



## ■ The 3rd SPARC Japan Seminar 2013

### Redefining the Impact of Research Outputs in the Age of Open Access: Current State of Reuse and Altmetrics

Friday, October 25, 2013: National Institute of Informatics, 12th floor conference room  
(Attendees: 107)

The third SPARC Japan Seminar of 2013, in keeping with this year's Open Access Week theme of "Redefining Impact," focused on the influence of research achievements. In Europe and America, there is intense debate concerning the reuse of data used to produce a scholarly article as efforts advance to mandate open access to research data. At the same time, interest is growing in "altmetrics," which differs from traditional methods of measuring impact by looking at a broad set of social sources, known as the social web, to gauge the reach of various scholarly output that has become openly available. These developments have made it necessary to expand the traditional definition of both research output and its impact. The seminar featured lectures from people on the front lines of each of these areas, followed by a panel discussion on the future of open access. Members of the audience played an active role, making for a lively seminar. We hope the seminar will help to stimulate broader discussion of these issues in Japan. The seminar was attended by 107 persons in all, including publishers, university librarians, and researchers. See the SPARC Japan website (<http://www.nii.ac.jp/sparc/event/2013/20131025.html>) for the handouts and other details. A summary of the seminar is given below.

#### Video letters

The seminar kicked off with messages in the form of video letters from Mark Patterson (eLife) and Peter Binfield (PeerJ), both of whom are front-runners in open access publishing and have participated as speakers at previous SPARC Japan seminars. In the videos, they talked about their recent activities and gave a message for the seminar. Mr. Binfield closed his video letter with the strong exhortation, "Go, Open Access!"

#### Lectures

#### Current Research Data Management Support for Researchers in the UK

Ui Ikeuchi (University of Tsukuba Graduate School)

The sharing and reuse of research data is growing in various fields, driven by the desire for greater research efficiency, the need to verify research results, and trends among the national government, research funding institutions, and others toward mandating the sharing of research data. Expectations are growing for programs by university libraries that assist with research data management to help achieve this data sharing. In the UK, each university library draws up its own policies and plans based on its scale and circumstances, and provides support services with the help of the Digital Curation Centre (DCC). If a Japanese institution is to embark upon its own research data management support programs in the future, it will be important to obtain an accurate

grasp of needs of researchers at each institution and to provide such services with a clear understanding of where support is needed. Some suggestions



toward these ends are to make use of training programs outside Japan, to use standardized metadata, and to take advantage of suitable outside resources.

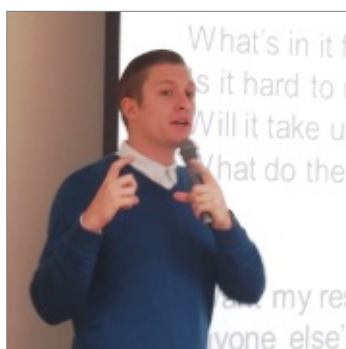
### The "Reuse Factor" and the Future of Credit for Research

Mark Hahnel (figshare)

Why do we need to make all kinds of research outputs fully open, and what impact will this have on future research work? A huge amount of data goes into producing a scholarly article, but only a small portion of that data is released when the paper is published. Making all your own research output open and accessible, including research data, is a way of showing how influential your own research is.

As open access to research output progresses, however, resulting in a swelling volume of information, it becomes necessary to measure the impact of that output and perform filtering. At figshare a DOI is assigned to each bit of content (video, dataset, figure, etc.), which is made citable. Registering content with figshare makes it more discoverable, giving it greater impact. Altmetrics makes it possible to measure that impact.

It has been reported that a gap exists between the



percentage of people wishing to use the research results of others and the percentage of those willing to share their own research results.

To eliminate this

gap, we will need incentives to encourage researchers to register their research output, along with a compelling force such as National Science Foundation (NSF) policy. Moves to make open access mandatory in the case of research data from funded research projects are spreading not only in the US but in Europe as well. Even if funding agencies set policy mandating research data openness, the corresponding support may not always be provided. This is where support by libraries becomes necessary.

We face many different issues, such as not knowing just how much data is output by a research institute; but what we need to focus on is the common desire shared by both researchers and research institutes that their research output will have a bigger impact, along with the question of how best to measure that impact. The number of citations to published papers is still a highly important measure, but by itself it is not enough. What I have proposed is to measure the impact not just of articles but of data, code, and all kinds of other research outputs by their "Reuse Factor." There are many different ways of measuring research output, among which altmetrics is an outstanding Web-native tool, usable not just for articles but for datasets, videos, and a variety of other research outputs. At figshare, we support the major altmetrics services and have also tied up with publishers.

The problem for data is that no citation markup format has been defined, and data not included in the list of references is likely to be missed by impact metrics. This kind of problem is being addressed by relevant agencies worldwide, but efforts to spread awareness among researchers have not been very successful. Here is another area where libraries should assume responsibility.

In the Web era, openness of both research content and research output is advancing. Thanks to this open research and altmetrics, the next generation of research should be carried out more efficiently.

### Altmetrics: The Next Step for Open Access

Jason Priem (ImpactStory)

The importance of open access goes without saying, but it is only the necessary first step on the way to the future of research. In scholarly communication, the Web is still not being used to full advantage. Currently the innovation made possible by the Web extends only to distribution, whereas it is possible to use the Web also for the four steps of data collection, data analysis, storytelling, and conversation.

Data can be published on the Web using a repository such as figshare or Dryad. Publishing it enables the data to be shared, analyzed, and duplicated. It then becomes possible to talk about the process of carrying out the research, a step called storytelling, making use not only of the conventional article and book formats but also videos, blogs, infographics, and many other ways. The Web being a tool well suited to encouraging conversation, collective knowledge can be utilized.

Thanks to the Web, we are no longer limited to the conventional means of publishing information but can share articles and data on our own, quite easily. Comments on something published on the Web can also take the place of traditional peer review. This does not mean that journals are obsolete, as there is still a need for filtering.

And for filtering, some kind of quantitative metric is needed. The conventional metric of the number of citations can make impact visible, but has the problem of reflecting only part of the story. Now that references to research output are moving to the Web, it is

becoming possible to gauge all impact from all aspects. These kinds of impact can be classified from the two



standpoints of audience (who is seeing) and engagement type (in what form is their involvement). This kind of measurement is called altmetrics, which gets its name from “alternative metrics.”

Right now attempts are being made to apply network-based filtering, like that of Google, to scholarly communication. This will have to be done in an open way, however. ImpactStory aims to build an open database covering all scientists, all products, and all impacts.

With the appearance of the Web, a second revolution is coming to scholarly communication. It is not yet clear what direction it will take, but its coming is a certainty.

## Technology Development of Database Integration in Life Science

Hidemasa Bono (Database Center for Life Science)

At the Database Center for Life Science (DBCLS), in collaboration with related institutions, we are integrating databases in the life science field to make them more usable.

Besides providing catalogs of databases (mainly those made in Japan) and database cross-search services,



we are taking over operation of databases that have become unmanageable.

Of the DBCLS initiatives, this presentation introduced (1) development of database integration technology and (2) production of reliable content. Regarding (1), database integration has been proceeding mainly by means of RDF (Research Description Framework). In the life science field, vast amounts of base sequence data are being produced by next-generation DNA sequencers, but due to ethical issues this data cannot always be made public in the case of human studies. Another problem is that metadata granularity is not necessarily standardized. To deal with the huge volume of data, a so-called yellow pages service called DBCLS SRA has been created and quality checking of individual data is performed. Also being provided is a service for mapping data with papers using that data, and a service that provides relevant data for each illness classified in the National Library of Medicine’s MeSH (Medical

Subject Headings).

Ambitious efforts are also being undertaken regarding (2). Newly arriving papers by Japanese authors that have been published in top journals are reviewed by the authors in their native language of Japanese. Contents can be reused under a Creative Commons Attribution (BY) license. Rather long transdisciplinary reviews are made by researchers in each discipline and are assigned a DOI. We also provide “Togo TV” as video tutorials on databases in the life science field.

While database integration in the life sciences is being undertaken by the DBCLS and others, the efforts are still at the stage of telling researchers about their existence and having them try the service.

Meanwhile, large amounts of data are being produced day after day. To encourage the sharing of data, it will be necessary to establish citation methods suited to the data and to take steps to deter improper use. Also necessary will be the provision of tracking functions, and the building up of a stock of success stories toward achieving wider awareness of the advantages of data distribution.

### Implementing an Altmetrics Service into the Okayama University Science Achievement Repository

Hayahiko Oozono (Okayama University Library/DRF)

Recently we introduced an altmetrics service in the Okayama University Scientific Achievement Repository (OUSAR). The reason for doing so is that, while currently the main content of the repository is research papers published in bulletins, we would like to have researchers upload other kinds of



content as well. In order to encourage open access to this end, we would like to promote the advantages of the repository by adding the new metrics of altmetrics.

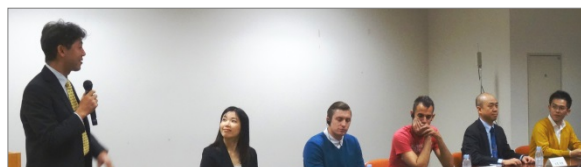
We adopted Almetric.com mainly because it is free, easy to implement, and allows librarians to use

the management tool Altmetric Explorer free of charge. Up to now the repository has been displaying the number of citations using Web of Science and SCOPUS, among others, but now altmetrics has been added. The Altmetric.com badge is displayed on the search results screen and detailed results screen. On screens that can be viewed only by administrators, more detailed information can be displayed by getting the API keys.

After introducing the service, we conducted a comparison of repository download numbers, Altmetric.com score, Mendeley Readership, and Web of Science citations, using journals published in the university as case examples. With this limited number of samples we did not find a correlation between Altmetric.com scores and repository use. We do believe, however, that there is significance to displaying diverse metrics.

In order to expand the repository content, it will be necessary to provide other incentives besides altmetrics to encourage researchers to upload content other than bulletins. Use of identifiers may be one such approach.

### Panel Discussion



Moderator: Kazuhiro Hayashi (National Institute of Science and Technology Policy)

Panel members: Ui Ikeuchi (University of Tsukuba Graduate School) / Mark Hahnel (figshare) / Jason Priem (ImpactStory) / Hidemasa Bono (Database Center for Life Science) / Hayahiko Oozono (Okayama University Library/DRF)

Mr. Hayashi, the moderator, started off by summarizing each of the presentations, and then led the discussion while fielding questions from the floor.

The panelists were asked who should be

responsible for assigning metadata necessary for data sharing, and how formatting should be approached. In response, Mr. Bono noted that even in the life sciences, where sharing is relatively advanced, there is a store of experience that has been given shape, but by no means has all data been standardized. He suggested that these matters should be decided between the people who want to use the data. Ms. Ikeuchi said the ideal would be for researchers to assign metadata, but noted that this would require data to be valued as highly as academic papers. Regarding the valuation of data sharing, an audience member mentioned data journals.

Another issue taken up was how to accurately gauge impact in social media. Mr. Priem noted that there are some impacts that cannot be captured for reasons such as lack of descriptors; but since that bias applies equally across all research outputs, it does not affect the degree of impact. In the long term, advanced data mining should become possible.

Addressing the issue of data and copyrights, Mr. Hahnel indicated that it would be difficult to claim rights to research output resulting from public funding; while Mr. Bono explained that the DBCLS adopts CC BY as the basic license because DBCLS is funded by public grants.

Responding to an audience member who asked about an institutional version of ImpactStory, Mr. Priem said he was hopeful he would be able to meet this need during the coming year. He was also asked about the penetration of altmetrics abroad, to which he explained that currently evidence of the usefulness of altmetrics was mounting. In this connection, Mr. Hayashi made an additional point about the potential of altmetrics, namely, that the

degree of impact of research data shared prior to publication of a paper might serve as a leading indicator of paper citations.

Speaking about the future role of libraries, Mr. Hahnel and Ms. Ikeuchi pointed to the major role in supporting and raising awareness among researchers, while an audience member suggested that the work of librarians was likely to undergo change in line with the needs of the times and of the organization.

One more issue raised from the floor was how to deter misbehavior on the part of researchers such as falsifying data. Mr. Hahnel said that making data public would allow others to reanalyze it, while Mr. Priem said that enabling data to be checked by many eyes and performing pattern analysis of natural numbers should aid in detecting irregularities.

A participant pointed out that various social media will continue to appear in the future, to which Mr. Priem responded by noting that, even if some data sources should disappear, ImpactStory will keep monitoring and collecting data from as many data sources as possible and the interpretation of the data will continue to be left up to recipients.

The audience was actively involved in the discussions from start to finish. The moderator concluded the panel discussion by noting that in the seminars on open access up to now, the discussions centered on how to achieve open access, whereas this one was more about the world after open access and how each of the relevant parties should deal with open access. In that sense, he felt we had come to a new stage, and he hoped each of the participants would put to good use the cutting-edge ideas and discussions.

## -----From attendees-----



### ■ Impressions of the seminar

(People affiliated with university libraries)

“ It was good to be brought up to date on the latest trends in metrics other than impact factor. I also found it quite useful to hear clear explanations of the current state of data

sharing. ”

“ This was very good in that it was the first time for me to learn both about storing data in repositories and about altmetrics. ”

“ Much of what we heard today was new to me; it was very educational. I was especially struck with how today the premise has become open

access, not journals.”

“I was hoping to hear about the current state of open access and altmetrics in Japan, but today the presentation was at the level of learning what altmetrics is.”

“Funding agencies in the UK and US appear to be aware of the value of data reuse, but today’s discussions made me think about how librarians can convey this to researchers who do not even realize that the Web is changing research. This is something I would like to think more about.”

“I get the feeling that Japan’s awareness of data repositories is still behind that of other countries [presumably concerning use of institutional repositories existing in a field]”

(Other/researchers)

“Having doubts about the abnormal use of impact factor, I was interested in learning about other metrics. I was generally satisfied with the seminar. I believe one approach would be to keep the unchanged original and make that clear.”

“This was very instructive regarding the current state of this field in Japan.”

“I thought the lack of data sharing was a problem in the life sciences, but was surprised to learn that the situation is even worse in

other fields.”

## Content/themes/speakers you would like to hear in the future

(Other/other)

“I hope there will be a workshop focusing on altmetrics implementation.”

## Other views and impressions of the seminar program

(Person affiliated with a university library and involved in academic journal editing)

“It will be necessary to put in motion a practice of collecting altmetrics for Japanese-language papers as well, and to educate researchers about it.”

(Person affiliated with a non-university library)

“The content was highly significant, but I am unable to connect it to my own work. Rather than a general discussion, I would like to see a project about actually giving altmetrics a try.”

(Other/other)

“In the cyberphysical field, the issue is obtaining high-reliability data. I have the feeling data repositories will come to play a major role in this regard.”

## -----Afterword-----

😊 Even though I was excited to learn about the future vision of open access, I feel as if I have been handed homework to study what libraries will need to do in order to realize this vision. As someone involving in putting on this seminar, I would like to thank the speakers who gave us their presentations (I’m really delighted we were able to put together this lineup!), Mr. Hayashi who led the project, and Ms. Matsumoto, and the other working group members. A Japanese-language version of altmetrics service was released the other day and I am looking forward to its deployment.

Yui Nishizono (Kagoshima University Library)

😊 My role was to break down the theme I was

given, open access+altmetrics → Open Access Week, and to coordinate with the speakers from overseas. Other than that I was confident in leaving things in the hand of Ms. Nishizono and Ms. Matsumoto. At the Singapore event when I was trying to persuade Jason Priem to make a presentation, I was a bit concerned about how the combination of ImpactStory and the already committed figshare could contribute to Open Access Week. Happily, those concerns turned out to be entirely groundless, as the themes provided by all the speakers meshed well with each other, combining to remind us again of the underlying potential of open access and connecting with the future.

Kazuhiro Hayashi



(National Institute of Science and Technology Policy)

😊 As someone who loves to learn about tools, I was very much looking forward to this seminar. With both figshare and ImpactStory there was a sense of speed, and both gave a highly positive and fresh impression. Just as with Mendeley, tools

that find broad acceptance are not just novel, but are backed by solid analysis and also bring with them a bit of a playful spirit. What will have come of these two next year? Will newcomers emerge? I plan to keep watching the trends with interest!

Risa Matsumoto  
(National Institute of Informatics)

OPEN  International  
ACCESS WEEK

October 21-27  
2013