■ The 1st SPARC Japan Seminar 2015

"How can we evaluate the work on Humanity and Social Sciences?

— From the scholarly communication point of view —"

Wednesday, September 30, 2015: National Institute of Informatics
12th floor conference room (Attendees: 95)

Participants in the 1st SPARC Japan Seminar of 2015 explored such topics as recent trends in humanities and social science research evaluation and Britain's research assessment initiative. They also discussed the role universities and libraries can play now and in the future in research evaluation and support in the social sciences and humanities, including development of infrastructure, in the light of various ongoing initiatives to promote scholarly communication.

A summary of the seminar is given below. See the SPARC Japan website (http://www.nii.ac.jp/sparc/event/2015/20150930.html) for handouts and other details.

Presentations

Fundamental Issue: What Are the Humanities and Social Sciences?

Hisashi Nakao
(Yamaguchi University)

Focusing on the nature of the humanities, the presentation highlights three core ideas: (1) The human sciences are by nature highly diverse. (2) It is wrong to frame the issue of research evaluation as a problem specific to the humanities and social sciences. (3) Who should be responsible for evaluation?

It has been suggested that the humanities have their own unique mission and should be approached as "slow science." But in reality, there is an important role for research papers as well as books, and there is no reason why research in the humanities should be uniformly slow. The humanities should be able to accommodate a variety of approaches, not just one.

In fact, there have always been lines of research that defy the traditional "humanities/social science/natural science" classification. Moreover, the boundaries between those categories are becoming increasingly blurred. How to establish criteria for good research amid this increasing diversity is a challenge confronting academia as a whole, not something to be approached in a compartmentalized manner, field by field.

Whether or not a given type of research can be evaluated quantitatively, researchers should not simply leave the criteria to the evaluators. We need to develop diverse assessment criteria through a process of dialogue, exchanging information as we go.

The Importance of Multifaceted Evaluation in the Social Sciences: Extrapolating from Political Science and Environmental Studies

Ko Nomura
(Nagoya University)

Research results in the field of political science are often more difficult to express numerically than those in the natural sciences. Owing to the academic culture, moreover, there is an emphasis on single-author books and academic society bulletins. Furthermore, digitization of research outputs is still a work in progress. And since papers in the social sciences tend to be long, single-author works and written in Japanese, both research outputs and the frequency of citations tend to be lower than in the natural sciences. Evaluation must take into account disparities in the way research is presented in different fields. If standards are slanted too heavily toward such quantitative criteria as the number of papers published, it could create incentives for doing research that yields quick results, which would be detrimental to the development of the field. Since an important aspect of social science is its impact—that is, its contribution to our understanding of society and our efforts to solve social issues—research also needs to be evaluated in terms of its contribution to the development of the field and its broader social impact. In the area of environmental studies, the existing criteria for university evaluation (ranking, etc.) have been criticized as discouraging activism, and some have called for assessment of research in terms of its social orientation and actual contribution. In the United States, some institutions have adopted Sustainability Tracking, Assessment & Rating

System (STARS) as an alternative assessment index. In the social sciences, research evaluation should be multifaceted, tailored to the character of each discipline, and oriented to concrete social issues.

What Is the Goal of Research Evaluation in the Humanities and Social Sciences?

Kiyonori Nagasaki
(International Institute for Digital Humanities)

Whether gauged by funding or by numbers of researchers, the humanities and social sciences represent only a small fraction of all academic research carried out in Japan. Yet judging from the Science and Technology Basic Plan and other official documents, the government expects quite a lot from the social sciences and humanities. One key issue when it comes to the evaluation of research in the humanities and social sciences is that of society's assessment versus the assessment of academia. The former may not take shape immediately. And the latter will vary depending on the purpose, whether it be review within the field, university personnel decisions, or world university rankings. Among the challenges we face are those of reconciling these divergent criteria with society's demands and of incorporating assessment standards that accommodate new research trends. We will need new and creative approaches to quantitative and qualitative assessment to meet these challenges. For publications carried by such digital libraries as J-Stage and CiNii Articles, providing citation information is one possibility. Interdisciplinary academic societies in the digital humanities field have been addressing these issues from various angles, and we should make the most of the resources they have to offer. American learned societies in the traditional disciplines of literature and history have published guidelines for evaluating digital research outputs, which may also be of use.

Research Evaluation and the Promotion of Responsible Research

Masaki Nakamura
(Osaka University)

This presentation explores the issue of research evaluation from the standpoint of research misconduct and research ethics education. As Japan takes steps to prevent research fraud and misconduct, guidelines published by the Ministry of Education, Culture, Sports, Science and Technology highlight the responsibility of universities and other research institutions, calling on them to adopt such organizational measures as research ethics education and systems for preservation and disclosure of research data. But are these measures sufficient? Unethical research practices include such behavior as multiple submissions of papers and false or misleading authorship practices, as well as specific instances of fabrication, falsification, and plagiarism (FFP). However, after much debate, the US federal government has adopted a definition of research misconduct that is limited to FFP and ignores other serious deviations from accepted standards of research practice. Nonetheless, debate continues on how to deal with questionable research practices other than FFP, and the focus is now shifting to the promotion of "responsible conduct of research" (RCR). Questionable research practices other than FFP are very widespread, although the frequency varies, and prevention needs to be approached via the basic factors that foster RCR: research environment, reward system, and educational process. We also need to consider policies to promote high-value research so as to reduce "research waste." In this context, debate is likely to focus on development of research systems to guarantee high-value research and design of evaluation systems that factor in the reputation of the researchers.

Research Assessment and Its Impact on Humanities and Social Science Research in the UK

Ikuya Sato
(Hitotsubashi University)

In the UK, efforts to assess research performance at a national level began in 1986 with the Research Selectivity Exercise. Subsequently, the Research Assessment Exercise (RAE) was conducted four times between 1992 and 2008, and the Research Excellence Framework (REF) was implemented in 2014. Most UK universities have taken part in the RAE/REF. In 2014, the program reviewed some 200,000 research outputs by about 50,000 researchers. The cost of the review has been extremely high. Since the assessment outcomes impact the allocation of funding, competition among institutions has intensified year by year. Currently it has reached the point where anything less than a perfect rating is considered worthless. The system has also been criticized for further increasing the functional division between teaching universities and research universities. High ratings give institutions an advantage in securing outside funding, which enables them to secure even higher ratings, and so forth.

Supporters of the RAE/REF argue that it (1) creates accountability regarding government expenditure on research, (2) applies meritocratic principles to support high-quality research, and (3) uses competition to encourage efficient research activity. Critics, meanwhile, contend that

institutions have begun "gaming the system" to enhance their ratings. Specifically, they claim that the framework has fueled rampant headhunting of star researchers, a bias toward easily publishable research, and massive output of unimaginative research lacking in novelty or creativity. Some have also complained that it deemphasizes educational and administrative functions of higher education that have no direct impact on ratings.

In the social sciences, the form of publication has been shifting from books to papers, suggesting that researchers may be placing priority on short papers in consideration of the assessments. The lesson that Japan should take away from Britain's experience is the need to (1) clarify the ultimate goals sought through "selectivity and concentration," (2) tailor the evaluation process (means) to the policy objectives (ends), (3) carefully consider the effects and unintended consequences of evaluation, (4) provide disclosure concerning the evaluation process and the allocation of resources based on evaluations, and (5) conduct "evaluations of the evaluation."

Panel Discussion
The Role and Potential of Universities and University Libraries

Moderator: Shoji Komai
          (Nara Institute of Science and Technology)
Panel members: Hiroya Takeuchi (Chiba University) / Hisashi Nakao (Yamaguchi University) / Ko Nomura (Nagoya University) / Kiyonori Nagasaki (International Institute for Digital Humanities) / Masaki Nakamura (Osaka University) / Ikuya Sato (Hitotsubashi University)

In the panel discussion, participants and attendees exchanged a range of views on such issues as the requirements for good evaluation and systems to enable new modes of evaluation. The following is a summary of their discussion.

SATO: It's important to consider the purpose of the evaluation, and for whom it's being carried out. We need to think about how to assess the qualitative aspects of research.

TAKEUCHI: I think a key problem is the base used for evaluation. In the STM disciplines, research competition focuses on the number of citations, and most people recognize that that's a pretty accurate reflection of research evaluation. But in the humanities and social sciences, the number of citations doesn't accurately reflect the academic assessment of research quality. When you try to come up with an indicator that most scholars in the humanities and social sciences would accept, you realize there are basic problems with the scholarly communication ecosystem. In Japan's humanities and social science disciplines, it isn't even a question of open access yet, since many of the scholarly journals haven't been digitized. A more developed scholarly information and communication ecosystem would pave the way for more diverse modes of evaluation.

KOMAI: In Japan, there's a tendency to move ahead with something only when we're certain that it can be implemented perfectly. But why shouldn't we start with what we're capable of doing now? We can also consider new modes of evaluation, such as the Facebook "Like" button or asking people to name papers they like, as in sociometric testing. And we need to think about ways of evaluating broader social impact. We need a platform that accommodates a range of evaluation methods. If anyone has any ideas on that, I'd love to hear them.

FROM THE FLOOR: I think environmental studies can be viewed as a sort of microcosm of scholarship in that they're multidisciplinary, comprising many different facets. I'm wondering how research in environmental studies is evaluated, taking into account the diverse nature of the constituent disciplines.

NOMURA: In environmental studies, we currently leave that to the judgment of experts in the individual disciplines. But there are efforts underway to conform to international guidelines, as by submitting papers to peer-reviewed journals, and a new peer-reviewed journal was even launched with this objective in mind.

FROM THE FLOOR: It seems to me that one can't even begin to discuss evaluation methods without access to the data to be evaluated. I think whoever generates the research should provide access to what they've done in list form, and then leave the evaluating to the evaluators.

FROM THE FLOOR: Facebook's "Like" function seems easy enough to implement from a technical standpoint, and it could probably be embedded in institutions' online repositories. Institutional repositories could also take over the citation tracking function that CiNii used to perform.

FROM THE FLOOR: I think we have to be careful because as soon as evaluators specify their assessment criteria, researchers start tailoring their activities to those criteria. For example, if they adopt the number of papers as a criterion, then researchers will start churning out papers, and if they use books as a criterion, they'll start writing books.

SATO: I'm very concerned that unless we clarify the purpose of evaluation, technology will run away with the whole process. In the field of art, university assessments take the form of exhibitions. Assessment is based on peer evaluation.

KOMAI: We're not going to come up with the answers here today, but it's clear we need to keep thinking about approaches to evaluation, so I hope we can continue holding these kinds of discussions.

--Attendee feedback---

It was a very stimulating seminar because so many of their points hit home. I seem to have a librarian mode and a researcher mode, and during this seminar I was in "researcher mode." Almost all the presenters were researchers, and it was a very good balance of viewpoints. Their analyses were spot-on, and I was very satisfied with the way they kept sight of both the positive and the negative sides of evaluation. It was especially interesting to hear the panelists argue that "there's such a thing as slow science in the natural sciences too, and standard evaluation criteria have helped science progress"

(Nakao); that "assessment begins to go off track as soon as you establish assessment criteria" (Sato, Adachi); that "there have also been problems in the humanities and social sciences, which haven't set clear evaluation criteria" (Takeuchi). The discussion impressed on me the dilemma between accommodating diversity and keeping things in hand. Judged in terms of the number of times the discussion really "clicked," I think this might have been the most stimulating SPARC Japan seminar yet. I'm really glad I attended.

(person affiliated with university library)

---------Afterword--

🙂 With the recent reorganization of Japan's higher education system raising questions about the need for the humanities, the topic raised universal issues of scholarship, in the humanities and sciences alike. Most scholars recognize that the phenomena they study are multidimensional and need to be approached from more than one angle. Surely the same principle applies to research evaluation. It seems to me that scholars need to take responsibility for promoting this sort of understanding and showing people how to probe beneath the surface of things.

Shoji Komai
(Nara Institute of Science and Technology)

🙂 Evaluation of research in the humanities and social sciences is a very important theme but a tricky one, and there were doubts during the planning stages as to whether it would be possible to hold a very meaningful seminar. But I think that the presentations, which represented a range of perspectives, along with the opinions expressed from the floor, have provided us with some

important clues on how to proceed. Developing assessment criteria that will satisfy people in every discipline is a difficult task, but I get the sense that it's important to start with what we are able to do now, tailoring assessment to the purpose, and to keep evaluating our evaluation systems to avoid stultifying rigidity.

Keiko Yokoi
(University Library, the University of Tokyo)

🙂 This year there has been a lot of talk in the media about evaluation of research in the humanities and social sciences, and as it has become an issue in my own department, I had a strong personal interest. It's not an issue that lends itself to easy answers, but I think it's important that we continue debating it in various forums and gradually formulate a new vision for and approach to the humanities and social sciences. I hope that this seminar contributes to that process.

Shinji Mine
(Mie University)