and interviews, most participants who share their own full texts say that they also upload their work into institutional or subject repositories (although we all know this varies widely by subject discipline.)

Informal sharing through email or an internal network or sharing print email by personal exchange was the most frequently mentioned method of sharing in both the United Kingdom and the United States. A focus group participant labeled this type of sharing as "bootleg" sharing, suggesting their concerns about copyright violations in this type of sharing. Many participants, particularly those from the focus groups, mentioned copyright concerns. Often, they may pass on the citation or link, but more rarely, the article. In light of these concerns, one participant said that he only shared open access journal articles.

Though "bootleg" sharing may not be strictly legal in regards to copyright or licensing agreements, convenience often trumped uneasiness for many participants. A social scientist often working in collaboration with others explains, "If I have got it, I will just share it. It is easier than having them go track it down, even if I have got the citation. The citation is relatively easy to find, but if I have got it, I will either just share it in a Dropbox [folder] or attach a file and send it. It saves them trouble." In fact, Dropbox was the most frequently mentioned tool for complete document sharing with collaborators or within a research team, and

Twitter was the most frequently mentioned social media tool for sharing links or citations.

We asked participants with whom they most commonly shared material. Not surprisingly, colleagues topped the list at 93%. The circle of colleagues is growing, as the number of coauthors grow and big science projects require big teams. Not far behind is sharing with other researchers. Over 80% of participants said they share articles or links with other researchers. Sharing material with students is also a common practice (79%).

Reasons for Sharing

Academics and researchers share material for a variety of reasons. After all, sharing is a natural part of scholarship and most participants maintained that they share material to further scientific and academic discovery. Participants also share what they feel is good or noteworthy—that is, to promote the reading of their own or someone else's work.

One participant in the social sciences explains his reasoning for sharing material, "I think sending articles is one of the ways to maintain a relationship with the people in my professional network. So if I find something I would send it to my PhD, someone in my PhD whom I have studied with or someone I met at a conference and I might say, 'You might find this interesting.' So, I think that is an excellent way to maintain rapport at the same time show that you care about their research and that you want to be part of their research." People will often then share relevant things with you. In other words, as he elaborated, you build a network that then becomes "your passive information seekers." Essentially, sharing material builds goodwill among potential future collaborators.

The second most common motive, again of an altruistic nature, was to fulfill an information need, especially for researchers who lack access. These top reasons, reflect the very nature of scholarship and both formal and informal systems that make it easier to do are widely adopted. One researcher noted, "I'll have people contact me separately, people outside the United States and internationally who don't have access . . . And I

feel very comfortable, and I don't know if I should, but I feel comfortable sharing my work." Another social scientist explained, "Generally, I have some people outside the country when they have difficulty accessing articles, so I can give them some satisfaction and if they cannot find it, they ask me [and] sometimes I send them a pdf."

Nearly all participants viewed having one's own work shared by others positively. One scientist contends, "It is a stamp of how you are measured and evaluated," while another said, "I would be delighted if my work was known by more." After all, as many note, it is good for citations and reputation. It is also good for feedback and critical review. Although one interviewee is unaware if his work has been shared, he wouldn't mind; in fact, he "would be pleased." Another is "comfortable and trusting of colleagues who share his material."

However, many participants expressed reservations about sharing. For example, can I share a pdf legally? Does it violate intellectual property rights or copyright? One social scientist asked, "So you are published in a journal so now do I have a right to put it in my CV, to put a link to it? I have a pdf of it, but can I share that?" But, the general feeling was even if they might be in violation, they do not think of it as purposefully violating copyright. As a social scientist in the United States explained, "You don't even think [that] usually you are breaking copyright laws. You're just thinking, 'Well, I just trying to be academic. I'm trying to promote this or whatever.'"

Some participants made a distinction between formats in discussing their sharing habits. Sharing their *articles* was viewed mostly positively, sharing their *books* was not. Some participants worried about the effect upon royalties and sales when sharing books. A social scientist noted, "I have a book out and I have to stop and think I want royalties too, right so . . . why would I download somebody's book for free when I wouldn't want that done to myself?"

Moving Forward

In conclusion, Beyond Downloads aims to define ways to measure usage beyond downloads, look at the relationship between COUNTER downloads and additional usage, and to develop practical ways to average or estimate total usage. Our goal, therefore, is to provide a more complete view of value. We also wish to initiate discussion across various communities reading these issues. Our overall question that we would like to posit: Is a COUNTER-like measure or calculator possible?

References

Shepherd, P. T. (2004). COUNTER: towards reliable vendor usage statistics. *VINE*, *34*(4), 184–189.

Harley, D., Acord, S. K., Earl-Novell, S., Lawrence, S., & King, C. J. (2010). Assessing the future landscape of scholarly communication: An exploration of faculty values and needs in seven disciplines. *Center for Studies in Higher Education*.

Cheng, M.-Y., Ho, J. S.-Y., & Lau, P. M. (2009). Knowledge Sharing in Academic Institutions: A study of Multimedia University Malaysia. *Electronic Journal of Knowledge Management, 7*(3), 313–324.

Beyond Downloads Research Team. (2014). Interviews/focus groups white paper. Center for Information and Communication Studies. (White paper).