



Beyond the Book®



Copyright, Open Access, Subscriptions and Permissions: What Editors Need to Know in the New Digital Publishing Environment

**Presented at Council of Science Editors Annual Meeting
Pittsburgh, May 4**

BERKLEY: My talk is in the public domain in the United States, so feel free to record it. But outside of the United States, it's not in the public domain. I didn't know if you guys know that but 17 USC 105 only prevents the establishment of copyright in the U.S., so if you want to ask our permission to publish this in Ukraine, you might have to. But it's OK. You can do it.

I'm actually from the Office of General Counsel of the Department of Health and Human Services and I'm a patent attorney and I pretend to be a copyright attorney when the need arises. And I was pulled into this PubMed Central issue a number of years ago when we were considered not to be so friendly because it was – when somebody talks about a mandate, that's not a nice word you want associated with your name, I guess.

But now, I want to present the friendly face of NIH to you and say a few words about the legal, about the public access policy.

And as John said, the PubMed Central essentially revolutionizes the accessibility and connectivity of publicly funded biomedical research and related scientific data. So it's hard to underestimate the value of PubMed Central as an actual research tool. It's almost like an MRI machine because for the first time, all government-funded biomedical research can be accessible from one place and the principle is – that you as taxpayers should consider is – that if the taxpayer funded the research, the taxpayer ought to be able to use the research for free.

And before PubMed Central, that wasn't true. PubMed has been online for quite a number of years and it's a citation website, so you go out there, you have a tumor in your head. And there was in fact just a story on *Marketplace*, if you listen to NPR, this week where they talked about this guy who had some very sad malady and he wanted to research it and he went out and put the search into PubMed and he pulled up all these wonderful articles and all he could get were the citations. He couldn't get to the full-text article. He would click on them and he realized that he had to pay 40 bucks for each one.

So that's what PubMed Central is designed to address. He's a taxpayer who contributed to the \$29 billion that NIH annually funds for biomedical research



Beyond the Book®



throughout the world and \$29 billion comes into Bethesda and we only keep about a billion. So almost all the \$29 billion for the – NIH is the largest biomedical government funder in the world so it's hard to talk to a biomedical investigator who hasn't received NIH funding and so the question that you always have to keep in mind – and it's hard to argue against – is if we pay for it, why shouldn't the public be able to access it freely?

John was talking about the business model, the historic business model for the publishers. Of course, the publishers weren't – and I don't want to say anything provocative because as John said, that fight, I think, is largely over.

But there are efforts out there in the political arena. There's something called a fair copyright act, law, that's being proposed by one of our representatives who is trying to essentially overturn the PubMed Central model, which I think would be a shame, but I think I'm biased about it a little bit.

The public access policy is now mandatory. It's the law. We tried at NIH to get the contribution of our grantees, their research papers, into PubMed Central in a voluntary way, and we got about a 20 to 30 percent compliance over a couple of years, so it didn't work.

So into a 2007 appropriation law went this very simple paragraph and it told the director of NIH – by the way, this is a paraphrase because it was a little bit longer than that – that if you get money from NIH, you have to deposit your final peer-reviewed manuscript into PubMed Central. Now, that's not the publisher's finished manuscript, of course. It's not the PDF. And you have to do it within 12 months of the actual publication date.

So this is the inductance that I think John was talking about in his paper. There's a lag time, which we think is quite generous, if we do say so ourselves, before you have to get the paper into PubMed Central where PubMed Central tags and hyperlinks all of the important concepts, the references. If there's a reference to a genomics database or if there's a reference to a protein structure database, which the National Library of Medicine will have available, then you just click on that hot link and it'll take you right to that structural database or genomic database, so it's an incredibly powerful research tool that is only possible because – as you guys had noted before – now we're talking about digital information and we're not just scanning pieces of paper into a PDF form and plopping them up on a website. You can hit every word in there with your mouse and sit at your desk and do research.



Beyond the Book®



This concept was actually introduced by I think Harold Varmus, who was one of the former NIH directors and David Lipman at the National Library of Medicine, who some of you may know.

I used a lot of red here. The bottom sentence is interesting. That was sort of a political, I understand because I wasn't involved in it. They keep the lawyers out of the political stuff. They only tell them later what was going on.

But it said that provided that the NIH shall implement the public access policy in a manner consistent with copyright law. And so that for a copyright attorney, one who pretends to be a copyright attorney or plays one on TV anyway, the question is, how could this possibly conflict with copyright law since the Copyright Act is not being amended?

There have been a lot of creative arguments out there that have been advanced to suggest that what NIH is doing does conflict with copyright law, but you'll see in a second – I really only need two slides to make my point, and that is that – you can ignore the rest of the 22 that I have up here. We'll go through them real quick because I want to give you time to ask questions about these guys who have much more interesting things to say.

But essentially, all we're doing is reserving the right we always had at NIH to use our grantee's manuscripts in any way we wanted to. There. I said it.

(laughter)

BERKLEY: Because in our regulations, 45 CFR 74.36, we reserve the right to use the manuscript in any way we want to, but we've never asserted it.

So what we do – let's see if I can – oh, here's a couple of definitions. Let me just –

PubMed Central, not to be confused with PubMed. PubMed is the citation, obviously, website that you go to when you first want to find anything on your subject and then you can go to PubMed Central and find the – what happened to that? Oh, that's interesting.

And a final peer-reviewed manuscript, that's different from the final published article, which of course the publisher is very jealous of. So any final manuscript that is peer reviewed is what public access policy applies to, so we're not talking about books or abstracts or anything like that.



Beyond the Book®



But here's what I was talking about. The HSS grant regulations reserve this royalty-free, nonexclusive license in the grantee's work for federal purposes. But nevertheless, we ask our grantees, who is essentially anybody who's working in biomedicine in a university in the U.S. and elsewhere in the world, to reserve the right to deposit with PubMed Central in their publishing agreement.

So that caused a little bit of work on the parts of all the universities and their tech transfer offices, for example, to make sure that in the publishing agreements that they signed between Elsevier and Nature and Science, etc., they explicitly reserve the right to deposit with PubMed Central, and we put that in our grant agreements to them.

So what happens if they don't do that? Well, then we talk to them about whether they're going to be funded the next year.

Because Congress told us to do this, we had to do it and we tell our grantees that they must reserve this right notwithstanding our right, we think, to put these manuscripts ourselves up on PubMed Central with all of the tagging, the metatagging, the hyperlinking, etc. that make PubMed Central such an important research tool.

And this just lists the ways that the publishers have arranged with the smart people at PubMed Central to go through the mechanics of getting the manuscript up on PubMed Central.

Method A is much the preferred one because the author has nothing to do with it. They take themselves out of the equation. For example, Elsevier and the people at PubMed Central, Neil Thacker, for example, who's going to speak later on at 3:00. If you're interested in this at all, you should see his talk because he's got a lot of cool slides that give you that update, the new compliance statistics with respect to PubMed Central, etc.

Method A is perfectly transparent to the author, and that's what we want to do.

Who does this apply to? It's all of our grantees and it's the government employees who are NIH investigators in Bethesda and elsewhere. We have 18,000 Nobel Prize-level – we like to think – investigators who work at Bethesda who have to comply with this requirement as well.

Method A has now been agreed to by over 550 journals, which Neil can tell you later on and it's growing. So feel free to encourage your favorite journal to join the growing crowd that are doing this.



Beyond the Book®



Now, how do we comply? This is a mechanical thing. I personally am the guy who had to approve or review hundreds of publishing agreements. I have a real interesting job. Yesterday I was working on pandemic flu and we were looking for patents that were related to antivirals because there's something called the swine flu that's spreading across the world and we were trying to figure out how we were going to break some patents that are related to that.

Is this going out? Oh, yeah. This is being recorded.

But anyway, this is my day job and my afternoon job is reviewing publishing agreements. You can imagine, they all look alike after a while and they all include provisions that the government can't agree to because the government is special. The government can't indemnify people. It can't provide certain warranties, for example. And we have this – everything that we do is in the public domain and so if you read – I don't know how many of you actually read a publishing agreement that your publisher puts out. Yes, you have.

So half of it we can't agree to, so we got tired of it and we took this opportunity to write our own what's called a cover sheet. We gave it a special name. I came up with that.

(laughter)

BERKLEY: That's my creative side. Before I went into the law, I was a physicist and that has nothing to do with it.

(laughter)

BERKLEY: But we developed this cover sheet that has what we call core provisions, which would have been in red, but there was too much red, so core provisions. And they essentially say, no, we're not going to indemnify you, and no, we're not going to warranty you, but we're going to represent certain things to you that we're happy to represent.

So the indemnification, the thing that prevents us from doing this – I'm really going to bore you now and that's why I'm going to go quick. This is compelled by the Antideficiency Act. In other words, we can't spend money that we don't have that Congress hasn't allowed us to spend. And if we hold you harmless, that means that there's a no-ceiling obligation on our part to warranty you.

We also put in our standard copyright clause in this cover sheet. I didn't explain what's going on.



Beyond the Book[®]



The publisher's agreement sends us to one of our – sends the – our investigator sends in their manuscript. The publisher sends the publishing agreement to the investigator. She signs it and sends it back, usually because we don't know anything about it. But some of them – two or three of them – send them to their office of general counsel and when they do that, we staple a cover sheet on it and we have them sign the cover sheet and send it back to the publisher.

And we've actually institutionalized this with Elsevier and lots of other publishers where they've just – they got tired of listening to us or arguing with us and they now put this thing up on their website and they say, OK, NIH. Click. And we're very happy to do this. Click license. You're going to agree to this copyright provision, which just says that –

And this is very generous. It allows the publisher, Chris, to go out and secure rights outside the United States. This is more generous than their own publisher's agreement, for example, because some publishers think that all government works are in the public domain everywhere, and that, of course, isn't true. So we help them out with that.

So we have this copyright provision. We have the PubMed Central license, which explicitly – so everybody knows what's going on – allows us to metatag and deposit the manuscript – the final peer-reviewed manuscript in PubMed Central. We don't put in the PDF unless they allow us to, unless they tell us to do that, but a lot of them do. A lot of them let us put in the –

By the way, there's a trademark issue with respect to that. You've got the math staff and so on and so forth, so we don't like to do that. And that's something that I'm not sure the publishers realize. But there's a tradadress issue with respect to just throwing a PDF up there.

These are the representations that we're happy to make that NIH is the owner, the work is original and all these other usual representations that are included on our cover sheet. And we agree to other reasonable publisher's policies.

And like I said, we sign that cover sheet. We don't sign the publisher's agreement. We staple it and we have our investigator send it back to the publisher. And that's the way we effect our compliance with the public access bill, public access law at NIH, and FDC and CDC, by the way.

So these are some milestones I just wanted to talk about for a second. It took eight or nine years here from the point where PubMed Central debuted through this voluntary regime where we asked our grantees to contribute their manuscripts to PubMed Central, and there was all kinds – there were at least two bills that



were proposed that would require us to do this, and finally when it was clear to Congress that the compliance was so low nothing was going to happen, in 2007, the policy was made mandatory and in April 2008, we put it into effect.

Here again is a timeline, the final report to Congress, which said, hey, this isn't working in 2006. We had all kinds of public meetings, literally hundreds of public meetings with publishers and other interested so-called stakeholders, which is a word I think should be deleted from the English language. But there it is.

And in 2005, you can see the voluntary policy took effect and had no effect.

I also want to show you how nice we are in response to public comment and after NIH consideration. We first proposed that there be a six-month embargo, so-called embargo period between the time when the paper was published by the publisher and when it got put up on PubMed Central. But we're such nice guys we stretched it out to 12 months.

So we have essentially the nicest – the less stringent – is that a word? – public access policy among all of the funders that mandate these public access policies. For example, the Wellcome Trust, they're bad guys and Howard Hughes, etc., six months. But we give you 12 months to make a profit and then we put these manuscripts up on PubMed Central.

MOORE: (inaudible).

BERKLEY: Two seconds?

MOORE: Two seconds.

BERKLEY: OK. Three seconds. These are some slides. I'm a copyright attorney, so I can steal slides. You didn't mention that. You're immune. You can just steal slides from whoever you want. These are from National Library of Medicine.

You can see that there's a disorganized science data and this is what pub – I love this picture because these things are literally at the click of your mouse at the National Library of Medicine. You've got the Entrez Protein, the Entrez Structure and you've got the Entrez Nucleotide and Entrez Genome. All of these databases are available at the National Library of Medicine in the little building that sits over on the NIH campus and essentially run by David Lipman, and Betsy Humphries is the director of the National Library of Medicine, and they're all available in one place.



Beyond the Book®



This is perfectly consistent with the wonderful talk that John gave about how important and the interconnectivity of these resources are, so this is one-stop shopping for – and this is a view of what these things look like if you go to your computer and pull them up, all the way from the paper – and here's the chemical structure and here's a 3-D view of the chemical and protein. You can do this this afternoon if you pick up any paper from PubMed and go to PubMed Central.

And I'm done. Thanks.

END OF PROGRAM