Chapter Six

A Measured Approach

Evaluating Altmetrics as a Library Service

Hui Zhang and Korey Jackson

The emergence of altmetrics has drawn the attention of academic libraries as a new and effective approach for capturing the types of impact often ignored by traditional citation-based metrics. At Oregon State University Libraries, before investing in a full-scale implementation of altmetrics services, we conducted an assessment survey of faculty and other researchers in order to determine whether such products were as valuable to researchers as traditional bibliometrics (including h-index and journal impact factor). Based on the survey results, this chapter seeks to understand best practices for the introduction and implementation of altmetrics services. The results reveal a mixture of both enthusiasm and suspicion toward altmetrics as an impact measure. Ultimately, we find that, while academic libraries are in the best position to act as intermediary providers of altmetrics services, there is much need for refinement of altmetrics evaluation techniques and a more robust set of best practices for institution-wide implementation.

INTRODUCTION

The term altmetrics has already become familiar to many librarians, and as more libraries and journal publishers add altmetrics data to their scholarly content, the term is finding currency among researchers as well. But the actual practice of altmetrics—which we define as the “tracking of multichannel, online use and conversation around a discrete piece of scholarship”—has yet to gain wide acceptance among scholars, in part because scholarly and administrative motivations for adoption have not yet appeared. Librarians, on
the other hand, especially those who have experienced the limitations of traditional bibliometrics, have been more apt to tout the potential benefits of the “alt” movement in impact metrics. These benefits have two primary beneficiaries: (1) scholarly venues outside the traditional academic journal (e.g., data and software repositories, blogs, non-peer-reviewed online publications) and (2) scholars who might be more likely to contribute to these venues (graduate students, researchers working in fields with clear public contribution mandates, software developers, and all scholars broadly invested in widening the focus of “what counts” as viable scholarship) (Hausstein et al. 2014; Sud and Thelwall 2014).

With these benefits in mind, Oregon State University (OSU) Libraries and Press has begun looking to add altmetrics tracking to our institutional repository services (more on this process later). While Chamberlain (2013) and Piwowar and Priem (2013) have reported that scholars are already beginning to include altmetrics data in their CVs, the practice is by no means widespread, and general acceptance of altmetrics impact indicators is by no means assured. Before committing funds to a particular service, we wanted to be sure that such a service would actually benefit our faculty members and other local researchers and that these stakeholders had a clear investment in the overall efficacy and value of altmetrics.

The purpose of this study is not only to assess whether publicizing altmetrics data would benefit campus researchers but also to better understand researcher awareness of altmetrics and overall attitudes toward this relatively new model of assessing impact; if scholars expressed high degrees of skepticism about (or ignorance of) altmetrics services, then the net gain from introducing yet more data into local scholarly assessment practices would be doubtful. While the study is locally motivated, we are equally interested in how this emergent form of impact data has taken hold among scholars more broadly and what it represents as a trend to researchers who find themselves increasingly subject to data-driven evaluation—evaluation that results in decisions as crucial as who is hired into a new position, who receives tenure or promotion, and who is terminated from a position. To summarize, our research questions are:

1. What is the overall perception of altmetrics services and data among researchers (research associates, junior and senior faculty)?
2. How do rank (faculty and other) and departmental affiliation affect support (or lack thereof) for alternative metrics?
3. What benefits would result from OSU Libraries subscribing to specific altmetrics services for its institutional repository content?

What follows is a brief history and explanation of altmetrics practices and services and an overview of contemporary use cases for institutional adopt-
tion of various services. We then move on to a discussion of our survey methodology, a presentation of our findings, and finally a discussion of those findings and how they have affected our decision to support and implement altmetrics at OSU.

[6.9] ALTMETRICS OVERVIEW

One of the major reasons for altmetrics’ perceived lack of trustworthiness is the basic fact that it is so new to the bibliometrics scene. The term *altmetrics* was coined in 2010 by Jason Priem, a graduate student at the University of North Carolina and cofounder of Impactstory, an altmetrics service provider. Early on, the term became a token of semiradical dissent, due in part to its popularization in “Altmetrics: A Manifesto” (Priem et al. 2010), where the authors pushed for an expansion of the definitions of both *scholarship* and *metrics*. Ultimately, the manifesto concludes that the new networked world has forced a dramatic change in scholarly communication—seen in the profusion of alternate sites for publication (blogs, social networks, disciplinary and institutional repositories) and types of scholarly output (blog posts, datasets, other online dialogs)—and that this change demands a commensurate expansion of the tools used to assess scholarly merit.

On the surface, the argument is uncontroversial. It seems only fair to expand metric tools as the realm of scholarly production expands. But altmetrics has met with a mixed reception in both the scholarly and bibliometrics communities. For conservative elements in academic and publishing spheres, the alt in altmetrics tends to imply a worrisome either/or relationship with more traditional citation metrics rather than a more measured both/and relationship. And for true-blue radicals in the academy, altmetrics looks uncomfortably like an extension of the same neoliberal fixation on “outcomes assessment” that gave rise to administrators’ overreliance on numbers to measure knowledge in the first place (Kansa 2014). The picture is not all hand-wringing and anxiety, however. As Haustein et al. (2014) have shown, librarians are starting to take notice: Of those they surveyed (all members of the 2012 Science and Technology Indicators Conference), a full “72% valued download counts, while a third saw potential in tracking articles’ influence in blogs, Wikipedia, reference managers, and social media” (1145). Numbers like these represent strong positive growth in the rate of adoption of various altmetrics indicators.

To better understand altmetrics as a practice, we can turn to Tananbaum (2013), who offers a pithy definition in “Article-Level Metrics: A SPARC Primer.” He writes that altmetrics (and specifically article-level metrics, which represent one subset of the genre) “open the door to measures of both the *immediacy* and the *socialization* of an article” (4, emphasis in original).
Unlike such instruments as h-index, which require long periods of citation history to be effective and accurate, altmetrics services tend to track scholarship’s immediacy—its initial entry into the scholarly conversation and its uptake in popular news, social media, online citation managers, and the like. This kind of circulation data also accounts for the idea of the socialization of scholarship: how it is shared not only by scholars but also by pundits, policy makers, and the public at large.

Like most categories of “process,” altmetrics is not any one general operation. It is much more accurate to define it as a diverse set of citation-tracking activities performed by a diverse set of actors. At the time of writing, the field of altmetrics services is dominated by three major players: Plum Analytics, Altmetric, and Impactstory. There are, of course, other services that make up the larger practice of alternative metrics tracking, particularly article-level metrics (ALM)—as the name suggests, ALMs represent similar kinds of usage data allied with a particular publisher’s corpus of articles. ALM services are now being offered by the likes of the Public Library of Science (PLOS), Elsevier’s citation database Scopus, Nature Publishing Group (publisher of the journal Nature, among others), and BioMed Central. But the three previously mentioned organizations represent the core of the market for non-publisher-specific services aggregating data on the use and discussion of various types of scholarship.

**Services**

It might be useful to explore these three altmetrics services briefly, as each has a slightly different mission and set of services to offer. There are of course macrolevel similarities: All provide tracking for different categories of user interaction with online scholarship (download counts, mentions in social media, references in online citation managers, etc.), and all offer some kind of visual shorthand for representing the numbers behind these various metrics.

For instance, Plum Analytics (which since 2014 has been part of EBSCO Information Services) has billed itself as a comprehensive metrics tracker, focusing not just on alternative sites for citation but also on typical in-journal citations. Plum Analytics’ PlumX Metrics service, which is promoted specifically for institutional repositories (IRs), cuts a wide swath through the crowded terrain of publication types, locations, and categories of use. In order to make sense of the types of interactions being tracked, Plum Analytics breaks them down into five categories:

1. **Usage:** including numbers of clicks, downloads, and views of an article within various public and institutional repositories (DSpace, EB-
SCO, ePrints, PLOS, figshare, etc.), as well as the number of libraries holding a particular publication.

2. **Captures**: including bookmarks and favorites from sources like Delicious, YouTube, and Slideshare, number of forks in GitHub, and number of saves by readers in scholarly networks like Mendeley and ResearchGate.

3. **Mentions**: including number of blog posts discussing an article, number of comments about an article in Facebook, and number of Wikipedia links to an article.

4. **Social media**: including number of times an article has been like or shared on Facebook and Google Plus and number of tweets mentioning an article.

5. **Citations**: PlumX currently tracks citations from a number of sources, including CrossRef, PubMed Central, Scopus, and Social Science Research Network (SSRN), among others.

Whereas Plum Analytics is primarily geared and marketed to institutions—providing large-scale synopses of departmental and university-wide research impact—Impactstory is focused squarely on the individual scholar. Indeed, their tagline, “Your CV, but better,” is a clear indicator of the audience they envision. Users begin building their scholarship profile by creating an account linked to author identifiers (such as ORCID) and can then upload scholarly works to that profile page. The resulting page organizes material into types, such as article, dataset, figure, and software. Impactstory also tags each uploaded scholarly contribution with a series of badges according to subgroup, including “Cited” and “Highly Cited” (based on Scopus data) or “Saved” and “Highly Saved” (based on Mendeley and Delicious data).

Having gotten its start as the product of a hackathon in 2011, Impactstory is also more forthright about seeking to change the academic incentives and rewards culture. As their “About” page states, one express aim of the Impactstory tool is to “build a new scholarly reward system that values and encourages web-native scholarship.” Though categorized less explicitly than on Plum, a similar set of research use and citation data is also available through the service. Impactstory also expands on the types of scholarship represented, featuring information about articles, datasets, figures, posters, slide decks, and software products. Finally, Impactstory differs from both Plum Analytics and Altmetric in that it is a 501(c)(3) nonprofit corporation, a status that contributes to its overall commitment to openness in terms of both its own transparency and its promotion of open science, open access, and open source materials.

Founded in 2011, Altmetric is probably best known for its signature “donut” scoring mechanism. The Altmetric donut represents a numeric score compiling several different inputs, including Twitter, Facebook, and other
social media sources; blogs; citation managers like Mendeley and CiteULike; major media outlets; and other sources like Wikipedia, F1000, and YouTube. Unlike Plum and Impactstory, Altmetric is attempting to cover all potential consumer bases, offering different service packages for institutions, researchers, and publishers: Almetric for Institutions is an application that displays impact figures for all research articles associated with a particular institution; Altmetric Explorer is their baseline system for tracking various mentions and uses of individual articles; and Altmetric API allows for embedding of an Altmetric score on any website associated with the particular article being tracked.

Games

One point worth mentioning relative to altmetrics’ overall trustworthiness is the potential for unscrupulous researchers to game altmetrics scores. Many social media sites and citation managers like Mendeley and ResearchGate allow users to upload and share academic publications. All of the previously mentioned altmetrics services extract such data as number of downloads, views, and registered readers from these sites—data that ultimately becomes part of the final altmetrics score. Because such indicators are harvested automatically from the web, most of them are vulnerable to gaming (i.e., activities used by authors for self-promotion or boosting popularity) and spam (i.e., usage statistics being polluted by a software agent or web bot).

Previous studies reveal how various social impact indicators have been manipulated. Thelwall (2012) argues that authors and even journal editors could inflate the usage statistics of individual articles by repeatedly downloading them with the help of either human or computer agents. Additionally, although captures, such as bookmarks of an article made by users in Mendeley, are a relatively reliable impact indicator used by altmetrics-scoring mechanisms, gaming these article bookmarking features with fake profiles or simply by asking colleagues and other users to bookmark articles is yet another low-tech method of gaming final scores. Another common gaming approach is a kind of link farming in which all occurrences of various types of article metadata (e.g., title, DOI) are extracted to dummy web pages and used to inflate the final tally. There is also the broader issue of counting such mentions as a significant research metric in the first place. The majority of article mentions appearing on research journal websites are found in tables of contents or other similar descriptive lists and thus have no fundamental citation impact (Kousha and Thelwall 2007; Vaughan and Shaw 2003). Finally, within the sphere of social media, altmetrics is especially prone to being gamed due in large part to the kinds of preexisting spam that already infect platforms like Twitter and Facebook. Mass following, creation of multiple
accounts or bot accounts, and repeated posting and duplicate updates are all examples of practices that produce misleading altmetrics.

[6.28] Fortunately, the National Information Standards Organization (NISO) is currently at work on strategies to detect altmetrics gaming and spamming techniques. NISO (2014) recommends “[m]aking all altmetrics data openly available via a standardized API and/or download, a centralized altmetrics data clearinghouse, and audits for altmetrics data” as a set of operations that would help to reduce the vulnerability of altmetrics data to gaming (9). While NISO’s efforts are still young and obviously rely on wholesale adoption by major players within the marketplace, there is reason to expect that the near future of altmetrics assessment will be a much more accurate one.

[6.29] Case Studies

[6.30] Rather than creating their own altmetrics harvesters for scholarship collections, most academic libraries and many publishers subscribe to services like the ones discussed earlier. We offer the following case studies as models of the ways in which both higher education institutions and publishers have implemented such third-party services.

[6.31] Along with the Smithsonian Institute, the University of Pittsburgh Library System was one of two pilot institutions for Plum Analytics’ harvester toolset PlumX (Howard 2013). Having successfully completed the pilot, the university now offers altmetrics tracking for all scholarly works deposited in their institutional repository, D-Scholarship, which houses a variety of material types, including documents, software, data, and images. At the bottom of each D-Scholarship publication landing page, a viewer can see associated altmetrics (e.g., number of downloads and Twitter mentions) embedded under descriptive metadata. At the time of writing, the University of Pittsburgh’s PlumX account covers a total of 74,874 artifacts in the IR, predominantly government documents (approximately 37,000), articles and other papers (approximately 20,000), and theses and dissertations (approximately 6,000) (University of Pittsburgh 2015). In an interview conducted in early 2013, Timothy Deliyanides, director of the Office of Scholarly Communication and Publishing and head of Information Technology for the University of Pittsburgh, emphasized that the purpose of providing altmetrics is to offer a complete view of the impact of all intellectual outputs, including “gray literatures,” such as presentation slides and software (Enis 2013). In this sense, PlumX was not implemented with any express purpose of changing how fundamental tenure and review processes work at the university but rather to provide a more granular snapshot of the impact of faculty and departmental scholarship in its many forms and in the many channels where it has taken hold. As Deliyanides puts it, “We’re not really on a crusade to change any of the university’s normal processes for tenure or review. . . . But
we hope people will think of new ways to use this data. We do feel it’s valid
data and something that hasn’t been gathered or reported before” (Howard
2013).

In a similar move, Wiley Journals, a major publisher of scholarly journal
content, began in May 2013 to pilot the use of Altmetric services for a
number of its subscription and open access journals (Warne 2014). By Wi-
ley’s own account, the results of the six-month pilot were positive—enough
so that they decided in 2014 to deploy Altmetric tracking across all of its
1,500 journals. As a publisher, Wiley’s primary motivations for adding alt-
metrics were less about measuring research impact and more about increas-
ing readership and author submission rates. For this reason, much like our
current study, Wiley was concerned with user attitudes about altmetrics im-
plementation. As Warne (2014) summarizes,

A major objective of our pilot was to assess reader and author views of altmet-
rics. . . . During the pilot we ran a poll of website visitors. 65% felt the metrics
were useful with a further 23% indicating that they were somewhat useful.
77% of readers responding to the survey agreed or strongly agreed that altmet-
rics enhanced the value of the journal article. 50% agreed, or strongly agreed
that they were more likely to submit a paper to a journal that supports altmet-
rics.

On the landing page of an article in Wiley’s Online Library, users can click
the “donut” thumbnail badge, which brings them to a dedicated metric page
with detailed data and an overall Altmetric score for that article. While the
Altmetric score is just one indicator of the quantity of attention the publica-
tion has received, it is an aggregate of several different sources. The “Score”
page offers several modes for understanding the score with more granularity,
including comparisons with all Wiley articles, with that particular journal’s
articles, and with articles of a similar age. The page also shows the geograph-
ic origins of the various input sources.

Starting in 2014, Oregon State University Libraries (OSUL) began adding
an embedded Altmetric badge to the web page of each open access journal
article deposited into ScholarsArchive@OSU, the university’s IR. The im-
plementation at OSUL first parses the article identifier (either as a DOI or a
DSpace handle) from a metadata record and extracts metrics using the Alt-
metric API. Only when the article has garnered enough attention to receive
an Altmetric score larger than zero will it trigger the display of the Altmetric
badge.

In 2014, OSUL had a pilot Altmetric Explorer test covering all scholarly
works deposited in ScholarsArchive@OSU—more than 50,000 repository
items in total. Altmetric Explorer targets the needs of the individual author to
measure and monitor the social impact of his or her works. The institutional
edition of Explorer, Altmetric for Institutions, allows librarians or adminis-
tutors to group authors by academic (i.e., departmental) affiliation in order to demonstrate the overall research impact of that group. During the pilot, library staff were able to use the tool to identify the top publications (those with the highest degree of web attention) at the university. The IR librarian notified corresponding authors in order to motivate further self-deposit of article content into the IR; to date, faculty self-deposit has increased following the introduction of altmetrics at OSUL. While this kind of outcome is certainly positive, there are still too many variables to trace direct causation. With this in mind, the next phase—which could take shape as a full subscription to Altmetric for Institutions or another similar service—will require further evidence to make a fully informed decision.

**Survey Methodology**

A total of 304 researchers were identified based on having authored publications indexed in Web of Science; this list was previously generated by library staff in an effort to promote open access and encourage authors to deposit their publications in the university’s IR. The researchers on the list come from various academic rankings— Instructor, assistant professor, associate professor, full professor, and emeritus faculty—and all have authored at least one journal article since 2013. The rank and normalized OSU school/college affiliation (i.e., all acronyms or abbreviations resolved to that school/college’s full name) of each person were obtained from the university database. The survey was conducted using the online survey software Qualtrics and issued by staff at OSU’s Survey Research Center (OSU-SRC). The solicitation of participants continued for five weeks.

The survey questionnaire was designed to be concise and free of field-specific jargon. For instance, the word *altmetrics* was replaced by *web usage*, a term defined by such examples as number of downloads, Twitter mentions, and inclusion in citation managers like Mendeley. The survey begins with six straightforward multiple-choice questions asking for participants’ familiarity with bibliometrics (journal impact factor and h-index in particular) and social web tools like Twitter, Facebook, and reference managers. Next, participants answer five multiple-choice questions designed to ascertain their perceptions of altmetrics and its relative importance to their fields of study. The questions cover various aspects of altmetrics, such as which types of scholarship (journal articles, software, datasets, etc.) are most effectively measured by altmetrics and which originating institutions for altmetrics data (institutional repository/library, publisher, funding agency, etc.) are most trustworthy. The questionnaire concludes with two open-ended questions asking for participants’ general comments, concerns, and opinions about altmetrics.

All the harvested survey responses were stored in Qualtrics and accessible to the two PIs and OSU-SRC staff. The answers to the multiple-choice ques-
tions were analyzed using cross-tabulation analysis, a popular method that is effective for categorical data. The columns and rows of a cross-tabulation table represent two different variables (such as “faculty rank” and answers to “How important is web usage as a measure of an article impact?” respectively), and the cells report the frequency counts and percentages corresponding to both variables. For instance, eight assistant professors consider web usage as a “somewhat important” measure, which constitutes 36.4 percent of all qualified respondents. The power of cross-tabulation analysis is its capacity to provide granular insight into the relationship between such variables as perception of altmetrics, faculty rank, and affiliation.

Results

Out of the 304 faculty members who were contacted to participate in the survey, 69 completed the survey, for a response rate of 22.7 percent. Faculty rank and affiliation in the e-mail list were used to group questionnaire responses for purposes of comparison. By affiliation, the greatest number of completed responses were received from the College of Agricultural Sciences (n = 16). An inadequate number of responses were received from Ocean and Atmospheric Sciences (n = 2); the College of Earth, Ocean, and Atmospheric Sciences (n = 1); the College of Education (n = 1); the College of Pharmacy (n = 3); and the College of Veterinary Science (n = 1). Therefore, these affiliations were not considered in the analysis. By title, the greatest number of responses were received from assistant professors (n = 22), followed by full professors (n = 17). For clarification, the designation “other researcher” refers to graduate students, postdoctorates, and research associates.

PERCEPTIONS OF BIBLIOMETRICS AND ALTMETRICS

Responses to the following five survey questions were used to answer our first research question concerning overall perception of altmetrics and bibliometrics:

Q1. How likely are you to rely on citation metrics, such as journal impact factor, when deciding which journals to publish in?
Q2. How likely are you to include your h-index or other measures of scholarly productivity in your CV or promotion dossier?
Q6. How important is citation count as a measure of an article’s impact in your field?
Q7. How important is web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) as a measure of an article’s impact in your field?
Q8. In your opinion, does including web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) along with citation count create a more trustworthy measure of an article’s impact?

Fifty percent of responding faculty reported they were “somewhat likely” to depend on journal impact factor in their selection of journals for research submission. Another 28 percent were “very likely” (see figure 6.1). Thirty-two percent were “very likely,” and 19 percent were “somewhat likely” to include their h-index or other measures of scholarly productivity in their CV (see figure 6.2). For purposes of measuring the impact of journal articles, 88 percent of respondents reported that citation count is “very important” or “somewhat important” (see figure 6.3), whereas only 37 percent reported that they consider altmetrics (i.e., web usage) either “very important” or “somewhat important” (see figure 6.4). Forty percent of researchers agreed that combining altmetrics with citation count creates a more trustworthy measure of impact than citation count alone (see figure 6.5). It is also worth noting that almost a quarter of respondents selected “not sure/does not apply,” indicating that many researchers are still uncertain or undecided about altmetrics, even in combination with more traditional bibliometrics.

Effects of Faculty Rank and Affiliation

Answers to the same five questions are separated into both rank and affiliation to provide more granular analysis. We designated the following four categories related to rank: assistant professor (asst), associate professor (asso), full professor (prof), and other researcher (other). School/college affiliations are subdivided into the following: College of Agriculture Science (coll agri sci), College of Engineering (coll engi), College of Liberal Arts (coll lib arts), College of Business (coll business), College of Forestry (coll forestry), College of Public Health and Human Sciences (coll PHHS), and College of Science (coll sci).

When asked about the importance of citation count, an overwhelming majority of faculty across all ranks reported that it was either a “very important” or “somewhat important” (results combined as “important” in figure 6.6) measure of article impact in their respective fields. The same trend is also observed across colleges (see figure 6.7).

The results also suggest that faculty rank has little impact on perceptions of altmetrics (see figure 6.8). There are only slight variations across rank, with a majority of faculty in all four ranks reporting that altmetrics are either “not too important” or “not at all important” (again combined as “not important” for the sake of analysis) for measuring article impact. Likewise, respondents from across the seven colleges reported that altmetrics were “not important” to their fields, with the exception of the College of Liberal Arts,
Hui Zhang and Korey Jackson

Figure 6.1. Q1. How likely are you to rely on citation metrics, such as journal impact factor, when deciding which journals to publish in?

where 57.1 percent (four out of seven respondents) considered altmetrics “important” (see figure 6.9).

Looking at the question of whether combining altmetrics with citation count would make for an overall more trustworthy impact measure, there are observable differences in answers by faculty rank. For instance, 47.1 percent of full professors responded that offering both types of metrics was a good idea, whereas only 36.4 percent of the other researchers agreed (see figure 6.10). Overall, however, this roughly eleven-point differential among the four ranks is not significant enough to make a strong assertion that academic ranking has an impact on perceptions of altmetrics. We can conclude, though, that there is slightly more acceptance of altmetrics’ importance among senior faculty when compared to colleagues in the ranks of graduate student, postdoctorate, or research associate.

The results suggest that academic affiliation has a stronger effect compared to professional rank: Scholars at three colleges favor the idea of combining altmetrics and bibliometrics, scholars at two colleges are against the idea, and scholars at the final two colleges are undecided (see figure 6.11). It is hard to conclude, however, that college affiliation (and, by extrapolation, research field) has a significant impact because its effects are relatively inconsistent. For instance, faculty from the College of Business and the College of Public Health and Human Sciences strongly disagree on Q8 (In your opinion, does including web usage [e.g., number of downloads, Twitter men-
Chapter 6

Figure 6.2. Q2. How likely are you to include your h-index or other measures of scholarly productivity in your CV or promotion dossier?

...tions, inclusion in citation managers like Mendeley] along with citation count create a more trustworthy measure of an article’s impact?); 80 percent of the former selected “No,” while 71.4 percent of the latter selected “Yes.” Despite this discrepancy, they are of similar opinions regarding the other questions (even Q6 and Q7, which address the isolated importance of bibliometrics and altmetrics).

Altmetrics Services at OSU

The survey takers were asked the following three questions as a means of understanding how the library might provide altmetrics as a service:

Q9. Which of the following types of scholarship are effectively measured by web usage data? Select all that apply: journal articles, software, datasets, slides and posters, books and book chapters, white papers and tutorials, and other.

Q10. Please indicate how much you trust each of the following sources as providers of web usage data: institutional repository (library), publisher, funding agency, research database, and other.

Q11. What would motivate you to include web usage data in your CV or dossier? Select all that apply: funding agency encourages use of this information, department encourages use of this information, university
Hui Zhang and Korey Jackson

Figure 6.3. Q6. How important is citation count as a measure of an article’s impact in your field?

promotion and tenure guidelines encourage use of this information, and peers engage in use of this information.

For Q9, 60.9 percent of the respondents selected “journal articles,” with “white papers and tutorials” and “books and books chapters” as the second- and third-most selected types of scholarship (30.4 percent and 29 percent selected, respectively). It is not surprising to see “software,” “datasets,” and “slides and posters” at the bottom of the content types that are effectively measured by altmetrics tracking (see figure 6.12), as these types of scholarship are likely still unfamiliar to some research fields.

Institutional repositories (library) were selected as the most trusted source of altmetrics, with the highest percentage of respondents (46.4 percent) saying they trust them a great deal. Next were publishers, with 31.9 percent of respondents reporting that they would trust this venue a great deal. Research databases and funding agencies were regarded as the least reliable providers of altmetrics, garnering “trust a great deal” responses 24.6 percent and 21.7 percent, respectively (see figure 6.13).

Faculty report that the largest motivator for inclusion of altmetrics in their CVs or dossiers would result from encouragement within university promotion and tenure guidelines (68.1 percent). Motivation from peers was the second-most influential factor, with a 56.5 percent selection rate (see figure 6.14).
Q7. How important is web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) as a measure of an article’s impact in your field?

**DISCUSSION**

**Limits of Study**

Graduate students and junior researchers, such as postdocs and research associates, are likely underrepresented in the survey due to how the survey participants were recruited. Among the 304 researchers at OSU who received the survey invitation, 202 were faculty members (assistant professors, associate professors, or full professors), 64 were graduate students or postdocs, and 25 were categorized as assistant or associate instructors (it is worth noting that the response rate for this last group was null). This recruitment approach produced a high response rate of 22.7 percent (69 out of 304 candidates), with the tradeoff being that the results predominantly reflect the perceptions of faculty rather than other types of researchers (graduate students, postdocs, etc.). It would be valuable to conduct a similar survey focusing on graduate students and compare their perceptions of altmetrics with results from faculty. It would be similarly valuable to explore responses from higher-level administrators (department heads, deans, and provosts), especially because these positions carry a great deal of weight when it comes to...
Q8. In your opinion, does including web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) along with citation count create a more trustworthy measure of an article’s impact?

Longevity and Trust

Most of the faculty members who participated in the survey considered traditional citation count the most important and trusted measure of impact, despite the well-documented weaknesses of this method (Roemer and Borchardt 2015; Wilsdon et al. 2015; Wouters et al. 2015). The limitations of
citation count include the lack of context (e.g., when, where, and why an article is cited), lack of capacity to credit newer types of scholarship (e.g., software, datasets), and lack of capacity for tracking the many ways a research product can be used outside of citation (e.g., sharing or hyperlinking in a blog post, mentions in news articles or on Wikipedia).

Given these limitations, we might ask why citation count is still considered the most reliable indicator of research impact. Based on our survey results, it appears that time is the key factor behind the dominance of traditional bibliometrics in scholarly communication. Figure 6.15 shows a correlation between the longevity and the perceived trustworthiness of a given metric strategy.

In this graph, the y axis represents respondents’ reported levels of trust in certain metric types. The x axis indicates how long a given metric has been actively available (in years). For instance, the graph demonstrates that the “citation” metric has been in existence for approximately fifty-one years and maintains a corresponding trustworthiness of 88 percent. The longevity of each metric has been determined using the following information:

- **Citations (citation count):** *Science Citation Index* became commercially available in 1964, fifty-one years prior to the study.
- **JIF (journal impact factor):** *Journal Citation Reports* began publication in 1975, forty years prior to the study.
H-index: Jorge Hirsch first published his paper “An Index to Quantify an Individual’s Scientific Research Output” in 2005, ten years prior to the study.

Altmetrics: “Altmetrics: A Manifesto” was published online in 2010, five years prior to the study.

The data do suggest a trend whereby the longer a bibliometrics service has been in existence, the more it is generally trusted by scholars. While this is obviously only a very general pattern and does not take into account variables like commercial viability and marketing, it does suggest that, if a new metric instrument can maintain active use and continue to demonstrate some kind of market value, then it stands a fair chance of being more readily adopted—and becoming increasingly trusted—within the academic community.

Faculty Concerns with Altmetrics

The survey contained two open-ended questions allowing respondents to address questions and concerns they had about the use of altmetrics impact measures (again, referred to as “web usage” in the survey). Overall, we were surprised by both the number of responses and the depth of the input these questions generated. Answers spanned nearly eighty comments and generat-
Figure 6.8. Q7. How important is web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) as a measure of an article’s impact in your field? (by faculty rank)

Self-Promotion

In general, researchers were highly suspect of the utility of social media when it comes to scholarly impact. Many suspected that scholars, publishers, and other institutions would take advantage of social media outlets like Twitter and Facebook, falsely inflating both their altmetrics scores and the corresponding appearance of impact. One respondent questioned “how Twitter mentions can be a measure of impact,” stating that there are “too many reasons why an article could have a Twitter mention, many of which are not related to scholarly impact.”

Researchers were also anxious that altmetrics would force them to self-promote. Because social media sites are important sources for altmetrics data, the implication was that a researcher should be active on several platforms (Twitter, Mendeley, ResearchGate, etc.) to demonstrate impact. There was concern that such activity would impinge on actual research time and would not, in the end, do much to reflect the quality and usefulness of the scholarship.
Gaming and Spamming

Likewise, there was a good deal of concern over the potential for altmetrics numbers to be gamed or otherwise falsified. One respondent brought up the particular issue of bot downloads used to inflate download scores; this same researcher opined that “web usage data are highly suspect, no matter what the source, as long as bots are not excluded from downloading.” In another comment, the respondent worried that “[l]arger groups and those with bigger public relations budgets will begin to dominate the market,” their scholarship moving to the top of the list simply because it has been “promoted through various means.”

Favor to Certain Fields over Others

There was also some apprehension about certain fields garnering more attention than others because of mainstream media popularity or general funding rates. One of the responses provided the following detailed scenario:

The size of a given research community is often directly related to the amount of funding available in that area. In this regard, people who work in areas of “politicized science” are often deemed more relevant, and given a higher profile, simply because they’re working on a more “popular” topic that catches the public’s attention and pushes a particular political agenda. If people who
Q. In your opinion, does including web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) along with citation count create a more trustworthy measure of an article’s impact? (by rank)

work in these areas receive more citations and downloads than someone in a less competitive field, what does this metric actually measure? . . . To measure all scientists, from both hard and soft disciplines on the same scale of productivity is ridiculous . . . Some structural biologists may spend years preparing a single manuscript, while a computational scientist may publish 12 papers in that time.

This last concern is especially acute when taking into account differences in the level of national funding and mainstream media attention between applied sciences and various forms of humanistic studies.

Lack of Meaningful Peer Review

The overall lack of peer review of online scholarship (e.g., presentations, blogs, and articles) was also a concern in the responses. As one respondent put it, the “lack of critical review of the majority of web-based materials prior to publication” could lead to the unfortunate scenario in which a “totally bogus article from any source” becomes accepted scholarship as a result of a high altmetrics score. The follow-on to this issue was general misgiving about altmetrics becoming a means of promoting and rewarding lower-quality research. Another respondent expressed “concern that papers that are
Q8. In your opinion, does including web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) along with citation count create a more trustworthy measure of an article’s impact? (by affiliation)

published in lower impact journals, because they are simply not as strong or rigorous, could go on to be perceived to be more important or ‘trustworthy’ due to high social media exposure.”

Providing Altmetrics as a Library Service

We were impressed with the level of knowledge many respondents had about altmetrics and the general operations of altmetrics services. The concerns expressed here are both legitimate and, in most cases, highly informed. That said, despite these concerns, there is still hope—and a good deal of it—for the future of altmetrics as a library service. Indeed, anxieties about altmetrics were counterbalanced by a fair amount of enthusiasm for new methods of measuring scholarly impact. Several respondents were hopeful about the fuller picture of impact offered by altmetrics indicators. One representative comment offered the opinion that altmetrics are a “good addition to more ‘traditional’ metrics,” while another respondent was “supportive of alternative web-based metrics to supplement traditional citation metrics.”
Q9. Which of the following types of scholarship are effectively measured by web usage data?

In addition to this kind of faith in the potential for altmetrics coming from the faculty and researcher ranks, there is also the fact that major national organizations and indexers have begun to support and refine the development of altmetrics. EBSCO, which acquired Plum Analytics in January 2014, has begun including usage statistics for articles and books across their various databases (EBSCO Information Services and EBSCO Discovery Service). EBSCO (2014) describes the benefits of the relationship: “The article-level data from these databases will allow Plum Analytics’ product PlumX to provide usage statistics on articles and books from tens of thousands of providers. This collaboration marks the first time the wealth of information about the actual usage per article such as abstract views, downloads, etc. can be measured across publishers.”

In a similar move, at the end of 2013, Elsevier began displaying Altmetric badges for the top three rated articles from thirty-three of their titles (Huggett and Taylor 2014). This kind of support, while obviously motivated by commercial interest, speaks volumes about publisher investment in these types of services and, as a corollary, their perception that altmetrics stand to become a useful value-add to authors, readers, and scholarly communication consumers in general.
Based on the results of the survey, analysis of the faculty comments, and the general trend in the larger scholarly communication environment, we recommend implementation of altmetrics services at OSU Libraries. While we do not single out a specific provider, in all likelihood such implementation will take the form of subscription to a tool like Impactstory, Altmetric Explorer, or Plum Analytics’ PlumX harvester (each of which maintains a back-end harvesting engine of various altmetrics data and a front-end display of an aggregate score based on this data). Acknowledging, however, that any implementation should be approached carefully, we offer the following best practices for the service:

- Display both altmetrics data and traditional citation indicators.
- Provide context for understanding the numbers behind a particular altmetrics score. Offer thoroughgoing answers to the “what and why” of altmetrics, including an introduction to what actually comprises altmetrics scores—where data are collected, how numbers are calculated, and how to interpret the results—and the overall rationale for implementation.
- Create a transparent interface for feedback about the service. We want to make sure that any researcher who has questions or concerns about the
display of altmetrics data or methods for harvesting this data has access to a librarian who can answer such questions.

- Customize the suite of altmetrics indicators based on feedback from the local scholarly community.

In general, these practices point to a need for proper introduction to and interpretation of altmetrics as a service from a neutral third party; right now, the academic library is best placed to be such a mediator. For one, libraries represent a natural intermediary, already standing as they do between users and researchers on the one hand and service providers on the other. Libraries also have the potential to be a neutral but invested partner in creating a customized product that better serves both its users and scholarly communication writ large. Academic librarians are also in the best position to fill the information gap—which is one source of the trust gap revealed in the responses to our open-ended questions—by providing detailed information about altmetrics without any necessary agenda regarding their eventual use (in spaces like promotion and tenure review). Our survey results support this proposition, with 46 percent of responding scholars seeing the university library as the most trustworthy source for altmetrics services.

Thinking for a moment about why the library is trusted to serve in such an intermediary capacity, we might point out the fact that libraries have long served as both a space for information expertise and a zone free from any
overt political or market-based agenda. Librarians may seek to change certain aspects of information dissemination, but ultimately such change is motivated by a desire to improve the accuracy, discoverability, and accessibility of information. To “improve” in this case means not only to widen the circle of discovery but also to create an information landscape that is culturally more inclusive. In the context of altmetrics, such inclusivity is about expanding both the genres of scholarship and the modes by which it is measured and evaluated.

CONCLUSION

In many ways, what we describe here is a desire to implement altmetrics “ahead of the curve”—to future-proof evaluation metrics in a dynamic scholarly environment by adopting and adapting tools that better serve scholars, their readers, and ultimately their careers. While faculty and other researchers involved in our study were not unilaterally in support of altmetrics, the fact that they expressed enthusiasm for new supplements to traditional citation count suggests the need to experiment with alternatives.

That said, librarians should not be satisfied with simply acting as passive consumers of vended products. As a neutral partner to the scholars we serve, we are well placed to voice concerns about everything from the accuracy of the source to ambiguity in the final score and vulnerability to nefarious actors. We suggest that more librarians be involved in informing policy.
around best practices, standards, and contextualization of altmetrics service products. While several members of the academic library community have served on standards committees sponsored by the likes of NISO, further involvement in this process can only help to strengthen altmetrics products and, finally, the ecosystem in which they can best succeed.

APPENDIX A: SURVEY QUESTIONNAIRE

Q1. How likely are you to rely on citation metrics, such as journal impact factor, when deciding which journals to publish in?

- Very likely (1)
- Somewhat likely (2)
- Not too likely (3)
- Not at all likely (4)
- Not sure/does not apply (5)

Q2. How likely are you to include your h-index or other measures of scholarly productivity in your CV or promotion dossier?

- Very likely (1)
- Somewhat likely (2)
- Not too likely (3)
- Not at all likely (4)
- Not sure/does not apply (5)

Q3. Do you maintain at least one active social media account (such as Facebook, Twitter, Instagram, etc.)?

- Yes (1)
- No (2)

Q4. Do you maintain at least one active reference manager account (such as ResearchGate, Mendeley, etc.)?

- Yes (1)
- No (2)

Q5. How important is open access to you and your scholarship?

- Very important (1)
- Somewhat important (2)
Hui Zhang and Korey Jackson

• Not too important (3) [6.127]
• Not at all important (4) [6.128]
• Not sure/does not apply (5) [6.129]

Q6. How important is citation count as a measure of an article’s impact in your field? [6.130]

• Very important (1) [6.131]
• Somewhat important (2) [6.132]
• Not too important (3) [6.133]
• Not at all important (4) [6.134]
• Not sure/does not apply (5) [6.135]

Q7. How important is web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) as a measure of an article’s impact in your field? [6.136]

• Very important (1) [6.137]
• Somewhat important (2) [6.138]
• Not too important (3) [6.139]
• Not at all important (4) [6.140]
• Not sure/does not apply (5) [6.141]

Q8. In your opinion, does including web usage (e.g., number of downloads, Twitter mentions, inclusion in citation managers like Mendeley) along with citation count create a more trustworthy measure of an article’s impact? [6.142]

• Yes, overall (1) [6.143]
• No, overall (2) [6.144]
• Not sure/does not apply (3) [6.145]

Q9. Which of the following types of scholarship are effectively measured by web usage data? Select all that apply. [6.146]

• Journal articles (1) [6.147]
• Software (2) [6.148]
• Data sets (3) [6.149]
• Slides and posters (4) [6.150]
• Books and book chapters (5) [6.151]
• White papers and tutorials (6) [6.152]
• Other (describe below) (7) [6.153]
Q10. Please indicate how much you trust each of the following sources as providers of web usage data (see figure 6.16).

Q11. What would motivate you to include web usage data in your CV or dossier? Select all that apply.

- Funding agency encourages use of this information (1)
- Department encourages use of this information (2)
- University promotion and tenure guidelines encourage use of this information (3)
- Peers engage in use of this information (4)

Q12. What concerns do you have about web usage data as a trustworthy measure of article impact?

Q13. What else would you like to say about the use of alternative web-based metrics as a supplement to traditional citation metrics?

APPENDIX B: RESPONSES TO OPEN-ENDED QUESTIONS

1. First, I think that web usage data are highly suspect, no matter what the source, as long as bots are not excluded from downloading.
2. I have no idea if the statistics on web usage are reliable. Just because someone visits a web-page does not mean the page content had any “impact.”
3. It could be manipulated too easily.
4. Just because an article or dataset gets a lot of hits does not make it useful. It could be that it’s receiving that attention because it’s ungodly...
Hui Zhang and Korey Jackson

1. Awful. Web usage data does nothing to inform how the article is being used.

2. Just because people download data and publications, it doesn’t mean they actually use it.

3. Might be possible to game the system. Numbers would have to be taken with a grain of salt.

4. Some articles in my field have their impact very slowly—Philosophy is, in general, a slow-moving field—and gain a reputation over time.

5. The size of a given research community is often directly related to the amount of funding available in that area. In this regard, people who work in areas of “politically relevant science” are often deemed more relevant, and given a higher profile, simply because they’re working on a more “popular” topic that catches the public’s attention and pushes a particular political agenda. If people who work in these areas receive more citations and downloads than someone in a less competitive field, what does this metric actually measure? Pragmatism? To measure all scientists, from both hard and soft disciplines on the same scale of productivity is ridiculous. . . . Some structural biologists may spend years preparing a single manuscript, while a computational scientist may publish 12 papers in that time. . . . A difference in the number of other researchers citing this work is indicative of multiple factors, none of which may accurately portray an equitable difference in the caliber or global import of the work being produced. Furthermore, in research . . . the impact of a paper should not always be measured on its immediate impact. . . . Paradigms and regimes of knowledge can change. . . . Good ideas can be missed, buried, forgotten, and then rediscovered.

6. We will increasingly be flooded with information promoting various people’s scholarship. Authors will essentially become promoters of their own individual scholarship. Larger groups and those with bigger public relations budgets will begin to dominate the market. “Good scholarship” will not necessarily emerge to the top but rather the scholarship that gets promoted through various means. I can also envision that there will be large cabals created to partake in activities like this.

7. Web usage data says more about self-promotion than about the scientific merit of a publication.

8. As a supplement I suppose it would be fine, but it would be a very concerning path for me if we were to start to emphasize web-based metrics over citation metrics in the absence of meaningful peer review.

9. I am very supportive of alternative web-based metrics to supplement traditional citation metrics.
13. I like the idea that they are used as a supplement to provide a bit more info. I also like that there is no delay as there is for citations dependent on the review and publication process. I would like to see public scholarship, such as blogs, have some (even if small) consideration in P&T.

14. I think it would be great to have additional measures of impact above and beyond citation indices and h-indices. However, a metric that is just a count of something is not likely to be very useful. A useful metric also has to carry some evaluatory weight. Number of downloads, for example, could become a measure of the popularity of a field but not tell you much about the impact of a particular study. On the other hand, something that tracks citations through time and across disciplines would be a nice complement to CI and h-index.

15. The lack of critical review of the majority of web-based materials prior to publication can fool a web-based metric into implying a totally bogus article from any source is a measure of exceptionally good performance. Hence, web-based metrics are by their nature intrinsically flawed due to a lack of actual thought involved in generating the measure of performance.

REFERENCES


Hui Zhang and Korey Jackson


