ARL

A BIMONTHLY REPORT ON RESEARCH LIBRARY ISSUES AND ACTIONS FROM ARL, CNI, AND SPARC

241 August 2005

Cornell Library/Press Collaboration 5
NLM Permanence Ratings for Web Documents 6
Global Perspective on Scholarly Communication 8

Expanding the Public Domain

by James Boyle, William Neal Reynolds Professor of Law and Faculty Co-Director of the Center for the Study of the Public Domain, Duke Law School

Remarks presented at the Association of Research Libraries 146th Membership Meeting, May 26, 2005.

During the last 10 years there has been a frenzied and intensive debate about the desirable limits of intellectual property policy. For much of that time, if you said amazingly bland and banal things like: we should have balance, or it's important to think about the inputs for creativity as well as protecting outputs, or we should not commoditize facts and ideas, you could be labeled as a communist, an anarchist, or, rather confusingly, both. So, what I think I'm going to do is produce a stunningly banal set of ideas. First, I'll discuss what we mean when we talk about the "public domain." Second, I'll explore a set of ideas about recent expansions in intellectual property policy. And third, I'll talk about public domain initiatives that we can undertake within private institutions.

A Richer Understanding of the Public Domain

Although my topic is the public domain, I want to stress something that I would like you to remember throughout my talk: the public domain is fed by intellectual property. It is not merely the opposite of intellectual property. The way to have more things in the public domain is not always to get rid of intellectual property. For example, if we were to get rid of the patent system, many inventions would end up covered by trade secret law and we might never get access to them. Intellectual property has an important role, and that is the premise of everything that I'm going to say. Preserving the balance between intellectual property and the public domain is not an attack on intellectual property; rather, it's about preserving a living ecosystem between intellectual property and the public domain.

First, I want to pull back and address a few definitional issues, which I think need to be clarified in order to talk about the "public domain." We all have a rich and complex understanding of "property." We understand that there are lots of things you can do with property: giving it, sharing it. We know that we can rent an apartment, and that it's still owned by someone else, but we nevertheless have rights over it. We are, in fact, immersed in a culture of property, and it's constantly maintained, constantly named, constantly refined, all the way from "that's mine, you can't have it" on the playground through signing your first college lease to your mortgage and your retirement plan.

We also live in a world of the public domain—the realm of material that is not covered by intellectual property, and is accessible for all to use. But it is not as well named, and not as well understood. When we talk about the public domain, for example, are we talking only about complete works that are completely free, such as Shakespeare plays and Mozart symphonies? These are in the public domain in the sense that the copyright has expired, and you can do whatever you want with them. You can make a new version of them, abridge them, base a new work on them. We could also be talking about things which are not, and never could have been, the subject of intellectual property, such as $E=mc^2$ or two times two equals four. Some people would include both the works of Shakespeare and Mozart, and the world of ideas and facts, in the public domain. Others might include the limitations and privileges within intellectual property as part of the public domain.

So, for example, my ability to criticize a book would mean that, for that particular use, it is in the public domain, or my ability to parody a song would mean the parody-able aspect of a song is in the public domain.

You might say, well, "what's in a word?" The point is that we need to develop as rich, complex, and varied a notion about the "public domain" as we have about "property." When we talk about the public domain, we need to ask, what is it that I want here? Is my claim that this whole thing should never be subject to rights? That this thing could legitimately be subject to rights, but at some point they should actually expire? That a particular use of an aspect of the thing should not be controllable? Often these ideas get conflated. It's not that we need a precise definition of what is in the public domain. Instead, we need a better analytic process of definition. We should ask, what is our purpose here? What is the mental work we are trying to get done? What definition will get us there? And then make clear the definition we have adopted, and the tasks it seeks to accomplish.

We also need a richer understanding of the notions of the "public domain" as opposed to the "commons." Until very recently, a lot of people would use these terms more or less interchangeably. But it's not clear that they're actually the same thing. Is open-source software in the public domain? No, not at all. It's strongly protected by copyright—that's, in fact, how open source can be maintained. It's because of copyright that I can say, "The terms of this general public license are attached to your use of this software." You may copy it freely, but if you wish to change it, you must add your new innovation to the 'commons,' not the public domain. You must make it available under this same license, which lets the future user, who also will add to the commons, use your innovation."

Now the point is, that's not the public domain. It focuses on many of the things that the library community cares about—access issues, sometimes price issues, sequential innovation issues—but it is built on the back of intellectual property rights.

In fact, there are currently developments in the scientific community, which some you may be aware of, where there is going to be a hard tactical choice along this front. For example, we're right at the beginning of "synthetic biology"—creating entirely new molecules, entirely new biological entities, using, effectively, DNA as a programming code the way someone might use C++. Most of the sequences are probably not copyrightable. But some of the scientists who passionately want this stuff to be openly available wish that they were. Why? Because they want to attach a General Public License-like condition that says, if you want to use my building block, my enabling technology, then you have to add your innovation to the commons. They're saying, this must be "property," so it can be free.

So when we're working on these types of issues, it's important to be clear about definitions. And to be honest, I think right now we have a better understanding of the public domain than someone contemplating a society with property rules for the first time—say an anthropologist who'd come to us from outer space and had never heard of this weird idea of property. We have some familiarity with it. We've used it, we've been embedded in it. But we simply don't have the richness and complexity, either of social uses, so that your kid would know what the public domain is all about, or even of philosophical, theoretical, and legal uses, so that we'd have a precise vocabulary and set of tools that would allow us to agree on particular definitions and goals, and get to work.

Expansions of Intellectual Property Rights

To move on to my second area of focus, as you know, intellectual property rights have expanded dramatically in recent years. They've been expanding in every field of intellectual property, and in every dimension: length, extent of penalties, scope, subject matter. The copyright term has been extended by 20 years, and copyright penalties have become more severe. New rights protect not only the copyrighted work, but the digital fence in which the copyright owner wraps it. Patent law covers things we never used to cover—gene sequences, business methods. In the European Union, database protection now covers unoriginal compilations of facts.

What arguments have been used to justify this expansion? One is what I call the "Internet threat" argument, which assumes that, as copying becomes cheaper, intellectual property protection must increase. The argument goes like this. If you have a monk with a manuscript in his scriptorium, copying a book out by hand, you don't need intellectual property protection, you just need to control a single copy of the manuscript. Copying would take months. Then along comes Gutenberg, and people can copy things quickly and more cheaply. We now have what economists call a public-goods problem, because we have a book that is non-rival and non-excludable. And now we see, for the first time, the need for intellectual property protection (which actually, somewhat confusingly, doesn't arrive for over 200 years after Gutenberg). And as we go on, every time the copying costs fall, the need for intellectual property protection goes up. Zero intellectual property protection at the monk. The Statute of Anne by Gutenberg (except that it's 200 years out of date), and as we go through the photocopier and the VCR, towards the world of Napster and Grokster, we need, effectively, perfect control. Because the Internet lowers the cost of copying to zero, and we have an infinitely leaky system.

Now, this is not a dumb argument, but it is wrong. It's not dumb in that there is a real problem. The Internet

does lower the cost of copying, so it will magnify the amount of illicit copying. But it will also magnify the amount of licit copying. And it expands the size of the market, makes it easier for you to distribute things, lowers your advertising costs. On balance, are intellectual property holders better off or worse off? Well, even economists don't think that you can decide that in the abstract. They say you actually need evidence, right?

Here's another remarkable thing about intellectual property policy over the last 10 or 15 years: it is almost evidence-free. People criticize the FDA about Vioxx. But if we were doing FDA drug approvals the way we approved intellectual property expansions, this is how the process would go. The drug company would say, "This is my friend. He took the pill and he feels better." Or sometimes even, "This is my friend, he needs to take a pill and he thinks it will make him better." And then they would offer a model about as complicated as a picture of the person with a mouth and the pill in their stomach and say, "See?" That's about as data-intensive as things have been.

Preserving the balance between intellectual property and the public domain is not an attack on intellectual property; rather, it's about preserving a living ecosystem between intellectual property and the public domain.

What if we had a test case where two regions adopted different intellectual-property policies, and we actually had evidence showing how these policies worked? Well, we actually do have such a case—in the area of database protection. In Europe, there is strong database protection under both copyrights and sui generis database rights. Many European governments also claim some kind of copyright over databases. And there is the idea that institutions, such as the Ordnance Survey or the weather companies, should recover their costs by charging users. The US tradition is totally different. In the US, there are no rights over data or unoriginal compilations of data. Any text produced by the government is free from copyright and passes immediately into the public domain. As for government-funded data, it is produced and distributed to the public with the idea, remarkably, that taxpayers have already paid for this, and shouldn't have to pay for it again.

Now, we actually have some good evidence about the effects of these different approaches. The United States database industry is considerably larger and more thriving, and has higher rates of return, than the European database industry. In fact, at the moment when Europe introduced sui generis database rights, there was a short one-time spike as database producers raced into the market, but then growth rates returned to previous levels, and many companies left the market.

And when did Reed Elsevier and Thomson enter the legal database market in the United States? It was after a case called *Feist*, which said that facts, and unoriginal compilations of facts, were uncopyrightable. That is to say, European companies chose to come into a classically public information field in the United States after they had found out, for sure, that they could get no copyright in unoriginal databases. Yet, even without database rights, they're getting high rates of return. So, we have evidence showing that less protection has been better for innovation than more protection. But you could spend days listening to arguments about database rights, and you'd never hear these facts mentioned.

Additional evidence shows that publicly generated data turns out to spur more economic activity if provided at marginal cost—close to zero—than if it is provided in order to recoup its cost of production. Europe puts into public weather-data generation about half of what we do in the US, and it gets a nice return of about a six- to eightfold boost in production. The US puts in twice as much, and gets back a 39-fold increase in production. Why? The information is initially provided for free, but a massive secondary industry—the private weather industry—takes the publicly funded data and adds value to it. They employ more people, pay more taxes, and are an enormous portion of the economy. Keeping public information free just works better. It's not even a close call, as with Vioxx and aspirin.

So I have discussed two themes: first, the Internet threat argument, which says that as the cost of copying goes down, we automatically need more protection. And second, the idea that we can make intellectual property policy without having any evidence. This idea is bizarre: other government subsidies are rigorously assessed in order to figure out whether they're worth it, but here the government is handing out heaping slices of monopoly rent in the form of intellectual property rights, without empirical evidence that these rights are necessary, or that they will do more good than harm. My points are: lowering copying costs brings benefits, as well as costs. And we need evidence before we make policy. Banal and boring, right? It is in that context that I think we need to look at the range of intellectual property expansions that have been put forward, because in many cases we'll find that underlying them is the Internet threat assumption, and that they were passed without evidence. The call for evidence-based policy is one that we can really wrap our arms around: it's a positive proposal, and it's very hard to object to it.

Initiatives for the Public Domain

Let me turn now to private initiatives—practical things that we can do, akin to the kinds of things that the environmental movement did with its "think global, act local" initiatives.

Identification and Labeling

One idea is to actually identify public domain materials as such, in order to make people aware that the public domain is there and that they are using it. We're already digitizing things and making them available online but, for many people, the legal conditions under which they get this material are completely opaque. So we ought to tell people why and how they got access to this material, because if they realize that "this poem is here because the copyright term's expired, and I'm glad about that because I can do something really useful with it"—then they can learn to value the public domain and what they're getting out of it.

Fuelling Demand

Along these lines, we also need to think more about the demand side of the public domain in general. We've thought a lot about the supply side—how to ensure availability and access. But what about the demand side? One of the things we found with Creative Commons is that an initial expenditure of time, effort, and money by people who cared could galvanize entire communities around public domain resources, or, in this case, resources made available under Creative Commons licenses. So, for example, we got David Byrne, the Beastie Boys, and other worthy musicians to put some of their music out under Creative Commons sampling licenses, which allow you to take snippets of a song, remix it, make your own song, and even, in some cases, sell it. Those musicians thought that would actually be great. So we ran a competition for all of the remixers out there who wanted to create something. They really got into it, and now there is a huge group of people, stretching well beyond those who were involved in the contest, who realize that there's all this material out there that is free for them to use and to remix. We can stimulate the demand side of the public domain, and of the commons, by initiatives and educational approaches that get people to use public domain material.

Education

At the university level, copyright education campaigns need to emphasize that copyright is a positive thing that people can use, rather than just something that gets in their way. These campaigns need to teach faculty and students about copyright, rather than merely telling them such untruths as there's no such thing as fair use, and so forth. Copyright education is being done. It should be, it is important, but it is being done badly and inaccurately. It is being done in a way that is entirely foreign to the critical intellectual tradition of the university. Librarians are ideally placed to develop strong, national, well designed, and visually attractive copyright education campaigns for schools and universities. Right now, copyright education consists of saying, "Don't download songs illicitly, and if you do, turn off the upload feature." That's it? Sometimes the claim is that "no copyrighted material may be used without consent." Really? What about fair use?

We need to be more serious about teaching our colleagues, our administrators, and our students about both sides of copyright. If we teach them that copyright serves valuable social goals, they might actually respect it more. If we explain the careful package of balances, the limitations, the boundaries of fair use, the things that can practically be done under existing laws, then we will begin to approach what copyright education should be all about. Copyright educators need to say, "These are the things you can do with the rights you have over your article, your materials. Here is what fair use allows. This is what you may not do, and this is what to think about. Here are the author agreements you sign. Do you want to? You have the right to self-archive. Are you doing it? Here's this resource called DSpace, and so on." Balanced copyright education is important.

Conclusion

My goal here has been to offer a theory, and a practice, of the public domain. The theory and practice come with a change in attitude. It's time to think about expanding the public domain, not just defending or salvaging it. Some of the decisions that have already been made were unfortunate. There was no need to extend the copyright terms, in my view. It was not economically justified, it didn't harmonize the law, and we've locked up 20 years of culture for no good reason. But the good news is, I don't think that the term extension would pass today. What we have to do now is to think of all of the ways in which we can use the wonderful technology that is available to us, and build a public domain that people can get access to practically, but also a public domain they are aware of. Because if people have a sense of this world of available, accessible information, and understand what they can do with it, not just as passive consumers, but as people who can actually use and build on it, then we will solve the theoretical problem I started out with. We will have our rich and complex idea of public domain because we will all be living it every day.