

# The 2nd SPARC Japan Seminar 2017

Preprint and Open Access

## Sharing and the Future of Open Access

Gregg Gordon

(Managing Director, SSRN)

### Abstract



Open access and digital repositories are important pieces of the research dialogue. The recent addition of Black OA to Green and Gold have asked questions that do not have simple answers. Sharing in general has changed significantly over the last twenty years and Gregg Gordon will provide an overview of the SSRN, from its start in the social sciences to its broad expansion across 30+ disciplines, including life and physical sciences. This talk will include real world experiences, examples of the problems that come from broad sharing, and a view into the future of Open Access and scholarly communications in general.



### Gregg Gordon

Prior to helping Michael C. Jensen, found SSRN in 1994, Gregg worked at KPMG and entrepreneurial companies in technology and healthcare. He speaks around the world and writes regularly about scholarly research and the changes needed to create innovative research faster. Most recently, he co-authored *The Question of Data Integrity in Article-Level Metrics*, published by PLOS Biology.

I am going to talk about two basic pieces of Elsevier, open science and SSRN, and how we see the world. I want to highlight one or two points and give you the big picture.

To get started, we will look at Elsevier in the broader context. Elsevier is part of RELX, and RELX includes various companies, but all of them are based on sharing information. RELX is truly a global company with over 66 offices in 24 countries. RELX also has some very well-established brands such as The Lancet and Cell. As publishing has changed, so has Elsevier. It is becoming a data and analytics company beyond just the publishing roots with a strong focus on open science.

### Open Science

Open science means a way of working together to make research more open, more collaborative, or more transparent. On this slide you see some of the pieces of Elsevier that contribute to the different



(Figure 1)

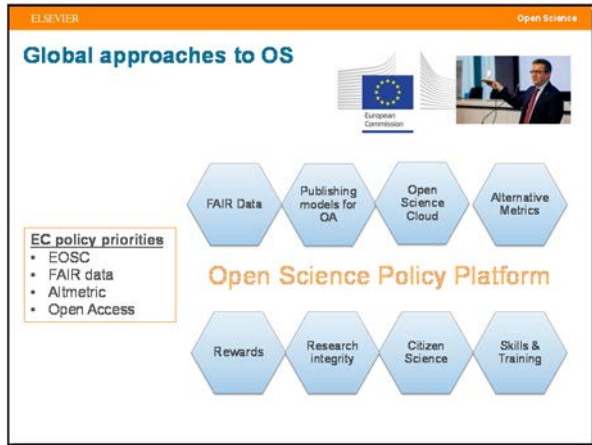
points of open, collaborative, and transparent (Figure 1). There are a lot of activities happening globally that involve two organizations such as SPARC, ORCID, or other entities across the globe. There are also global approaches toward open science in

the UK, Japan, and the United States (Figure 2, Figure 3).

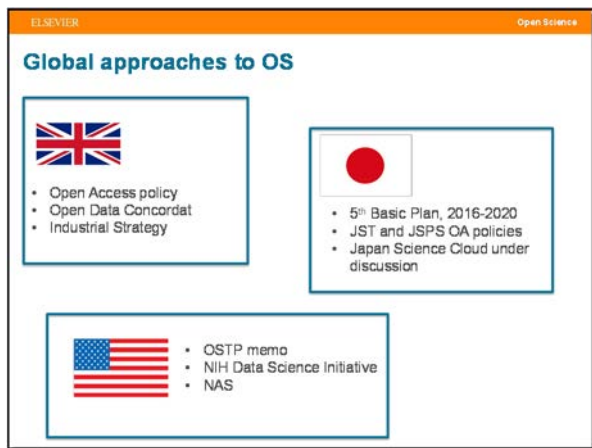
**More Open**

Let us talk for a moment about being more open. The simple definition of open access within Elsevier is free and permanent access to scholarly research for everyone. In the simplest sense, there was gold and green open access. This chart shows the growth of open access and these different types (Figure 4). However, the majority of publications are still subscription based. This map shows the worldwide use of green open access (Figure 5). Many of these countries also have gold open access, but green is available and widespread worldwide.

I chose to take China as an example because it



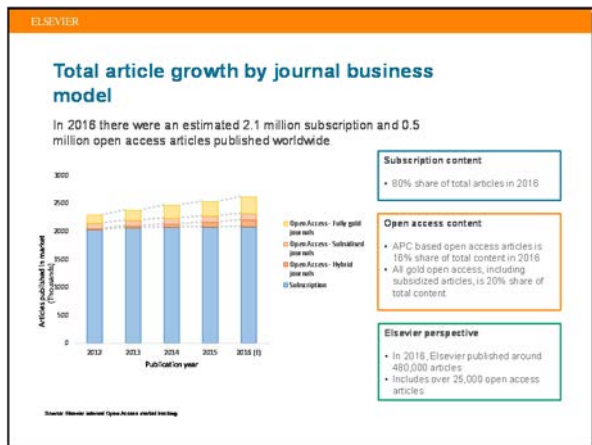
(Figure 2)



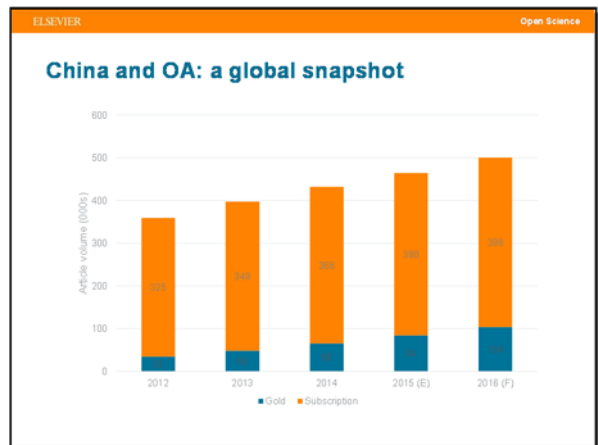
(Figure 3)



(Figure 5)



(Figure 4)



(Figure 6)

has shown greater growth than most other Asian countries (Figure 6). However, growth has still been limited. I view this as an opportunity for all of us to do more sharing. Here are some data regarding Elsevier's open access in both gold and green (Figure 7).

One of the ways that we can share more is to take advantage of the Elsevier application programming interface (API) that allows for both metadata and content to be placed in an institutional repository (Figure 8). Another alternative that Elsevier is working on is CHORUS, where they are currently experimenting with US agencies and piloting services in Japan and Australia to provide content to broader institutional repositories. Another way to share more is to start sharing data.

We often see the data being used by more than one researcher, and the more we share it, the more researchers benefit (Figure 9).

Linking the data to the article is very beneficial, especially if it resides in a specific domain data repository. It is now possible to share data within SSRN, or to publish data articles in many fields. A recent pilot is to start sharing open peer reviews. One of the problems with peer review is acknowledgement and recognition for reviewers. By providing this recognition we believe more reviewers will review (Figure 10).

### More Collaborative

Let us talk about collaboration. I spoke with the Chinese Academy of Sciences and we are seeing them

**Elsevier and open access**

**Gold open access**

- 2nd largest gold open access publisher
- Publish over 170 fully open access journals
- Publish over 1,850 hybrid journals
- Published over 25,000 open access articles in 2016
- Choice of either a commercial (CC BY) or non-commercial (CC-BY-NC-ND) user license.
- Article publishing charges (APCs) range from \$150- \$5000 (US Dollars)

**Green open access**

- Largest publisher enabler of green open access
- All 2500+ journals provide a green open access option
- Participate in CHORUS and support pilots with institutions and international funders
- Free API program to fuel repositories
- Share link service provides 50 days free access to recently published research
- Open archives in 108 journals, including Cell Press titles after 12 months

(Figure 7)

**Storing research data**

- Submitting authors can store research data in Mendeley data repository and link this to their article on ScienceDirect
- Open access, free-to-use repository
- Hosts all formats (raw, processed data, tables, code and software)
- 10GB free storage and CC0 license option
- Assigns data DOI to make it independently citable
- Partnership with DANS for long term storage and preservation

**MENDELEY DATA**

(Figure 9)

**Fueling Institutional repositories: Free API program**

Get metadata & abstracts	Show the final version	Keep users on your IR	Make manuscripts public
Search API can extract the metadata and abstracts of your affiliated authors to help populate your repository, enhance discoverability and create links or embed final version	Entitlements API** can ensure subscribed users see the final version and visitors can see the accepted manuscript hosted on your repository	Article Retrieval API embeds the final article in your repository so subscribers don't leave your site to access the article	Hosting permissions API can retrieve embargo end dates so you can make manuscripts automatically available to the public after embargo

**Scopus** Did you know? Scopus customers can also integrate Scopus APIs to retrieve metadata and abstracts across all publishers

Please Note: \*\*Minimum part of the program [/dev.elsevier.com/tocdoc\\_sd\\_ir\\_integration.html](https://dev.elsevier.com/tocdoc_sd_ir_integration.html)

(Figure 8)

**Open peer review reports**

Pilot in 5 journals to experiment with publishing peer review reports as openly available research outputs with their own DOI.

Editors	Reviewers	Authors
<ul style="list-style-type: none"> <li>• 70% of those surveyed believe that reports are more in depth and constructive i.e. improved and more useful to authors, as a result of being made open</li> <li>• Some concern it may make it harder to find reviewers.</li> </ul>	<ul style="list-style-type: none"> <li>• Like the ability to get recognition</li> <li>• Over 95% said publishing review reports didn't influence their recommendation nor would it influence them accepting further review invites for the journal.</li> </ul>	<p>Generally positive about having open peer review reports next to their article</p> <p>Little impact on their journal choice, indeed slightly more likely to choose a journal offering this.</p>

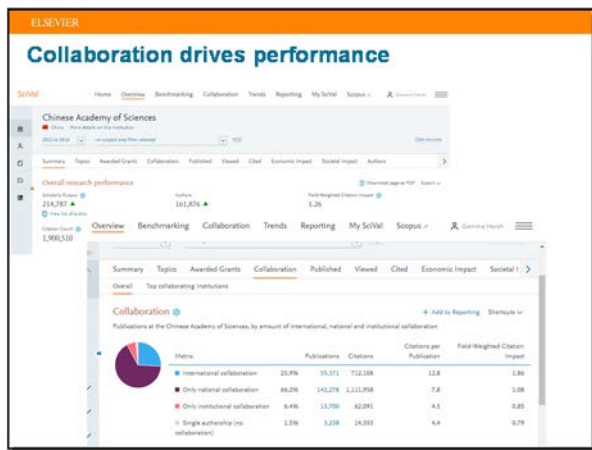
**Forward look:** Extend to other interested journals and provide more data about the peer review process on ScienceDirect

(Figure 10)

collaborating with a number of other international agencies (Figure 11). SSRN is part of this collaboration effort, as you can see by the number of users, papers, and authors in the SSRN system (Figure 12). Mendeley is another example of collaboration with early-stage sharing of references and other materials. Regarding working with others, we have the Atlas Award that recognizes societal impact. AudioSlides provides a video and a five-minute presentation for practitioners or researchers in other fields. Figure 13 shows a recent Elsevier report, Gender in the Global Research Landscape. SSRN this week launched its own Women's and Gender Studies Research Network for sharing early-stage research.

**More Transparent**

Moving on transparency, we have joined a number of other publishers in these data guidelines for making research transparent and open with the community of science. We have also enabled easier searching for databases in Scopus. Scopus is a very important tool, and we think that the ability to use it more broadly is much better. DataSearch is an interesting tool in that it does not just search the metadata, but allows you to go in and search the data itself so that you can determine if you are interested in a dataset before you download it in its entirety. We have also worked with FORCE 11 to enhance citation of data. It is a different type of reference, but we think it is important to acknowledge dataset creators. We also have enhanced reporting of information with negative results so that we can learn from the work of others, as well as specific journals dedicated to the publication of just data. As I mentioned at the beginning, Elsevier is becoming a data and informatics company.



(Figure 11)



(Figure 12)



(Figure 13)

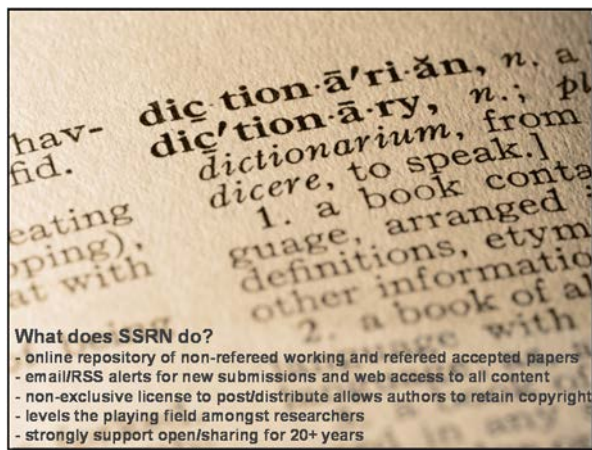
**About SSRN**

Shifting to SSRN, SSRN started about 22 years ago, and here is some basic information about it (Figure 14). I believe the most important of these points is fourth bullet of leveling the playing field amongst researchers. It does not matter whether you are at Harvard University or Monroe Community College in Rochester, New York where I live. It does not matter what country or region you are from, and it should not matter what language you write your research in.

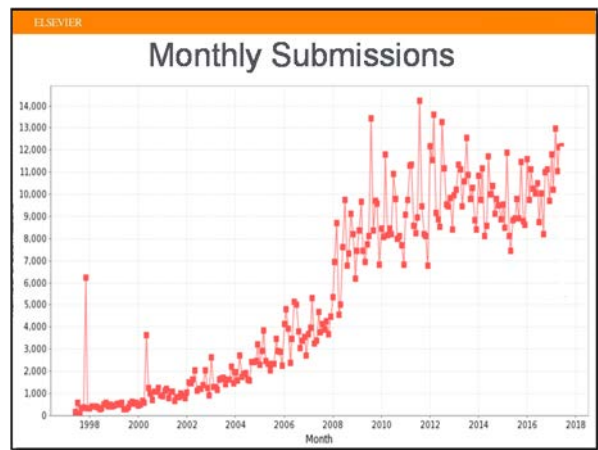
We have about 350,000 authors, and the spikes in this graph indicate where other entities have joined SSRN (Figure 15). This graph shows submissions over time (Figure 16). You see similar spikes and then significant growth in the last sev-

eral years. I think this is an important slide (Figure 17). The blue dots are the submissions by day, and the red dots are revisions by day. In November of 2007 we started to allow papers in the SSRN library to be revised/updated. This chart ends in April, 2012. On this day in April, 2012, we had more revisions than new submissions. This was an important day in SSRN's evolution because it meant that SSRN was a living, breathing body of information. We want research to evolve and get better over time and sharing early helps it happen faster.

Getting access to this early-stage research increases the amount of research that is available and relevant to you. The researcher's idea available to you, as well as the draft, working paper, con-



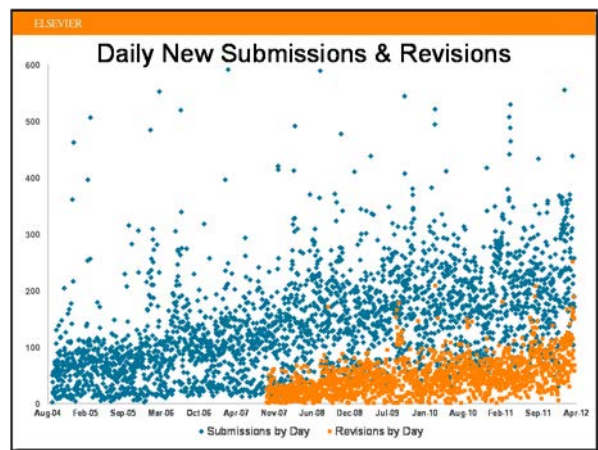
(Figure 14)



(Figure 16)



(Figure 15)

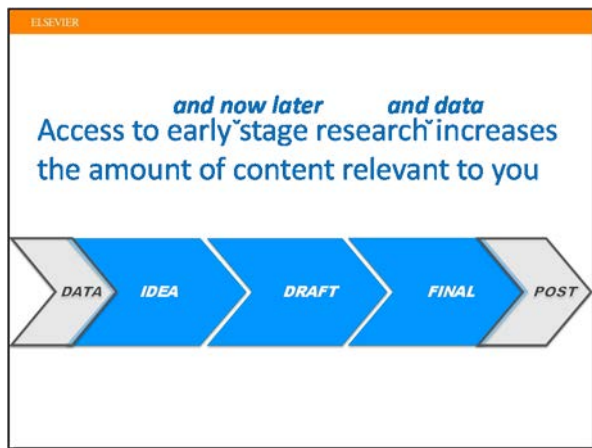


(Figure 17)

ference proceeding, or preprint, as is the final published article. What used to be one version is now three (Figure 18). We have now added post publication versions to update the research. We have also added data, so we have gone from one to now



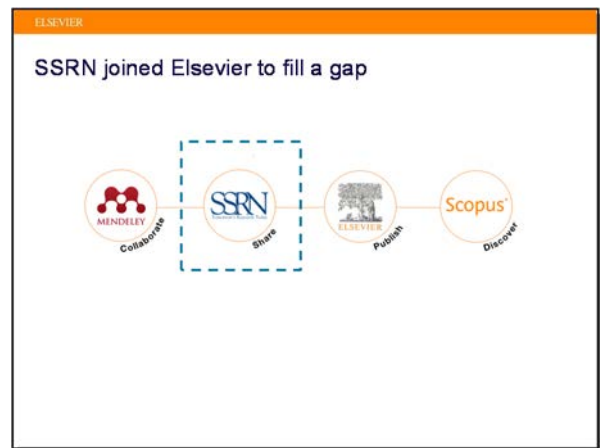
(Figure 18)



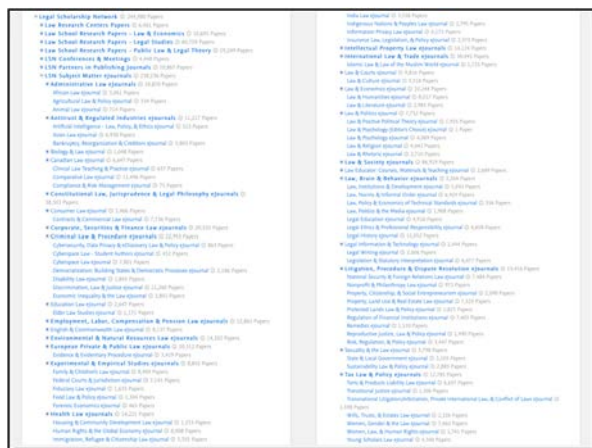
(Figure 19)

possibly five versions of a research paper (Figure 19). Managing these versions is complex. This slide is an example of the complexity in just the legal research areas within SSRN (Figure 20).

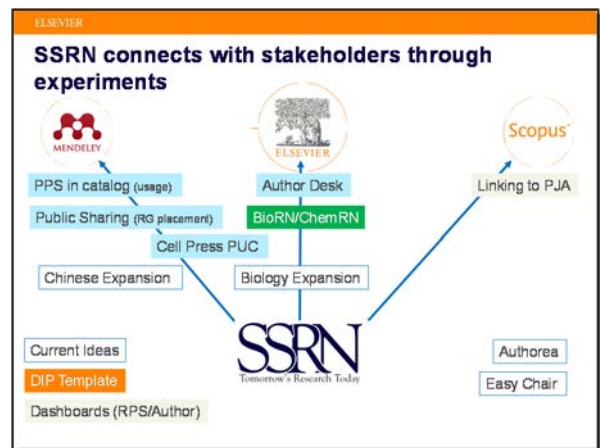
SSRN started in the social sciences, then added humanities, biology, chemistry, and women and gender studies this week. We will add engineering next month. Our slogan, our mantra is 'Tomorrow's Research Today'. We looked at papers published in the Journal of Financial Economics for 2014 and SSRN's eLibrary database had every paper but one three-and-a-half years in advance. When we looked at the world, we realized we could not fulfill our destiny without a partner, so we looked to Elsevier and felt we were filling a void (Figure 21). Now, we have gone from 'Tomorrow's



(Figure 21)



(Figure 20)



(Figure 22)

Research Today' to helping researchers connect, develop their ideas, and share early research.

My simple definition of innovation is to create new things by being exposed to a broader, deeper set of existing things. We want SSRN to be the broadest, deepest set of research things. This is to enhance research performance, or help researchers create cool, new innovative research faster. Whereas we thought we were filling a gap, we are actually connecting the different pieces with experiments and pilots to move research forward (Figure 22). I recently posted on our blog, SSRNblog.com, to celebrate Open Access Week, and I talked about how open access has evolved over the last 15 years. I talked about some of the good and some of the things we need to do better regarding open access. I encourage everyone to read it.

### SSRN Statistics from China, Korea, and Japan

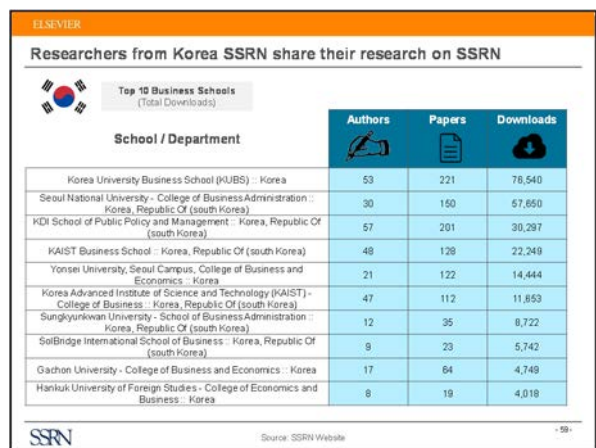
I was in Beijing, Seoul, and now Tokyo over the last week and a half. We have looked at the ways in which different researchers at different schools are using SSRN. These Figures show the statistics for China (Figure 23), Korea (Figure 24), and Japan (Figure 25). We realized that we need to do a better job of letting researchers know they can write, research, and share in their native language. We also need to let them know that SSRN works with institutional repositories, disciplinary repositories, and broader country repositories.

●Floor 1 I am from the National Bioscience Database Center. The number of Japanese users seems significantly fewer than those from China and Korea. Have you analyzed the reasons for this?

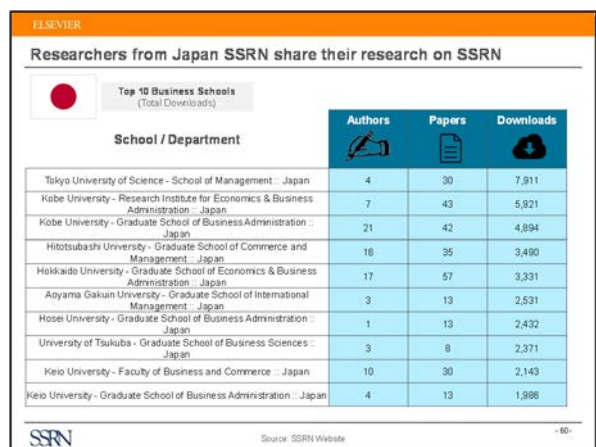
●Gordon We had this conversation at lunch with other speakers. China has more content, but Japan used SSRN earlier. I found research in Japan to be more sophisticated earlier than those other two countries. I am here to try to understand and



(Figure 23)



(Figure 24)



(Figure 25)

learn, but I think we are at a generational shift. Many of the more progressive Japanese researchers used SSRN. Now we are starting to see younger researchers coming up through the ranks using SSRN. I am here to help them be as successful as possible as quickly as possible.

●Floor 2 My name is Fukagai from Yokohama National University. I am also a member of the governing board of SPARC Japan. I think a commercial-based service and a mechanism that was started by academics and researchers such as arXiv.org may not fit well with each other. I believe it is important to watch what will happen in the future with regards to this.

If we are in a generational shift as you mentioned, the percentage of people using SSRN or other tools provided by publishers will certainly increase. This means that the number of young people who can use such convenient tools will increase, but the progress in academia and the increase in young users are two different matters. Although there have been many technological innovations since the Industrial Revolution, simply using something because it is new is not what is important. Rather, the kind of culture that we can create is important. Currently, the meaning of new technologies is questioned from the perspective of our daily lives and their place in society. We must also consider what they mean for the academic world as well.

Who will provide the tools required for open science? Will we do it ourselves? What is commercially provided is indeed convenient, and we are thankful for it. However, I believe that we must consider separately what kind of research is fostered in such an environment and how that

knowledge contributes to society.

●Gordon The question you ask is a very good one. I would make it a little bit smaller. I do not see the distinction as simply commercial versus academic. We would at least have to include governments and funding bodies.

I am friends with Simeon Warner and Paul Ginsparg. Simeon came to Rochester when arXiv was trying to figure out a sustainable business model. We ran through a number of scenarios of what they could charge, how they could charge, or most importantly how they can fund themselves. The commercial question gets even more complicated when you look at something like bioRxiv or the Open Science Framework. bioRxiv now is funded by the Chan Zuckerberg Initiative (CZI), Zuckerberg founded Facebook. The Open Science Framework is funded by the Laura and John Arnold Foundation. Arnold is a billionaire hedge fund manager.

Therefore, one of our choices is government, which has to justify their activities to the voters, members of their parties, or special interests. Another choice includes bioRxiv and OSF where we have billionaire businessmen funding not-for-profits and making decisions through those avenues. The third choice are commercial entities that have to provide value or their customers will leave. I do not think any of them are inherently good or bad. They all have their flaws. I get frustrated when people do not acknowledge that each of them have flaws.

I am excited that SSRN has joined Elsevier, a huge commercial entity with some not so good things in its past. This sharing game has gotten very expensive and realized we could not be suc-



cessful on our own. We were fortunate to partner with a large entity that believed in what we were doing and fully supports us. We would not have biology, chemistry, women and gender studies, or engineering all this year if we did not have the right partner. For me, I trust an entity that has to continue to provide value in a situation where everybody could walk away rather than relying on entities that can change their minds and stop funding. We just need to make sure we have the right controls in place and security for our information. That is what I have tried to do as best as possible within SSRN to make sure that we will be able to continue working for a very long time. But, just like I came here for the weekend to enjoy the good weather of Tokyo, there are no guarantees.

●Floor 3 My name is Hayashi. I am from NISTEP and a member of the governing board of SPARC Japan. In 2013 there was an informal meeting called the Fiesole Retreat where many publishers gathered, and Gregg and I were invited as lecturers. I had the opportunity to talk with him, although despite being in different sessions. I would like to highlight an episode from that meeting and ask some simple questions.

When I met the CEO of Mendeley at the SPARC Japan Seminar in 2011, I asked him, “What will you do from now on? Is there a chance that Elsevier will buy Mendeley?” He apparently became upset. Later, Mendeley was acquired by Elsevier.

I asked Gregg in 2013, “What do you think about the sustainability of the business? Is there a possibility that you will seek commercial capital?” He answered at the time that he would place importance on the SSRN brand, but that it is now under the umbrella of Elsevier.

From a very neutral standpoint, I understand that who patronizes the system is very important, especially since we need to conduct a variety of trials despite the fact that no one can foresee what will happen in the future.

Based on this, I have two questions. First, how will the sustainability of SSRN be maintained by Elsevier? Second, considering your current situation, how do you evaluate the sustainability of the arXiv.org business model? What is your perspective on the membership business model (which could be called a taxation model) where fees are collected from members and information is provided to the public free of charge?

●Gordon First, it is worth saying that Mendeley and SSRN had plenty of options. To be honest, there were better offers financially, but they were not better for SSRN’s evolution. How many people in the room are doing the same job for over 20 years? I see very few hands. When you do something that long you do it because you love it. Despite the challenges, despite the frustrations, you want it to continue. We structured a deal that as best as possible guaranteed SSRN continues as it is.

If you look at it objectively, you will see that Elsevier and all the other publishers are changing. They are not stupid people. They do not sit in smoke-filled rooms making decisions. Everything is happening in the open. Everybody is doing different things, and none of us know what the future will be like in 20 years, 10 years, or maybe even five years. I say to my team and everybody else that I talk to within Elsevier, “This is a crazy time. Either embrace it or do something else. However, the great thing is that there are no wrong answers. We have to Figure it out every day.”

To get back to your original question, I spent a lot of time with societies and other membership organizations. They have difficulty maintaining their memberships. As an accountant by training, I would not pick that business model. However, as a person who travels hundreds of thousands of miles a year to stand in rooms and listen to smart people like everyone here, I know I certainly could be wrong. More importantly, I really hope that they are successful.