

I feel the dread hand of copyright hanging Mordor-like over the whole area of scholarly publishing. I heard to my horror in PennState that one University had embargoed all its theses in case they violated copyright. So I tested this in my talk and asked “are there repositories that embargo all their content for fear of copyright?” and got a few nodding heads. So I am taking this as fact, and asking:

Why is no-one except me angry about the way that copyright (or exaggerated fear of it) is stifling electronic innovation in academia?

So for example, I asked one speaker who proudly talked about their thesis aggregator “how many of your theses are available under CC-BY or equivalent and can I download all of these and data mine them?” Apparently they weren't his theses – he just aggregated metadata and I would have to approach the author of every single thesis to find out what the permissions are.

The whole meeting seems to be asleep about the urgency to liberate these theses into the digital Open. I am depressed. I don't think it's going to happen any time soon. Providing access to single humans for single views on single theses is all that seems available. Maybe some commercial company does some full-text indexing somewhere, but that's no use to me. We could process 10,000 theses tomorrow and extract the chemistry. The recall won't be great but it will still be thousands of results. But there is no way that anyone seems interested in this. Theses are precious jewels which require a priesthood to access – not for my robots.

So I am angry. It is not just the fault of the libraries – faculty, especially senior faculty, bear half the blame. But we are sitting on a goldmine of scientific information in academic theses and we are deterred from using them by copyright FUD. There is an implicit assumption that copyright is one of the god-given commandments – it seems almost revered here.

So let's abandon copyright in science. What does it gain us? Almost nothing, unless you author a successful textbook. Nowhere else is copyright the slightest use to a scientist and it stands in their way at every step. There are faculty who can't use their own research work in teaching. There are libraries which can't let the world see their theses. There are librarians who spend their time negotiating deals with publishers so they can access their own work. Even Laputa could not have designed the bizarre copyright system.

I stress that this is for SCIENCE. I agree that if you are working in creative arts you may wish to protect your work. But scientists don't. And they are being held back by the assumptions that apply to creative works.

So I would urge that science declares war on copyright. It is fiendishly complex, and wastes vast amounts of time. I know that, theoretically, we can't as copyright is a legal thing. But since the British Library is trying to get the UK government (what is left of it) to change the law, why don't we just assume that the spirit of the law and the future letter is that copyright in science should be effectively irrelevant.

SO AS A FIRST STEP LET'S JUST PUBLISH ALL OUR **SCIENCE** THESES OPENLY AND ALLOW UNRESTRICTED DOWNLOADING AND RE-USE?

I can't see any reason why not. Any publisher who sues a university for publishing work done and that University and for no financial profit has to show that this is not fair use. OK, it's then down to the lawyers, but I suspect that few publishers will relish suing large prestigious institutions.

And in that way the faculty-library complex can regain some of the sense of urgency demanded by students and researchers.