Commonplace •

More to Open than Access

Peter Murray-Rust, Sarah Kearns

Published on: Nov 15, 2021

License: [Creative Commons Attribution 4.0 International License \(CC-BY 4.0\)](https://creativecommons.org/licenses/by/4.0/)

In this episode, we'll be chatting with Peter Murray-Rust, a chemist at the University of Cambridge, about open beyond just access, and the challenges of making changes that will have a global impact on Scholarly Communications. Hope you enjoy.

SK: So maybe we could start off with you introducing yourself a little bit more. And also, in addition to sort of why and how you got into, and arguably contributed to, so much of the beginnings of the open access movement.

PMR: Okay. I'm Peter, Murray-Rust. I'm a chemist, I'm very much interested in modern information technology and communication, and I take a global view of things. I'm probably not really so much "open access" as "open". And by open, I mean all sorts of things, which, if you take an open approach—meaning that you share with people, that you're immediate—you change the world. We'll probably talk about that later, but open access doesn't yet have all the aspects of open. Now, I got into open probably about the late 1980s, with working on chat groups and things like that, on the... before the internet. Well, not before the internet, before the World Wide Web, and so on. And getting involved in online communities, and developing software and so on.

SK: I guess, could you go into a little bit more about sort of like, the differences between like the open movement and sort of open access, there? Like the mindset shift between the two?

PMR: Right. So open probably started roundabout the mid '80s. The seminal thing was actually "free", with Richard Stallman freedoms for source code; the freedom to read the code, to run the code, to modify the code and redistribute it. Stallman doesn't like the word open, but it's a general thing at the moment. And it does mean that it's about freedoms, it's about sharing, and about building a greater community by doing it. So I got in, really, because I did a lot of coding, and I met these various groups. And at the beginning of the 2000s, met up with the Open Knowledge Foundation [OKF], particularly Rufus Pollock, who started it. And that really catalyzed a great deal of open, and open was applied to documents, to code, to processes, to maps, to open corporate: all sorts of things. And we had meetings—OKF ran meetings—two or three times a year where all of these open things were discussed, and that was sort of "red hot" in the middle of the 2000s.

SK: Yeah, I feel like I remember... it seems to me that sort of like “open source” and sort of “open”, if we're using data in that sort of code-based way, seemed much more... easier to apply. Or like, much like less hesitation to adopt. So it makes me wonder, like, what sort of differences you you've seen between open code and like open scholarship, because it seems like there's like a much harder, harder way to have more scholarship... have that sort of same like ideology of transparency and sharing and reusing.

PMR: Well, open doesn't automatically apply to everything. And in each particular domain, you have to work out what works and what doesn't. The most important thing is to work out what your values are: why are you doing it, what do you hope to achieve, and so on. So, “open source” was very much about the fact that your code was part of something larger. You would take other people's code, you would add to it, you would republish it. And by this way, you would build something larger. It's also the idea that by going through this process, you may get better, because each person would correct problems from earlier versions—add more to it, and so on. And based on the idea that you often ended up with a small number of, you know, implementations of a tool, without everybody trying to do the same thing—so it was about using the commonality there. And the problem with open access is that that hasn't happened. There's really very little public statement as to why open access is valuable. There's an assumption that it is. But people don't look at the way in which other groups in open interact with this. There's this idea that you get scholarly knowledge published to the world, but not necessarily seeing what the benefits are. Does that make sense, Sarah?

SK: Maybe it does and maybe it doesn't... Because, now, since the open access, open, or open access-slash-open movement has sort of changed over the past many years, since it sort of has started... so maybe, maybe I'm just in a framework of like, for, and I'm also in a position where I'm like: “I think open access does have value and like, does improve scholarship.

PMR: It tends to make it more immediate. But it doesn't make it as immediate as it could be. Who says that, if you like, the mainstream approach of open access is, people do some work; they keep it to themselves for, let's say, a year or two while they're doing the work; then they decide to publish it, and then it might spend some months going through a publisher; and finally, you get a static report at the end of what they did. If you take open code, then what happens is the first bit of code you write goes straight on the web, where the whole world can see it. So it's immediate. It builds. People don't mind getting it wrong, because other people contribute to it. And so you

have a dynamic system, where as many people as can get involved in creating this, refining it, using it, and so on. And that process of reuse often has a major effect on how valuable this is. That doesn't happen in open access, which is very much “we are going to create this great work; we're going to take a considerable time to do it; then we're going to put it out, and people are going to be able to read it for free.” And that doesn't mean you have the spirit of open—of collaboration and involvement. Now, there are exceptions to that, and we can have a look at some of them. But generally, most open access papers do not describe an open process. They simply publish the result of that process.

SK: And it can be hard for some sciences, too. Like, I have a biochemistry background myself, and just some experiments are just not worth sharing immediately. A challenge, maybe, for a particular field—or I know like, astronomy; I'm reading this book now about black holes, where it just takes forever to analyze that that data.

PMR: Are you aware of Galaxy Zoo, for example? That is an example of an open process, where the Oxford group including Arthur Schmidt said that they would put, or they would put the task of indexing and cataloging a million galaxies out to the world in general, see who volunteered to help. And they speeded up the process by, you know, literally 20 times or more. And they also found new things as they did it: it was quicker, more immediate. And it was very much a communal approach to doing a data-driven science. Whereas the classical approach would be that people hide away in their laboratories, or they talk to the people in their field. And then at some stage, they release results.

SK: To scholars who are more siloed, what would you say to them that would, like, help them see that open access could be useful—or just the idea of open, beyond just access?

PMR: I wish I knew. I mean, I've been trying to do this for 20 years to try and come across with examples of where open benefits people and a relatively small number of crudely opens studies that you can point to. One of the best in our field is Open Source Malaria. So Matt Todd works in Sydney; he's now moved to London, and his he was a chemist making antimalarial drugs or potential anti malarial compounds. And he decided to make everything open on the web as soon as he did it. So literally, within a day of making a compound it would be on the web. Now, in my view, that actually makes it better science, because you've gotten a discipline of making sure that it is done properly. You know, there's an... I've done this myself, not with experimental science, but with information science: that there's a greater responsibility to make

sure it's right at the time. And you get into the habit of doing this, not being afraid if you make mistakes, and correcting them afterwards. That's not a crime itself. It's a positive. Matt and his group put this out, and as a result, they actually found people contributing to this in a way that they wouldn't have contributed before. There are clearly some areas of science where, if you like, a certain amount of secrecy in research, or non-publication is valuable. I think it applies to the cases where people are doing medical research, and it is important not to go out with half baked results where people can misinterpret them, and so on. And we've seen a bit of this during the Coronavirus, and ideas of drugs which might be valuable, and people pick them up and so on. But I'm a great believer in the preprint ethos of saying: we put this out before it's published, the world can review it, they can comment, and then we can decide whether it's worthy to have a stamp put on it that says since it's validated by some process. Now, I'm actually *not* a great believer in the pre-review process. I think that much of this can be done with publishing it, and letting post-review by the community review it. If it's not important, nobody will comment on it. If it is important, you'll get intense comments on it. And a consensus or a dichotomy will arise as a result of doing it.

SK: Some of that, too, might be influenced by, like, algorithms and other types of... not not data manipulation, but just like hype that one can have on the internet. I think that maybe some good science could get lost in that hype. Coming back to the idea of preprints and preprint servers, I think another thing that sort of has become like sort of emerging and sort of goes one step of the process before peer review, is open grants, too. I think that that's a something that's like, you can have, like, open your grant up, or open your grant submission to the public before you even start the research, and get comments and feedback before you even start the experiments. I think that that's a cool thing that's emerged as well.

PMR: I agree. And we've done a bit of this, where you put out what you want to do to the world and say, "What do you think about it?" I think at the moment, there's not been a huge engagement on this yet. But I think it's a good way to do it. And, of course, part of the problem here is that academia is now hypercompetitive. And if you release ideas, then other people kind of claim them as their own. So there's a disincentive to do that. Now, that in my view is not a criticism of open. It's a criticism of the values that are within academic research at the moment. What we've got in the information revolution—and it is a revolution—is we've got potentially all parts of the process available to everybody to take part in. So, you can take part in data creation, you can take part in analysis, you can take part in writing code, you can work with

people to provide different parts of this. And to come up with the particular... I'll call them artifacts; products; whatever down the line that people want. Now, there needs to be structure that in that, and so on. But if I tell you a little bit about what I'm doing at the moment, and this is working with a group in Delhi, who are looking at plants, and we're looking at the genomes of plants: the National Institute for plant genome research in Delhi. And what we're looking at is, how can we find out what enzymes are produced from the genome and what chemicals those enzymes produce? And we're using the literature. So, we're analyzing the existing literature and pulling out mentions of plants, plants and their genomes, and trying to tie that up. There's a lot of practical interest in it, because this deals with different crop plants, and that's some of the main thing. And over the last five years, I'd say, roughly, there's been a huge explosion in the ability to analyze and understand plant genomes. So we've got the real opportunity of both understanding this, and in many cases of modifying the way plants work. Now we're doing this on a very small scale, so we're doing it with undergraduates and Masters' students in Plant Science and Information Technology. But we're building the tools that allow us to read the plant literature, pull the parts of the papers out—which are about genes and their products—systematize this, because there's a lot of problem in nomenclature, different ways that people describe things, and then hopefully coming up with more information thanx is already available in databases, because a lot of the information is published in text form in the literature, in a very poorly structured manner. And what we're doing is trying to pull this out, and make sense of it. The amount of information is exploding, so, you know, we probably come across at least 50 model plant species, and many of them have 50 different enzymes in this area. They do all sorts of things themselves: they attract insects, or they repel insects, or they are growth hormones, or whatever. But also, they're the building blocks for other chemicals in the plants. And so, lots of this is in the literature, in a text form. And we're trying to pull this out using software that we and others have written to do that.

SK: So the project, and sort of like the global connection of this project, of like this natural product project, also... it just makes me think about, and wonder... It makes me think about how the access to literature and the access of information and data is, how do you say, disproportionate.

PMR: So, let me pick up on that.

SK: Sure.

PMR: So, the idea of open access to literature has been around for a long time, probably about 30 years. And the ideal for many of us is that all science is open to everybody, no matter where they live, whether they are an academic, or they work in companies, or they're a citizen scientist, or whatever. Now, unfortunately, that isn't the case at the moment. And it's a long way from being the case. And there are a number of restrictions. The first is that this information is technically copyrighted. And many people see copyright as a way of making money out of it. Over the last 30 years, a number of publishers have made lots of money by being the single gatekeeper for publication, the best known is probably *Nature*. And if you're a scientist and you publish in *Nature*, then you get much more respect than if you publish in some lesser-known journal. I don't believe this is a good judgment, but that's how it is. And so, you get to a stage where people say: Okay, well, we must publish this in high prestige journals, because that way, you know, our career will be advanced. So, you've got a lot of couplings which are not present in the other areas of open. You've got the fact that publication is not just or even primarily for dissemination of knowledge, and the collective use of knowledge. It is for personal advancement, or the advancement of the organization you work for. And those are big conflicts. Scientists would publish in prestige journals, because that way their career was advanced. We all agree that it's important to publish your results, and publication, or “publish or perish” was, has been, a phrase for many years. But it's, in my view, gone too far, where the important thing is not the dissemination of knowledge but the advancement of the person who's published it. It's a hybrid; it depends very much on the field, the results, and so on. But there's no doubt that this is a huge industry. Publishing is probably worth about \$20 billion a year. Billion with a “B”. And that's larger than the music industry, for example. And when you get a huge industry like that, it runs it for its own benefit, and not for the benefits of the people involved. And so, what you end up with is an industry which is self-serving, and the primary objective is to make money. And so they will develop the product in a way that makes money for them, rather than is the ideal product for disseminating knowledge. And we're locked into that. We have something like... oh, well, probably 25,000 different journals. There are probably 300 major publishers. And unraveling all of this is extremely difficult. We're locked into the fact that we publish in journals, because this is convenient for the publisher. And they're able to say, “This journal is a better journal to publish in that journal,” rather than, this is a better paper than that paper. And that's wrong. Simply wrong. Because the value of a paper should be not where it's published, but what's actually in the paper. We've got this thing that we honor the container rather than the content.

How do you change a system which has a \$20 billion turnover, without any regulator? If you take comparable industries, they have regulators who decide what's ethical, what's a product and so on. Here, we have total lack of regulation. I call it the publisher-academic complex, because it's out of control. And, and how will it change? I don't know. I'm trying to change it in a very small way, including, for example, by doing this podcast with you. So hopefully, that helps a message to spread a little. It may be that at some stage, the world gets fed up with the present position, though I doubt it. But we have... if we take COVID, for example, we have the situation where 80% of the relevant literature is still behind paywalls. Now people will say, "oh, the publishers released all their content." They didn't. They released content with the letter C-O, V-I, D, in it. And for a limited period. But they didn't release that picture with face masks in. If you want to find out whether face masks are valuable, you will find that between say, Elsevier and Taylor and Francis, 80 to 90% of the literature is behind paywalls. That is simply unacceptable, because it means that critically important studies are not available. I don't know how we change that, because those papers will always be closed access. Plan S won't change the existing papers, it will only change new papers. And so we will have a huge backlog of papers, which are always behind paywalls. And some of those are very valuable. There's a paper that predicted the Ebola outbreak. It was written 40 years ago. It's behind a paywall. It's still behind a paywall, and its prophecies have come true. But if we have that situation where the scientific literature is behind paywalls, then we're not using it properly. So Plan S does not aim for total liberation of the literature. It aims for liberation of perhaps 25 to 30% of it, something like that. Maybe 10 years down the line, we will have 50% open, but the backlog will always be out. And you'll have a hybrid system, which is appallingly difficult to work with. Building tools to read this with machines is extremely difficult, let alone the fact that publishers actually actively prevent people from using machines to read the literature.

SK: I guess with plan S, it seems like it's almost like an investment into the future. Like right now it's 80% behind the paywall, but maybe in like 30, 50 years, new science papers will all have access. What what would be like a better balance between, or better a better investment, I suppose. What would be like a better investment?

PMR: Well, this is a political problem, not a technical problem, right? And until we solve the political problem, it's difficult to know what to invest in it. I mean, I can tell you my dream of an open infrastructure for knowledge. And I'm building it, but there's

only me and one or two of my colleagues in India at the moment doing this, and some scattered people, and the reuse of Python code and so on. But the, um, the dream of a better infrastructure would be where everything was published as preprints immediately, where we had indexing tools, and these tools would not just index the words they'd index the tables, and index the diagrams, index the chemistry. I want to say, tell me today, which papers have been published for eukaryotic for—sorry—diaphoretic leaves in plants, which contain more than 50 terpene synthases, which have Vera Niall diphosphate, as their... the feedstock. I can do that on my own system (right?) already. And the reason I can do that is because we're incorporating a major new ontology, or set of dictionaries from Wikidata. So Wikipedia now has a classification of nearly every branch of science, and if you want to know about a plant, the first place I go to is Wikidata. It's all open, and if we could index the whole of the world's literature using Wikidata, then we would have a system which was miles ahead of where we've got at the moment. But at the moment, what's happening is, the publishers will try and come up with search engines, which they choose, obviously, the less they can invest in the search engines, the better. So you end up with poor text indexing of the literature, rather than content, full content indexing. If you try and index chemistry, the American Chemical Society will send its lawyers to stop you indexing chemistry. They tried to shut down the Public Index of Chemistry—Pub Chem—from the NIH, about 25 years ago, because it would threaten their closed business model. So we have major players, who have a major interest in stopping open development of infrastructure and software.

SK: And I think too that, kind of what you said about, like, how we've just bought into journals... like it's hard to convince people to sort of change their ways, even if it could be a better way. What, like, economic models could be sort of like maybe like a stepping stone, in between, like, the dream and the reality?

PMR: That's a very good question. Clearly, there is a cost involved to publishing. The cost is about somewhere between, probably, \$100, \$500 per paper, something like that. If you look at the fact that the average paper probably describes about a quarter of a million dollars of funded research, something like that, then you're looking at a very small gearing. And the Wellcome Trust, for example, has said that it will pay 2% of research funding for dissemination costs, and so on. So if you get that, it would totally cover everything you need—if it is done on the basis of cost, rather than a glamour-based price. So if we say \$500, and I think for many cases, it's less than that. There are many journals, which don't charge anything either to readers or authors. But if we take \$500, then we can do that across the board. But if you choose, say—*Nature*, for

example, is \$12,000. And that goes into all sorts of things like, you know, a huge rejection rate, a glamour presentation, et cetera, et cetera... then it becomes more difficult to find the costs. So if you did it at cost, and if you had a body that decided what was a fair economic cost... I keep using the word cost, which means what it pays to get something out, rather than price, which represents how much you can get out of the market. I mean, you know, these publishers have profit margins of 30% or more, and that's pure profit. Also, I think that their business model is incredibly inefficient, so a lot of that goes into funding the corporate nature of publishers. But if you took the costs, and allow for that in the system, then you would find that funders like Wellcome—national funders—could do this. And if you take South America, much of the publication there is funded for... funded by the public purse. There are initiatives such as SciELO in Brazil, and Redalyc, run from Mexico, and so on. And these generally come out of public funding. It may not be easy to find it, but the idea in much of South America is that the purpose of publication is for the public good, rather than for the advancement of the author. So, the money should not be a problem. But if you have a... if you take \$25 billion as the market, whatever it is, then there's no difficulty in finding the money. But you will not be able to tear it away from this, this set of publishers, which have got their hooks into people at the moment. And it's a very nice way of getting a huge income for very little effort.

SK. And in that description, you mentioned that some of the... was it, like, the journals or the research that's coming out of South America and Mexico? You said, or I guess you mentioned, that was hard to find. I guess, in what way?

PMR: Did I say it was hard to find? It's not actually hard to find it. But it's often regarded as second-class by indexing tools. So, if you take: how do we find papers? We use Web of Science, we use Google, we might use CrossRef—which I'll come to, and which is a good institution. But if you take Web of Science, they don't index much of the non-English language, because they don't regard it as saleable. Remember, Web of Science, or Clarivate, or whoever does it, are only interested in what they can sell. So their market is to rich, Western, English-speaking organisations and not organisations in non-English or the Global South. And here, for example, there's a standard feeling well, this is second class science. It isn't. This science is as well done as Northern science, in most cases, and if you're interested in things like global pandemics, then it's pretty important to have science coming out of all the nations of the world. If you're interested in the species that harbor viruses, it would be foolish to neglect countries, which have many of these organisms. Ebola came out of Africa. You want to know about bats. You want to know about Zika. You want to know about where the

malarial, or...malaria. Where the Zika mosquitoes come from. Much of that will be published in CCs, for example, and in the countries where the research was actually done. But it won't be indexed.

SK: Is that then a problem of like the journals then that they're publishing in? I guess, this whole conversation—and I mean, and beyond, even—has gotten me... it's like, sort of directing towards like a non-journal-centric publication platform method. But my concern, with that would be like, if there's a lack of journal, how do you have like a structure to have that indexing even occur?

PMR: Well, let's look at where research is currently done. And let's concentrate mainly on academia, although you mustn't forget things like startup companies, and let's say biodiversity groups, and so on. You know, so: local scientists organizations. But if we concentrate on universities, research starts at the Master's level. A lot is done in PhDs. And then a lot is done by research groups, in universities. And of course, a lot is done in national laboratories. And in some subjects, big science is done as international projects and so on. And that's often less of a problem than, you know, the laboratory subjects such as biology and chemistry. But much of this work will start off in PhD CCs, and my view is that it's technically straightforward to index all the world's PhD theses. And there will be lots of stuff in that that is valuable, but it isn't published or whatever. You take a PhD student; they've done this work; they then leave. The supervisor then tries to get the student to help write it up while they're looking for their jobs, and so forth. And often, it doesn't get written up or whatever. And so, the place you'll find it is in the thesis. There's nothing wrong with that, so long theses are public documents. It's been quite a battle to actually get many universities to publish their theses, but it's becoming common now, and so forth. So I would say the first thing to do is index all the theses in the world, which is not too difficult in many countries, because many countries have natural indexes of theses. France does it, Netherlands does it, many of the Global South countries do it and so on. But in the UK, much of this is done on a university-by-university basis. So there's no framework for searching this. You know, I've tried to build one, but it's a lot of work, because these systems are not designed for indexing. And in the US it's worse, because there are companies such as ProQuest who actually take over the management of theses and own the distribution rights and will not allow you to index the theses. It's their income to index theses. And so long as you get companies, which are revenue-driven controlling information, you will end up with these distortions. So, I'd start there. I'd try and show that there's lots of valuable science in theses, and that couldn't we also do this with preprints. I would also I take this idea of combining theses and preprints, and index them and show that you got

high quality information at an earlier stage in the process. Because sometimes it takes years for publications to go through pre-review, pre-publication peer review, and sometimes it fails. Whereas if you do theses and preprints, you can get anything from two to five years earlier than it would normally come out. So that's a model I'd look at.

SK: How could also we... maybe, also, help convince scholars that like, the information behind these things, is not just even just the scholars, but the scholars and the publishers: that there's information that's locked away that's important, and like, we can even argue, like revenue-driving? Like, that are in places that are being hidden. So I guess I'm thinking of, like you said, like, oh, we should know more about like Zika, which will come out of research in Africa, or we should know more about bats, which have come out of research in China. Not like America, not Europe. So, beyond having a huge breakout of Zika, or a breakout of COVID, sort of what could help drive the motives and values of researchers and scholars to sort of appreciate, or to like index more or to like have systems-level changes that make knowledge generation and like data curation more, like, equitable?

PMR: I wish I knew.

SK: I know that's a huge question.

PMR: I've been thinking of it for the last 20 years. If we look at some success stories... So, okay. Let me go back to the problem. The problem is you have a \$20 billion industry, which is going to fight to the last tooth to control it. That's right. They have no interest in changing journal systems, they have no interest in making this material open. They will fight all the way. I mean, I've done this for 10 years with companies like Wiley, Springer, and Elsevier, and you know, they resist everything. They are not our friends and collaborators. They are the problem, not the solution. I think, you know, we have to get rid of commercial publishers. Now this is going to be very, very difficult if we look at history, there have been revolutions, for example, in newspaper publishing, where the printing press came along and changed it. And so on. I wouldn't say it's made it better, but it changed from a hot metal press to electronic and and so on. Technology sometimes changes things. But here, I don't think technology is going to change things. There isn't that sort of drive to do it. The place where it won was where the open source movement beat Microsoft. In the 90s, Microsoft had a near monopoly of software. Their mantra was "extend, embrace, extend, exterminate". So they would ruthlessly destroy any contra... any competition. But what beat them was the fact that open was faster and better. And so, people were able to create tools which move faster, and ultimately Microsoft have come around to the view, that open is a key part of their

message. The community and the technology won. But in scholarly knowledge, you can't do that. Because each journal is unique. Each paper is unique: you can't replace one paper with another. So, you cannot use technology to change the current situation for the better. You've got to change the whole system. And that requires political will, from somewhere or other. Political will sometimes comes from on top, where... someone comes along, or some organization, says this is unacceptable, and ruthlessly changes it. And there are big fights. So what Plan S has done is to show actually, that the model is working with the publishers. And so long as you work with the publishers, you're actually in a hostage situation. You are working with your captors. So there's no way that Plan S is going to change that. They're going to put a lot of money into a small section of it, which the publishers will still control. And what the publishers ultimately want to do is to control the infrastructure. Now, the only way that will change is either if some radically new technology comes along, which cannot be monopolized by people—and I don't see that at the moment—or whether it is a cultural change, which might come from young people saying they're so fed up with a system that they're not going to play, and there will be a revolution. And I don't see that happening either. So I'm actually pretty despondent. If government organizations said, “We are not going to work with this system, because it doesn't represent what we want and it's unacceptable value,” you might change it. But when the NIH, as I say, tried to do this, they were fought by the American Chemical Society, and the American Chemical Society can go to senators, and they lobbied the Senator to have PubMed free, which was free chemical information, closed. And they very nearly won. So, I don't have a magic model.

SK: Yeah, no. I don't think anyone really does. Or we'd probably see those changes more than we than we do now. It's super challenging to try and change like an economic model, a political model, and a cultural model, sort of like simultaneously around the world. So yeah, how do you how do you do that? I don't know: slowly and steadily and steadfastly. Maybe I'm a little too like idealistic about it, where I'm just like, yeah, everything will be fine one day. Everything is gonna be great one day. So it's nice to have that sort of tempered sometimes, for sure.

PMR: So you talked about, essentially, utility, there. Open source has shown the value of open source, so that Europe last week did an assessment of how valuable open source had been, and they reckon it's over 90 billion—with a B—dollars per year: that's the value to the European community of open source. Now, nobody has even started to do that for open access. If we were reusing open access in that same way, and we were saying that open access gives you this value, then you might convince people. But at

the moment, there's nothing that says that open is more valuable than anything else, in a way where you can measure it. Now, I believe as an act of faith that public investment in projects like this is almost by definition valuable. The best known example is the Human Genome Project, where they measured that the downstream value of \$4 billion of investment was about \$180 billion of value. Now, that I think is inflated, because it doesn't take into account all the other technology. That's the sort of study that we need to, say, if you publish an open access paper, it is worth this amount of downstream value in reused research, but you've got to measure the reuse of the research. And at the moment, nobody's interested in showing the value of open access. Nobody even measures the amount the papers are read! And if you don't do that, how can you show how valuable they are? People measure how often they're cited, but that's a ridiculous measure. It's only used because it gives authors glory, and because it's technically possible. So that would be one of the things I would put in place, is can we actually measure how valuable open access is, and give credit for value at that state?