

*Opening remarks by Clifford Lynch at "Open Access Publishing: A Panel Discussion"  
sponsored by University Libraries in recognition of Open Access Week 2010*

*October 20, 2010*

Thank you. It really is a great pleasure to be here.

I'm going to take about 20 minutes to start here, to give you a kind of a broad perspective on open access.

Now I know many of you here are familiar with various aspects of open access, particularly open-access publishing and indeed I think we're going to hear a lot about open access publishing but I want to put this in a broader context. So, what is open access and where does this come from?

The way I think about it is open access is kind of a pervasive philosophy, it's a way of thinking about the mission of the academy and particularly the part of the mission that encompasses not just creation of new knowledge but the transmission of knowledge and the stewardship of knowledge. It's about asking questions about how can we most effectively advance scholarship? How can we most effectively open up access to scholarship to the largest number of people? How can we accelerate the pace of discovery? How can we make our investments in research and scholarship yield the greatest returns? Those are the kind of big, complex issues and motivations that I think drive open access.

Open access, of course, is also used as a shorthand for a very specific set of issues around the belief that articles in scholarly journals should be freely available and there are lots of details and nuances of that which I think we'll get into later. This can be accomplished in many different ways. One way, obviously, is to publish in journals, select publication venues as an author that make articles available without barriers and there are now a rather large number of scholarly journals covering a tremendous range of disciplines that do exactly that. They are financed by methods other than traditional barriers and subscriptions. Those are sometimes called "gold journals."

Another approach is to take a copy of the article you've written and place it somewhere like an institutional repository and this gets into some complexities about how the journal you're publishing in feels about that and whether it's cooperative in such a venture. Friendly journals of that type are sometimes

called "green journals" and that practice is often referred as author self-archiving.

There are actually, although it's not talked about as much, a number of other sort of more limited initiatives toward open access. I'm not one to let the perfect be the enemy of the good, it's worth recognizing that many of the traditional restricted access scholarly journals now open up their back files after a few years..it's sort of a moving wall where they keep the most recent year or five years locked up so that they can stay with their existing subscription business model, they make their back files publicly available which, you know, is a very substantive contribution to opening up access to scholarship in it's own right.

I think that the panel is going to get into a lot more detail on that but I wanted to sketch out that terrain and leave you with one final observation, specifically about open-access journal publishing.

The kind of archetypal case here--the picture has always been one of human readers. About, I'm a student somewhere or I'm just a member of the general public trying to find out about a disease that has suddenly become a passionate interest to me because one of my parents has it or something...but it's about an individual trying to get access to scholarly articles. That is the important and driving case but it's only one of two important cases we need to be mindful of.

Increasingly, the scholarly literature is becoming a target of computation as well as something that humans read. In fact, we're seeing tremendous progress in tools and algorithms and systems that can extract knowledge from large corpora of text, make connections, generate hypotheses and we can already point to some cases where there have actually been significant scientific hypotheses or discoveries mapped out and facilitated by that sort of text processing. So when we think about access to the literature, I urge you to think not just about human access whether it should go forward but also about the importance of us all being able to compute on that literature. And for me that is, I think, increasingly one of the driving forces that will push us toward more and more open access.

So let me conclude my comments about open access publishing specifically then with just a word about "why would you do it?" It's interesting, there are at least three kinds of classes of reasons why scholars might choose to publish in open access venues:

One class of reasoning is really about self interest. There are a number of studies and while this is a matter of some, I think still serious and legitimate scholarly debate, there is at least some evidence that works published in open access venues get more citations, they gather more readers, they have a greater impact quicker on the field. Now obviously, this is a very mushy thing to measure because each work is unique and we really don't have any kind of simple, linear metrics for quality of work which is clearly an important factor here, but there are a number of sort of pragmatic and self-interested reasons why scholars might choose these venues in order to make a greater impact--to be heard more widely.

The second set of reasons are really about sort of philosophical alignment with the broad principles of open access. The sort of simple statement "well, I choose open access channels because it's the right thing to do."

The third is because someone is making you do it. Increasingly, that's becoming a factor. We are starting to see funder mandates, both governmental and private foundations, which call for open access to the products of research and here you see, for example, arguments that seem pretty reasonable to me, that say 'If we're paying for research with collectively funded federal dollars that we all pay into through our taxes it's reasonable to think that the results of that research ought to be publicly available.' There is no case where there that's more compelling than in the biomedical areas where we spend enormous amounts of public money through agencies like the National Institutes of Health and yet you regularly see both physicians and individuals who are trying to make decisions about their own care having difficulty getting access to the latest results of those research investments.

Private foundations are in some cases reaching the conclusion that open access actually is a good tactic for accelerating research and discovery and it's consistent with the objectives of those private funders and they're writing in as well, as part of their expectations for grantees.

We're also seeing some institutions make collective decisions that they want to do open access as a kind of institutional norm or practice and this takes the form most commonly of the collective faculty mandates that you've seen from a number of institutions over the last few years.

So those are, you know, the sort of three big motivators: the philosophical, self-interest, and requirements from funders or institutions that you belong to.

Obviously, it's not uncommon for people to act from some mixture of all three and not just one.

But, let me having set that kind of scene a little bit, specifically about open-access publishing, step back as I indicated I would, and spend my last few minutes on some other dimensions of open access and how they might show up in individual choices by scholars or institutional choices by institutions of higher education. Other kinds of policies that advance the same kind of objectives about improving scholarship and the dissemination of knowledge.

One area that's receiving a lot of attention right now is the question of data sharing, data archiving, data stewardship and data reuse. Publications are just one part of scholarship, the other part is data: the observational data or the experimental data or other kinds of data that's generated in the course of doing research. And increasingly we're recognizing this as a major asset in some right and that by sharing and reusing this data we can really do great things to advance the pace of scholarship and to gain new kinds of research perspectives. So you're starting again here to see funders asking for data management and data sharing plans; you're starting to see journals and scholarly societies in these areas having conversations about norms and expectations about data sharing and data reuse and obviously there are numerous issues to balance here.

One complex of issues is around the people who went to the trouble of amassing that data getting some credit when it's reused but there are also real limitations on the ways in which data can be reused, especially data about people. You start running into some very difficult issues about how you balance off ideas of informed consent with ideas of open ended data reuse, for example. But this is an area that's clearly going to be of growing importance.

Another related area is open source software production. A tremendous amount of software is produced in the course of scholarly work now and to the extent that that can be made available for reuse and building upon, that's a very valuable thing to do. It helps to extend the value of the investments; it helps to insure that software doesn't automatically die out just because someone loses interest in it and moves on to something else.

There are a couple more...at many of our educational institutions we have great libraries, we have museums, we have archives, we have research collections at the departmental level. A tremendous amount of that material is out of copyright.

It should, in my view generally, be freely available in digitized form for reuse for scholarly purposes. And indeed, a few institutions are establishing collections like this. The Cornell libraries have recently said if the material is public domain we view ourselves as stewards of it and want to see it widely reused. We don't want to put up barriers there. You can contrast that with the historic practice that many libraries have followed of trying to charge for use of this material. Museums as well have behaved less than ideally in this area.

These collections are held in trust for us all and in the digital world they can be opened up and reused in ways hitherto unimaginable in just the same way that we can imagine that journal articles which used to be rooted in an economy of paper and shipping paper around and therefore had non-trivial marginal costs for additional copies. In the digital world, have basically almost zero incremental costs for very, very wide distribution once you create the material in the first case.

There are other areas...many institutions are thinking hard about their role in the production and dissemination of open educational resources of various kinds. Initiatives like MIT's OpenCourseWare. Another place that's very important is electronic theses and dissertations and even beyond that, opening up things like masters projects or capstone projects for various degrees to public access.

This is very consistent with the historic practice about dissertations being filed in libraries, for example, but it takes it to a much broader level and is helpful in showcasing the scholarly output and the contributions of the institution and the students who are part of it and the faculty work with those students.

I could give you more....aggressive support for fair use in scholarship. Moving away from a set of practices where publishers, including university presses, want you to get explicit permission for everything you quote and where they start with fifty dollar usage charges in order to get that letter saying it's ok to quote those sentences...which seem to me to be covered under fair use. Making policies and supporting people who take an aggressive view of fair use in the practice of scholarship would seem to be another area. I'm sure you can think of more and I hope that in our conversations you will suggest other areas.

But I mention this list to really suggest to you that while open access publishing, specifically in the journal area, has been a key flashpoint in the last couple of years for a number of very good reasons, the broader sort of philosophy of open access as a tool for advancing the work of scholarship is really one that has a pervasive set of implications in the policy choices and practices that individual

scholars may choose to follow; that communities of scholars may adopt as norms; that institutions may adopt as norms and policies; and I hope that from time to time you'll step back and consider those kind of broader implications and the way they may manifest themselves in your own work and actions.

Thank you.