

国立情報学研究所 国立大学図書館協会 共催シンポジウム

「大学からの研究成果オープンアクセス化方針を考える」

ーハーバード大学, レディング大学, 北海道大学を事例にー

The Harvard Open-Access Policies

Stuart M. Shieber

(Welch Professor of Computer Science, School of Engineering and Applied Sciences
Director, Officer for Scholarly Communication Harvard University)

Abstract

Systemic problems in the scholarly publishing systems that have arisen over the past decades, especially journal publishing, have led to a decreasing ability to access research results, and pressure on library budgets that have wreaked havoc on all of the missions of the library. I will discuss a set of policies and actions taken at Harvard to address both the symptoms and the causes of the dysfunction in journal publishing, in particular (i) the open access policies enacted by several schools at Harvard intended to promote the broadest access to the university's scholarly writings by retaining rights to distribute scholarly articles according to the principles of open access, and (ii) an open-access "compact" to found a sustainable business model for open-access journals.



Stuart M. Shieber

Stuart Shieber is James O. Welch, Jr. and Virginia B. Welch Professor of Computer Science in the School of Engineering and Applied Sciences at Harvard University. His primary research field is computational linguistics, the study of human languages from the perspective of computer science. His research contributions have covered a broad range of areas of inquiry beyond that field as well, extending to theoretical linguistics, natural-language processing, computer-human interaction, automated graphic design, the philosophy of artificial intelligence, computer privacy and security, and computational biology. He was the founding director of the Center for Research on Computation and Society and is the Director of the Harvard University Office for Scholarly Communication and a faculty co-director of the Berkman Center for Internet and Society.

Professor Shieber received an AB in applied mathematics summa cum laude from Harvard College in 1981 and a PhD in computer science from Stanford University in 1989. He was awarded a Presidential Young Investigator award in 1991, and was named a Presidential Faculty Fellow in 1993, one of only thirty in the country in all areas of science and engineering. He has been awarded two honorary chairs: the John L. Loeb Associate Professorship in Natural Sciences in 1993 and the Harvard College Professorship in 2001. He was named a fellow of the American Association for Artificial Intelligence in 2004, and the Benjamin White Whitney Scholar at the Radcliffe Institute for 2006-07.

Outline

Today I want to talk about the need for open-access policies and about particular policies that we have worked on at Harvard. I will talk about the flaws in the current system for scholarly communication and then what policies can be instituted to begin to cure the flaws and that will take up the remainder of my talk.

I will start by reviewing the goals that we all share to distribute information, the systemic failure of the current system, and then describe two policies that we have spearheaded at Harvard. One is a short-term policy to address the symptom of reduced access and the other is a long-term policy to move towards a solution to the underlying problem.

The Goal

First, I will talk about the goal that we all share. I mentioned the goal that universities are pursuing by engaging in scholarly publishing, and in my own school, which is the Faculty of Arts and Sciences of Harvard University, this goal is expressed in our formal research policy (Figure 1). I have a segment of that policy on the slide and you can see it calls for the widest possible dissemination of the scholarly output of the university.

There is a very strong agreement on this goal, which is the broadest possible access to the research outputs of university, ideally completely open access, if that is possible.

Traditionally, from the days of the printing of issue journals, the dissemination of the scholarly output of university researchers was provided by publishers. These publishers also provided a range of other services; they include logistics of the peer review process, managing the peer review process of production services such as copy-editing,

“First, the [FAS research] policy should encourage the notion that ideas or creative works produced at the University should be used for the greatest possible public benefit. This would normally mean the *widest possible dissemination* and use of such ideas or materials.”

Harvard Faculty of Arts and Sciences “Grey Book”

Office for Scholarly Communication
Harvard University

(Figure 1)

typesetting, and so forth, and distribution services. These services are absolutely central to the scholarly enterprise. They need to be preserved in an economically-sustainable fashion.

Therefore, the scholarly publishing system should be a partner in furthering this universal goal of broadest possible access to scholarly literature, but there is a strong evidence that the scholarly publishing system the way it is set up now is systemically and intrinsically flawed. I will show a couple of examples to make this point.

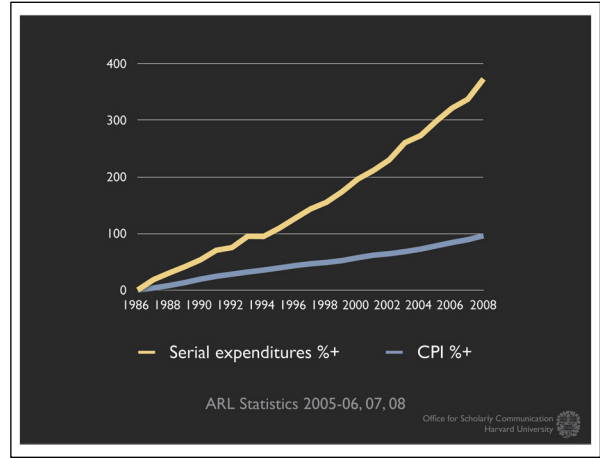
The Problem

I will start with the symptom of direct importance to the goal of broad access. For decades, we have seen a steady, consistent hyperinflation in journal prices. Expenditure for journals has been going up at several times the rate of inflation and the result is what Dr. Ojima referred to in his comments as the so-called serials crisis. The slide shows in blue line relative inflation rate and in yellow line serials expenditures, and you can see this hyperinflation in the divergence between these two lines with serials expenditure going up at many times the rate of inflation for decades now (Figure 2). Whenever we

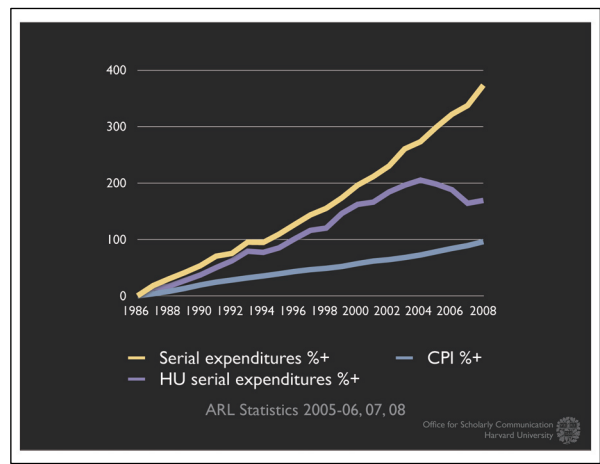
have exponential growth of crisis that cannot go on forever, something has to give in, and this giving in effect in libraries is reduced book collection, serials cancellations, and a reduction in access to articles in journals.

I am privileged to work at a university with substantial resources, but even at Harvard University, we are not immune from this problem. On this slide, I added a line for Harvard's expenditure increase over the same period (Figure 3). You can see that we attempted to keep up with this hyperinflation in serials expenditure over a period of some years, and then around 2004, we gave up. We are not the only ones who have given up. Giving up is inevitable because no budget can keep up with exponential increases, and it is these exponential increases that are the first piece of evidence that there is an intrinsic flaw in the scholarly publishing system that I want to talk about.

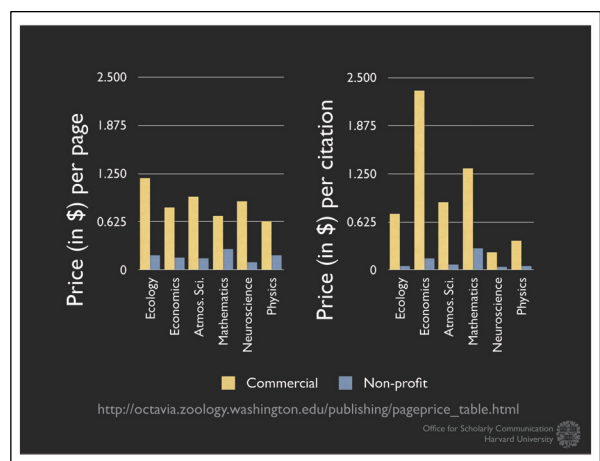
Next, I will turn to the second piece of evidence. You might think of alternative explanations for this hyperinflation other than failure of the scholarly journal market. For instance, may be publishing journals is increasingly expensive. This turns out to be false because if we look at the cost per page of journals published by commercial publishers versus scholarly society or nonprofit publishers, there is a differential factor of 6 between the two—commercial publishers are 6 times more expensive than nonprofit publishers (Figure 4). Perhaps the reason for this difference is because the commercial publishers publish journals that are much better than the nonprofit publishers and therefore are more expensive to run. We can test this theory by looking not at cost per page but cost per citation. If we look at cost per citation, we find



(Figure 2)



(Figure 3)

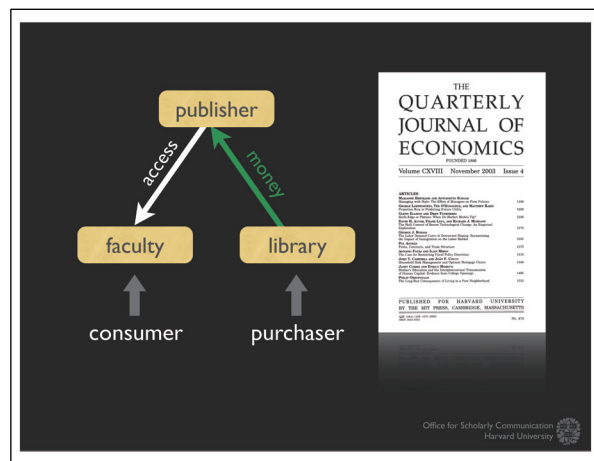


(Figure 4)

commercial publishers are not 6 times more expensive, but 16 times more expensive than nonprofit publishers. What can we conclude from this very large differential between the cost of the commercial publishers and the nonprofit publishers? Andrew Odlyzko in his paper on *The Economics of Electric Journals* says, “The great disparity in costs among journals is a sign of an industry that has not had to worry about efficiency,” that is an industry that is in a dysfunctional market.

To summarize, something is wrong in the scholarly publishing system and this underlying problem has especially bad side effect in that fewer people, researchers and general public alike, can get access to research results. Something needs to be done to restore this access that systemic market failure in scholarly journal publishing has led to. Therefore, I will talk about two things that can be done: one, addressing the short-term problem of access and the other one, addressing the long-term problem of market failure. These two things involve the establishment of certain kinds of policies at universities. A few years ago, they were completely untested, but now a small set of universities have tried these approaches, and we now have good information about how they work. In particular, we can address some of the primary worries that people had originally about these policies.

I will start with a policy to address the short-term problem of access, but before I do that, I want to say a bit about what is the underlying cause of this market failure that we are trying to compensate. There are a couple of factors that caused this market failure. The first is the fact that the product, the good that publishers sell, is access, and



(Figure 5)

of all the services they provide, that is the only one they typically charge for. The ability to sell access is based on monopolistic ownership through copyright law, so the product being sold is monopolistically owned and, therefore, is not subject to price competition.

Second is the phenomenon of what the economists call moral hazard; this is the phenomenon where consumers who are protected from the cost of a good tend to over-consume it. If you think about a normal case for subscribing to a popular magazine, the consumer of the good, that is, the reader is also the purchaser of the goods. The consumer, that is, the purchaser provides money to the publisher and in return the publisher provides access to the consumer. By contrast, for scholarly journals, the purchaser is typically a research library and the purchased goods is access to the article, so the users of that access are faculty, students, and patrons of the library. They receive the access, but they are not the purchasers of the access (Figure 5). This is a recipe for what economists call it a moral hazard, and, therefore, we would predict to see over-consumption, inelasticity of demand, and hyperinflation. So, inelasticity of demand and

hyperinflation are exactly what we are seeing. This is all very depressing.

Having talked about the problem and its root causes, I want to turn now to the policies that may be able to mitigate the problems. I will talk about two such policies.

The Short Term Approach

The first is a short-term strategy that is intended just to address the symptom of the problem that is, decreasing access to scholarly articles. From 2008, at Harvard, we began to establish faculty-based policies to promote open access to scholarly writings by granting license to distribute our articles through an open-access repository. It required faculty to make a copy of the final version of each of their articles available to be distributed.

This policy has three main parts. First, faculty members grant permission to the University to distribute their scholarly articles. Technically, this is a nonexclusive, noncommercial transferable license granted to the University. Because the license is transferable, it can be transferred back to the author so that the authors can distribute their articles as they see fit. That is the first part of the policy.

The second part of the policy is to make sure that the policy itself cannot stand in the way of the best interest of the authors; a waiver of the policy will be issued for any article at the sole discretion of the author. The author retains the decision as to whether to have this license or not.

Therefore, the first part of the policy is permission and the second part is the free waiver. These two parts combined together mean that faculty authors preserve their choice to retain the rights to their articles. The difference now due to this policy is

that the default has changed. Before the policy was enacted, authors did not retain rights unless they expressly opted in by engaging in negotiation with publishers, whereas after the policy was enacted, authors retained the rights by default unless they expressly opted out.

The third part of the policy is that the University can now take advantage of this permission the policy enables to the faculty and the faculty makes their articles available by depositing them into the Harvard Institutional Repository.

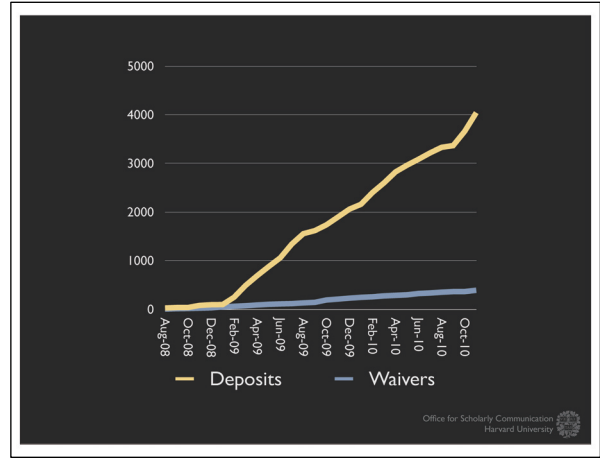
So, all the three parts, permission, waiver, and deposit, combined together make up the policy that my own school, the Faculty of Arts and Sciences, at Harvard voted in February of 2008. Since then, five other schools at Harvard have voted this policy as well as other institutions including the Massachusetts Institute of Technology, Duke University, the Stanford School of Education, and a dozen others in the United States.

Hence, this policy has a number of good effects and I will mention a few of them. First, the policy makes a collective statement of principle that the university supports the broadest access to our scholarly output. Second, it completely clarifies the rights situation for every article because either in a normal case the policy says that the university and the author have rights to distribute the article or there is an explicit waiver and we can track that waiver and know that that article only has whatever rights the publisher has provided. Further, this policy allows the university to facilitate the process of depositing articles into the repository and to negotiate with publishers collectively on behalf of the entire faculty because the university itself becomes a participant in the rights of distribution. Finally, hope was that by

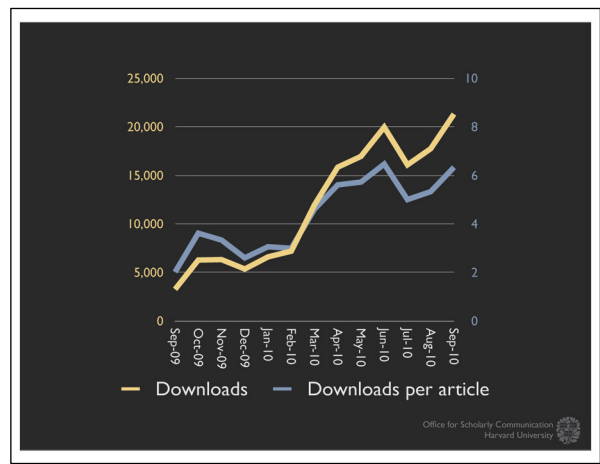
moving from an opt-in system to an opt-out system for retaining rights, we would increase the amount of rights retention and that has indeed turn out to be the case. I should mention that one of the attractive properties of this kind of opt-out policy is that it leverages the natural laziness of our faculty because if they do the least amount of effort, then we retain the most amount of rights. So, this was in February 2008 when faculties started voting these policies.

Now, I want to get a sense of what has happened since then in the following 2.5 years. Here is a graph of some statistics from our institutional repository at Harvard (Figure 6). The repository is called DASH, which stands for Digital Access to Scholarship at Harvard. The yellow line shows deposits of articles into the DASH Repository. We launched DASH internally within the Harvard campus in the middle of 2008, and in September 2009, we opened up the repository to the rest of the world. That was the external launch of the repository. You may wonder what happened in January 2009. That was the time when we started employing students or open-access fellows to work with the faculty in order to provide their articles to the repository as required by the policy.

So, we have seen very steady deposits into the repository. A very large percentage of the members of the Faculty of Arts and Sciences now have at least one article in the repository, and there are over 4,000 articles now. On this slide, this blue line is the cumulative number of waivers of the policy over time, and the notable point about the number of waivers is that it is low. In fact, we have had very few waivers of the policy. It is hard to know exactly, but it seems roughly 5% of the articles receive waivers. At Massachusetts Institute of Technology



(Figure 6)



(Figure 7)

(MIT), they have more accurate numbers for the waiver rate. The rate seems to be about 1.5%. So, it is probably even lower than ours.

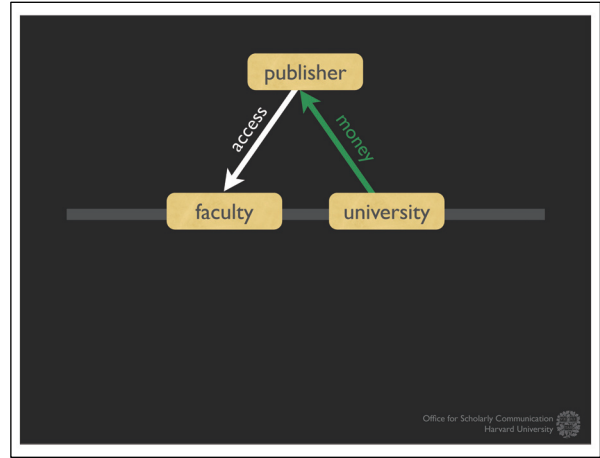
We have also seen extraordinary usage of the collection in repository that has been enabled by this policy (Figure 7). We are seeing tens of thousands of downloads per month and that number is increasing over time, which is represented by the yellow line. The average downloads per article is also quite strong and is increasing over time, which is represented by the blue line.

Now, because of the policy, we are able to retain

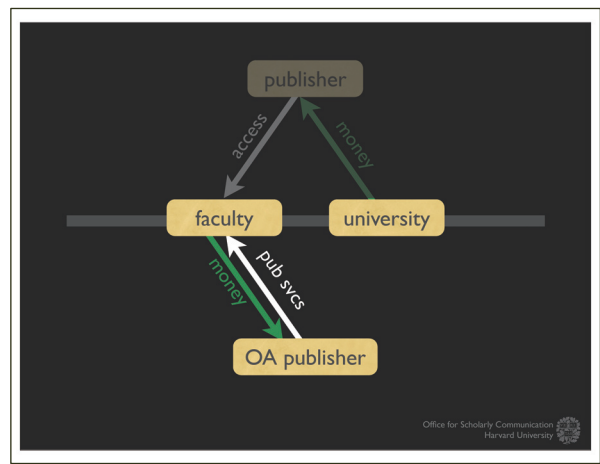
rights for distribution of vast majority of our faculty's articles and we are obtaining copies of those articles to freely distribute to anyone who wants to access them, but let me remind you what this policy does not do and was not intended to do. This policy does not serve as a replacement for journals or journal publication. We still need journals and we still need services that publishers provide through the process of publishing journals. It also does not threaten the viability of journals. The economic viability of journals does not really get affected one way or another. It is not a replacement for journals, but a supplement to the access that journals provide. Finally, it does not address this underlying market failure that has led to the problems in the traditional subscription-based scholarly publishing area.

The Long Term Approach

Hence, this short-term approach is really a measure to improve accessibility in the short term for our communities' writings. But in the long term, we would like an approach that would provide an alternative sustainable business model for journal publishing, one that is not subject to market failures of traditional subscription-based system. What would an alternate business model look like? Here, I am showing again the traditional model where universities and their libraries pay to the publishers to provide access for the faculty and researchers (Figure 8). This is the traditional subscription-based model. The natural alternative business model is the one that Dr. Ojiro earlier referred to as the gold open access approach, that is, the publishers would receive fees not to provide access, but directly for the services that they provide to authors (Figure 9). In this approach,



(Figure 8)



(Figure 9)

instead of the university paying the publisher for access, the authors are paying an open-access publisher directly for publisher services. Therefore, the faculty members are paying based on their role as authors instead of on their role as readers (Figure 10).

We might like to have this alternative business model for journals, but perhaps you can see the problem with this. An author choosing to select a journal to publish in would have to pay a substantial publication fee of, let us say, \$1,000, \$1,500, to publish in an open-access journal, whereas in the traditional journal, the author pays

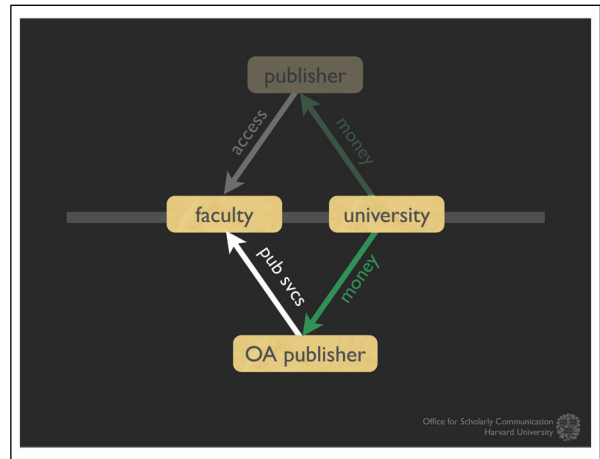
nothing. One obvious solution to that problem maybe this, the universities should be willing to pay publication fees just as they are willing to pay the access fees to the publishers.

This is not a new idea. There is a history of research on the economics of scholarly publishing as early as the early 1990s with the economist Roger Noll, who, looking at the scholarly publishing system, said that the best means for accomplishing the social good is subsidizing the cost of publication of these publisher services. Not only is it the best way to achieve social good, it is also necessary in order to put open-access journals that charge publication fees on an equal footing with subscription journals that charge subscription fees.

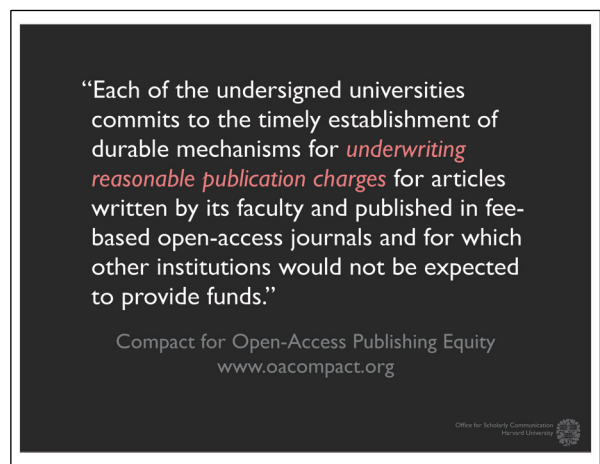
Compact for Open-Access Publishing Equity (COPE)

With this kind of subsidization in mind, we set up a kind of compact for open-access publishing equity. A group of universities developed this compact to place the open-access business model on a more level playing field with the subscription model. This is the key sentence from the compact, and it says that the universities that sign on to this compact commit to underwriting reasonable publication fees for articles written by their faculties (Figure 11). We essentially are saying if publishers are willing to move their journals to this new business model, we are willing to pay the fees necessary to operate in that business model.

So, I think all universities should make this commitment, urge your universities to sign on to the compact. Unfortunately, not all universities have. It started with a group of universities, Cornell, Dartmouth, Harvard, MIT, and the University of California, Berkeley. They were the



(Figure 10)



(Figure 11)

initial five signatories of the compact, and since then, a set of other institutions have signed on. Recently, CERN in Europe has also signed on as a signatory of the compact.

This compact, which we call COPE or Compact for Open-Access Publishing Equity, has an impressive group of supporters. In addition to the signatory institutions, there are other supporters for COPE; for instance, a group of over dozen Nobel Prize-winning scientists. Also, many open-access leaders are supporters of COPE. Various institutions, publishers, scholarly societies, and funding agencies are supporters of COPE as well.

Questions (& answers)

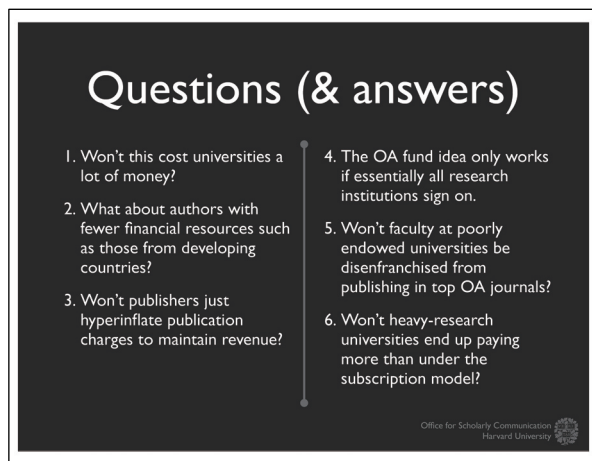
Now when people hear about the idea of this open-access publication fee business model, the so-called gold open access, and the compact for universities to support these kinds of journals, many questions come up, some of which are very legitimate, and I have listed some of them here (Figure 12). To the question of how this kind of commitment would be implemented, people worry about whether this will cost universities a lot of money in the short term. People worry about authors who have fewer financial resources such as those from developing countries. So, you may worry about what prevents publishers from hyperinflating publication fees, just as they have been hyperinflating subscription fees for several decades now. There are some more questions on the slide and other questions you may be thinking of yourself. I do want to assure you that there are positive answers to all of these questions, and if anyone is interested in what they are, I am happy to answer those questions.

Conclusion

Let me just conclude by making the following points; first, open access is and should be our goal. Second, there is a systemic failure in the scholarly publishing market based on subscription journals that are preventing us from reaching this goal.

In the short term, open-access policies like the kind of rights retention policy that we have instituted at Harvard and at other universities can help to mitigate the symptoms of the underlying market failure.

But in the long term, we are going to need an alternative business model, one that is efficient, effective, and sustainable for scholarly publishing,



(Figure 12)

and we should begin looking at alternatives now and supporting them now by policy.

Thank you very much for your attention.

◆ ◆
● Q 1 That was a very informative speech, thank you.

I think it is necessary to consider the differences in social structure between Japan and the U.S. At the beginning of the slide is written, “the greatest possible public benefit.” It will perhaps be simpler for Japan to achieve this, because in Japan most things are covered by public investment. In America’s case there is a great deal of investment from corporations, which I feel might pose some difficulty. On this point, as an American university, how is this dealt with? I would be grateful if you could share your thoughts on this matter.

I also have another comment. Again in the area of social structure, I am currently working for a government group think tank and I have been an administrative official for the State over the past 20-odd years. Based on this experience I feel that initiatives such as COPE are really excellent, but in

Japan some universities act independently, and this makes it difficult for the government to issue special financing to these universities. I think it might be simpler if the secretariat of COPE was an international organization like UNESCO, for example, whereby the Japanese government could contribute to the organization. Contributions from the Japanese government might well depend on the style of the structure.

●Shieber I understand what you are trying to convey. If at the beginning you were saying because of the funding situation in Japan, it should be easier to work towards open access, then that is good and I hope that that is true. The second comment was about SCOAP³ versus COPE, and I would not put them as alternatives. There is nothing inconsistent about pursuing SCOAP³ and COPE approach. They are both attempting to achieve the same goal of support for the gold open access in different ways. SCOAP³ is attempting to do that by putting together a critical mass of institutions to manage a group of journals in a closed field at once, in particular, in particle physics. On the other hand, COPE works on an article at a time but is not limited to a single field. It is much refined and much easily deployed.

●Q 2 I would like to hear from your university, the Harvard Medical School, how they work because they manage the famous journal, *The New England Journal of Medicine* and they get money.

●Shieber *The New England Journal of Medicine* is not published by Harvard Medical School, but by Massachusetts Medical Society. However, it is physically located at Harvard Medical School.

Harvard Medical School is not on the list of schools at Harvard that have instituted an open-access policy. The different schools at Harvard are very separately run and have different faculty sizes. My own school, Faculty of Arts and Sciences, is the second largest school with around 750 faculty members. With 750 faculty members, it took us about 2 years of discussion on campus to ensure that the entire faculty was comfortable with this policy, understood it, and realized that they would retain their free choice as to whether to retain rights or not. This was prior to February 2008.

Now, the largest school at Harvard is the Harvard Medical School with approximately 10,000 faculty members. It was hard to bring together 700 faculty members, so bringing together 10,000 faculty members is more difficult but not impossible. The vast majority of researches at the medical school are funded by the National Institutes of Health, which has its own open-access policy. So, the medical school faculty finds it less urgent, although it would be helpful to have this kind of policy. But we are continuing to work with the medical school and the other schools at Harvard. In fact, there are three or four schools at Harvard that we are working with now.