

Symposium: Collective Management of Copyright: Solution or Sacrifice?

A Digital Public Library of America?: Collective Management's Implications for Privacy, Private Use and Fair Use

John Palfrey*

The purpose of this essay is to examine collective licensing operations, with particular emphasis on some of the consequences for privacy, private use and fair use. I examine these issues in the context of a specific idea that might draw, at least in part, on a collective management approach: an emergent proposal to create a Digital Public Library of America ("DPLA"), which I think is an important undertaking.

In the digital environment, media are not all the same. It makes sense analytically to disaggregate, for instance, music from movies and from books in various ways.¹ I start here with books, though one might extend this argument, with a few twists, to cover works in other media formats. At Harvard Law School, I am responsible for our library. But I do not make this argument from the perspective of libraries per se; we should take up this issue with the broad public interest in mind, rather than by adopting any single perspective. Most of the time, this discussion is dominated by voices from a specific industry or a particular consumer-oriented group. In terms of further disclosure, I take, as a prior normative commitment, the position that it is important to seize what is and can be great about the digital age for the purpose of spreading information through networks. I start with the general posture that it is a good thing for us to try to make more information, especially that which is published in books, available broadly to many people.

Given all those prior normative commitments, one might infer that I would argue in favor of a particular, well known collective licensing arrangement for books, the proposed Google Book settlement. But I hesitate to take this last

* Henry N. Ess III Professor of Law and Vice Dean for Library and Information Services, Harvard Law School. In these roles at Harvard Law School, Prof. Palfrey is director of the school's library and cochair of the IT committee. He is also a faculty codirector of the Berkman Center for Internet & Society. His research and teaching focus on Internet law, intellectual property and the potential of new technologies to strengthen democracies locally and around the world. Prof. Palfrey holds degrees from Harvard College, the University of Cambridge and Harvard Law School. He is also the chairman of the Steering Committee that is working to establish a Digital Public Library of America, and has received a grant from the Alfred P. Sloan Foundation to begin the planning process.

1. See Scott Martin, *Alternatives to Collective Management: DRMs and Other Business/Technology Options*, 34 COLUM. J.L. & ARTS 675 (2011).

analytical step. I have concerns, as Pam Samuelson does, about the Google Book settlement approach, which is the leading effort in this broad area.² I think that we might be able to do better if we take a new approach to achieving similar ends, from a public-interested perspective, which might build upon or amend in various ways the Google approach to this problem and opportunity.

The history behind the idea of a DPLA, roughly, started in the 1990s. For more than fifteen years, librarians and others have been talking about collaborating on a major digital project to make more of our cultural and scientific heritage available than ever before. There have been many important efforts toward that goal in the last fifteen years or so. But no single idea or approach has gained much traction. In the fall of 2010, a group of us—including Pam Samuelson, Jane Ginsburg, Harvard's Robert Darnton, the Sloan Foundation's Doron Weber and others—met to think about whether it was possible to resuscitate this big idea. Could we pull together a very big tent full of people, who might come together around a joint project to create such a digital public library in the United States?

Since that meeting, there has been a modest groundswell of interest and support, over the past several months, for working towards the creation of a Digital Public Library of America. The initial idea is captured in a single sentence, to which the early discussants agreed: “an open, distributed network of comprehensive online resources that would draw on the nation's living heritage from libraries, universities, archives, and museums in order to educate, inform and empower everyone in the current and future generations.”³

The simplest instantiation of this notion, for the purposes of demonstrating the idea's potential, would be to start with things that are in the public domain. These materials could be made more broadly available than they are today for the public to access, from anywhere, through a relatively simple interface. It might even mean thinking about this problem at the protocol layer, with a proposed interface—just one of many possible interfaces—created to spark the imagination. Put another way, we might think of this DPLA merely at the level of technical interoperability or standards that make it possible for people to access these materials. And then it might get more complex with other kinds of works—ones that would need to be in something like a collective management organization approach or otherwise licensed into this system. The notion would not, to be clear, be to cut authors or publishers out of such a digital public library. We would presumably think about materials differently, depending upon their copyright status: we would think about in-copyright, in-print books as different from materials that are out of print and in copyright and orphans and public domain works and so forth. And getting somewhat broader, one might think about this in a cultural heritage framework, not just in the context of books in particular.

2. See Pamela Samuelson, *Legislative Alternatives to the Google Book Settlement*, 34 COLUM. J.L. & ARTS 529, 533 (2011).

3. *Sign On—Digital Library of America Project*, DIGITAL PUB. LIBR. AM., http://cyber.law.harvard.edu/dpla/Sign_On (last visited Apr. 13, 2011); see also DIGITAL PUB. LIBR. AM., http://cyber.law.harvard.edu/dpla/Main_Page (last visited April 13, 2011).

Despite its relatively long history as an attractive idea, the DPLA itself, circa 2011, is nascent. We are just at the beginning, at the concept stage. What I think is important about this particular frame is you have the ability to come at it from a design perspective. You can say, “What are the things that we are seeking to ensure happen, and what are the things on the other side we want to ensure do not happen?” Since I have been assigned to look at fair use, private use and privacy, let us examine these concerns in the context of a proposed DPLA.

Why, one might ask to begin with, would one go down the road of trying to create a Digital Public Library of America? The point is that one would want to make more works—books for starters, but other media to be sure over time—accessible to more people through this digital public library regime. One could imagine, in fact, that a DPLA might be revenue enhancing for those who write and those who publish. But it also certainly would be democratically enhancing in terms of benefits for access. I would underscore many of the things Pam Samuelson said earlier about access for disabled people in particular, and other things you can do in a digital format that might be benefits as compared to the current system.⁴

We have not decided on an approach for the DPLA. One could easily imagine that a way to do this would be through a collective management organization (“CMO”). One could imagine that the way to accomplish this—knowing that there are a lot of legal impediments, particularly in copyright, between here and setting up a digital public library mechanism—would be an “opt in” collective licensing approach. One could also pursue a strategy involving a mandatory license.

Once we have decided upon an opt in license and a collective management approach as the basic assumption for works still in copyright, it is easy to see that law reform of various sorts could help a great deal. Orphan works legislation is the most important possible single change. Better yet, though perhaps wishful thinking, the Congress could pass a whole package of reforms that would make this DPLA more likely to become a reality in the United States. Pam Samuelson disaggregates two things when she talks about the approach we ought to take: one is an orphan works approach and the second is her complete package of reforms.⁵ I am hoping we might get a complete package, but you could imagine some subset of that, and you could imagine the establishment of this organization.

A CMO, in the context of a DPLA, would help to manage the licensing of rights from authors and publishers to the public at large, via an intermediary organization. It might work formally through libraries that then pay those licenses. It might be more broadly organized, outside of the library system. It might have three sets of rules for different kinds of content; it might have many more such sets of rules. The basic idea, though, would be to put the rights together in this format.

Let us turn to the problems that may arise and explore these three possible concerns. With fair use, private use and privacy, for the sake of simplicity, I am going to lump together fair use and private use. These issues are analytically

4. See Samuelson, *supra* note 2.

5. See *id.*

distinct in some ways, but for the purposes of keeping this explication short, place those two in one bucket, and in the other bucket place privacy.

None of these three issues is necessarily a concern with collective licensing on its face. There is no reason that collective licensing has to be seen as a “bad thing” for the purpose of fair use, private use or privacy. By saying that one is aggregating a bundle of rights in one place and then enabling somebody to license those rights in various ways does not mean, inherently, that one is violating individual interests in any of these ways. But fair use and certain private uses—and I am for the moment calling these “exceptions and limitations” to copyright—might or might not be affected on the basis of how you implement the collective management organization.

These issues ultimately come down to a critically important design question for any CMO, and certainly for a DPLA. There are important design choices that we make, in implementing a CMO, that will have variable effects in these three issue areas. One of the reasons I have concerns about the Google Book settlement, even as amended, is the notion that a court that approves such a regime would be creating an enormous amount of private law on top of the public law, and in a manner that favors a few specific private parties.⁶ I fear that the Google Book settlement would undercut some things that we care about in the background public law. There is a corresponding advantage at this moment in designing something like a digital public library—where in fact one could anticipate these problems at the outset.

This is where I link up to the discussion of digital rights management.⁷ One of the ways in which a CMO may lead to concerns is if we have to rely on a strong form of digital rights management in order to ensure that the rights are all properly vindicated in this system. It seems logical that one would need to go this route. In lending something for a certain period of time through a CMO, one might have to say to an end user, “Ms. Library User, you have access to this book for a certain period of time, after which it will go poof.” One might need a digital rights management system to ensure that a certain library does not lend out the same book too many times. Alternatively, if the point is that one can lend a book out as many times as one wants and for as long as one wants, one will have to count the uses in order to compensate the copyright holders adequately.

One might end up implementing a system in which one restricts some uses of a work that might otherwise be deemed to be fair uses. There are many reasons why this might be the case. It is well-known that the combination of DRM-type measures with other aspects of U.S. law, such as § 1201 of the DMCA, which restricts certain circumventions of digital rights management—might restrict some uses that we deem to be socially beneficial, and which we privilege as fair under the Copyright Act.⁸ I do not mean to say that it is inherently the case that we would

6. See Amended Settlement Agreement, *Authors Guild, Inc. v. Google, Inc.*, 93 U.S.P.Q.2d 1159 (S.D.N.Y. 2009) (No. 05 CV 8136 (DC)).

7. See Martin, *supra* note 1 (discussing digital rights management).

8. 17 U.S.C. § 1201 (2006).

undercut public interests like fair use if we were to implement a DPLA. Rather, I want to emphasize that we need to analyze the intersections of this system with other aspects of U.S. law and other background rights that we care about. This is the first big concern.

The second concern is privacy. This issue is plainly distinct from the first, analytically, and it is a little more clear cut. There is almost certainly an effect on individual privacy, most likely for the worse, in going down this road. Again, it is less about the collective management scheme itself, or the way in which we manage these rights into and out of a certain central body, and more about the fact that this would be done in a digital environment. From the perspective of the end user, you might have to give up more information about yourself in order to use the digital public library lending process than if you went into a library in the ordinary way. If you wanted to walk into any public library in the country on a Saturday morning, you could browse through the stacks, reading as much as you liked, without anyone knowing who you were. Even if you were to check out materials, the number of hands that would be able to touch data about you and your history of check outs would be sharply limited. And it turns out that librarians are wonderfully fierce about protecting individual privacy.

In order to have a DPLA, one might imagine the need to authenticate users into the system and to track their usage in ways that are more extensive than in the traditional analog-world library model. One might imagine implementing an elaborate system of tracking the usage of an individual, relative to the books they interact with or “check out,” in the service of ensuring that the interests of authors or the publishers are vindicated in the end. The risk of going down this road might be that, from an individual perspective, there is a net decline in informational privacy.

Again, this privacy concern does not hinge on the nature of the CMO itself; it hinges upon the way in which the system is implemented. There are a couple of ways in which this might be so. In order to authenticate into a system, one would need to identify who a given user is to the system. Depending on how the payments are arranged to get back out to the authors or publishers, one might need to have certain access rights in various ways to certain kinds of material. Finally, one might need a mechanism for counting the usages and tracking them based on different variables.

My point is not that privacy concerns cannot be overcome; nor is it to say that a DPLA must have the forgoing elements; nor is it to say that one would not accept some additional privacy risk in order to enjoy the public benefits of a DPLA. I think of these concerns as rebuttable presumptions. There may be the possibility of ensuring that these measures are not so privacy restrictive; however, in the first instance one might say that a DPLA is likely to be net negative from a privacy perspective.

There are limiting factors here that cut in favor of personal informational privacy in the context of books and public libraries. The good news is that librarians are amazing when it comes to privacy, particularly in terms of protecting their patrons. It is this aspect of libraries—that they do not tell what people are

asking about—which lessens my worry about the privacy implications of potentially implementing some kind of a DRM-like wrapper in the digital public library zone. Privacy may not be such a big problem because there are people who would be implementing this scheme who care so deeply about this particular public interest or public right.

One further problem with privacy in the context of a potential DPLA is that I do not think one can see it only in this limited sense. One needs to see this potential privacy risk in the context of all these other privacy shifts that are occurring around us. We are giving away more and more information about ourselves all the time, sometimes voluntarily, sometimes involuntarily. In the digital age, our activities are more frequently and consistently monitored than ever before in human history. This trend appears to be continuing, and I think you do have to see a net decrease of informational privacy as part and parcel of other things that are going on.

Does it then make sense to not pursue a digital public library idea because it might be net negative for things like fair use and privacy? My view is: absolutely not. The key point is that we need to look at how we implement any such system—particularly if we are using a collective licensing approach—and we need to figure out how you design it to mitigate these potential harms. We need to ask the hard questions, at the outset, about how we design a DPLA, from the beginning, to vindicate these sometimes-competing sets of interests. The transformations that are possible at this moment in the digital era, in terms of making available much more information to many more people for these pro-democratic purposes—and in the process paying authors and publishers fairly for their works—trump the marginal risks. I would rather see us go down this road and work on broader privacy legislation than become bogged down with privacy solely in this area. We should deal with privacy in the context of social networks, and not worry so much that librarians are going to give away some marginal bit of information about us. And when it comes to copyright law, the process of designing and building a DPLA may well be a way to illuminate the need for reform, such as orphan works legislation. These prospective worries should not stop us from moving forward with CMOs and any DPLA models, but rather should function as important design considerations as we set about our work, in the broad public interest.