Session 3: To What Extent Should Libraries Be Permitted to Engage in Mass Digitization of Published Works, and for What Purposes?

Moderator: Karyn Temple Claggett, U.S. Copyright Office

Speakers:

Paul Aiken, Authors Guild*
Eric Harbeson, University of Colorado at Boulder**
Janice Pilch, Rutgers University***
Gloria Phares, Patterson Belknap Webb & Tyler LLP****

- * Paul Aiken, a 1985 graduate of Cornell Law School, has been the executive director of the Authors Guild since 1996. Mr. Aiken testified before the White House Task Force on Copyright and the Internet, participated in the Conference on Fair Use, and has testified before congressional committees on fair use, competitiveness in a digital age and the need for a small claims court for copyright infringement. His commentary on the publishing industry has been published in *Publishing Research Quarterly* and the *New York Times*.
- ** Eric Harbeson is the music special collections librarian at the University of Colorado at Boulder. He holds master's degrees in music from Cleveland State University and in Library and Information Science from the University of Illinois at Urbana-Champaign. He is chair of the Music Library Association's Legislation Committee, and has advocated on behalf of that organization on a variety of copyright issues and in a variety of fora. His research currently focuses on copyright ownership issues surrounding sound recordings, especially as they pertain to student recitals and other sound recordings produced in the daily business of educational institutions. Mr. Harbeson recently spoke on the subject at the International Federation of Library Association's (IFLA) annual conference in Helsinki, Finland, for which he was the inaugural recipient of the American Library Association's Robert L. Oakley Memorial Scholarship.
- *** Janice T. Pilch is the Copyright and Licensing Librarian and a member of the library faculty at Rutgers, The State University of New Jersey. Since 2001 she has researched and published on U.S. and international copyright issues. From 2007 to 2011 she served as an international copyright advocate for the Library Copyright Alliance, a coalition of three major U.S. library organizations—American Library Association, Association of Research Libraries, and Association of College and Research Libraries—at the World Intellectual Property Organization, and other international organizations, to promote fair and equitable access to information. Janice was appointed Visiting Program Officer on International Copyright for the Association of Research Libraries for 2009 to 2010, is a member of the Committee on Copyright and Other Legal Matters within the International Federation of Library Associations and Institutions, and chairs a permanent committee on copyright issues within the Association of Slavic, East European and Eurasian Studies. She has also served as chair of the Copyright Education Subcommittee of the American Library Association's Office for Information Technology Policy.
- **** Gloria C. Phares is a partner at Patterson Belknap Webb & Tyler LLP, where she practices in copyright, trademark and other intellectual property and related issues, representing publishers, libraries, museums and other cultural institutions, colleges and universities, technology and web-based companies and authors. Ms. Phares advised several not-for-profit publishers and others in connection with the Google Book Settlement and submitted a brief on behalf of the International Publishers Association, et al., amici curiae, in *Golan v. Holder*. She served on the Legal Advisory Boards of the 2011 Code of Best Practices in Fair Use for Poetry and the 2005 Documentary Filmmakers' Statement of Best

Karyn Temple Claggett: Columbia was a different school when I was here, actually, because we didn't have the wonderful new building that we have now, but I love coming back, so thank you, Pippa. As mentioned, our panel is going to discuss mass digitization. This is an issue that was briefly touched upon in the last panel, where a framework for the discussion was ably led by Professor Ginsburg. Our panel will explore the issue of mass digitization in the context of section 108 reform. Mass digitization of course is one of the most obvious gaps in section 108.¹ It's not addressed there, and though mass digitization has now become one of the most pressing issues on the copyright legal landscape, with cases such as *Google Books*² and *HathiTrust*³ (which were alluded to in other panels as well), the *Section 108 Study Group Report* did not explore this issue in much detail, 4 so there weren't any specific recommendations in the context of how to deal with mass digitization with respect to section 108. So it is both very timely and critical that we explore this issue now as we consider section 108 reform more broadly.

This follows, of course, on the mass digitization report that the Copyright Office issued in October of 2011,⁵ in which we raised a number of issues for public discussion relating to mass digitization, including: Is mass digitization a public policy goal? Should it be supported by copyright exceptions and limitations? And who should be able to do it? As we noted in our mass digitization report, mass digitization as a term has no singular definition. It can include anything from a hundred works to a thousand works to a million works. For purposes of our discussion today, we will just consider it to be a large-scale scanning process where there is a methodology or approach to scanning a wide range of works and perhaps making them available to others as well.

We have a very exciting panel that is going to speak with you about all of these issues today. I won't go into details in terms of their bios, as their bios are all in the materials that you were provided, but I will just briefly introduce them. On my far right is Eric Harbeson. He is the Music Special Collections librarian at the University of Colorado at Boulder. Right next to me is Paul Aiken. He is the executive director of the Authors Guild and has been since 1996. Right next to me on my left side is Janice Pilch. She is a copyright and licensing librarian and a

Practices in Fair Use (both sponsored by American University's Center for Social Media). Ms. Phares lectures on copyright licensing issues for Practising Law Institute's Advanced Copyright Program (1999–2013) and has written articles on various copyright issues, including restoration and art appropriation. She is a former trustee and officer of the Copyright Society of the U.S.A. and a former Chair of the New York City Bar's Committee on Copyright and Literary Property. Ms. Phares is a graduate of the College and the Law School of the University of Chicago.

- 1. See 17 U.S.C. § 108 (2012).
- 2. Authors Guild v. Google, Inc., 770 F. Supp. 2d 666 (S.D.N.Y. 2011).
- Authors Guild, Inc. v. HathiTrust, No. 11 CV 6351(HB), 2012 WL 4808939 (S.D.N.Y. Oct. 10, 2012).
- 4. See Section 108 Study Grp., The Section 108 Study Group Report (2008), available at http://www.section108.gov/docs/Sec108StudyGroupReport.pdf.
- 5. U.S. COPYRIGHT OFFICE, LEGAL ISSUES IN MASS DIGITIZATION: A PRELIMINARY ANALYSIS AND DISCUSSION DOCUMENT (2011), available at http://www.copyright.gov/docs/massdigitization/USCOMassDigitization_October2011.pdf.

568

member of the library faculty at Rutgers University. And then, finally, Gloria Phares is a partner at Patterson Belknap Webb & Tyler here in New York City.

I will start off by asking each one of them to briefly introduce themselves more fully and describe what their interest is in section 108 reform. Specifically, how section 108 reform might need to address mass digitization. I will start with Eric.

Eric Harbeson: So, I am Eric Harbeson. I am the Music Special Collections librarian at the University of Colorado. Despite my association with that institution and with the Music Library Association, I should say to start that my comments are my own and I am not representing either of them today. I'd also like to thank the Kernochan Center and the Copyright Office for sponsoring this and for inviting me to speak because this is a real pleasure for me. So my interest in section 108 is as a practicing archivist and librarian who happens to study copyright law as well. My interest stems primarily from the frustration that I see from my colleagues who are trying to do a job and run into legal roadblocks. Now, the archivists and librarians whom I speak with care deeply about copyright. They are concerned about rights holders; they are rights holders themselves. They advocate and teach copyright actually beyond what is expected or required by the law. At the same time, as I have mentioned, they have a job to do, and when the job conflicts with the law, that is when the problem occurs. One of the reasons that we are here today is because we have a law that was designed to help the librarians to do their jobs, but it hasn't aged well. So I am here because (a) I believe in the law, and (b) I think that it can be fixed to be more effective for everyone.

Since we are talking about mass digitization—and this may be a slightly different take than what Karyn proposed—I would like to offer an alternate way of thinking about mass digitization. And that is that mass digitization is, I think, more accurately seen as digitization projects where the decisions are made at the collection or repository level or some other level than at the item level. And this would include any copyright decisions. So if you are going through and you are making copyright decisions based on every individual work, you are no longer talking about mass digitization; you are just talking about digitization, which is not what we are talking about on this panel. That means that a mass digitization project could be as simple as, "I want to digitize those ten records that I have sitting in my basement," or it could mean a larger scale project, but it could also mean that a project to digitize fifty thousand photographs would not be mass digitization. So that's really all I had to say, except I want to make one note on terminology, and that is, I am both an archivist and a librarian, and I believe very strongly that both they and the museums ought be included in section 108, but I may refer periodically to one or the other, so I'd appreciate the assumption that that was out of laziness rather than an attempt to slight archivists or librarians.

Temple Claggett: Thank you, Eric. Paul?

Paul Aiken: Thank you. As I think some of you know, we have a rather keen interest in mass digitization issues at the Authors Guild. And this is a longstanding interest in what digital libraries will eventually be, what the rules are around them, and how it is that we make those rules work for everybody. We have to keep that in mind. However this is going to work, we've got to make sure that it's not just

backward-looking: how do we harvest what's been produced in the past and already fills our libraries and so it gets broadly out there? We also have to make sure we maintain the system so that authors really can afford to take time off from their other work. Most authors are not big authors who make huge amounts of money and are self-supporting. They've got to take leaves of absence. They need advances from publishers; they need to be able to do these things in order to create the books that everyone, including us, is so eager to get out there to the world so that people can have access to them. So we have to be sure that whatever we come up with works broadly for everyone's interest. And it's really a constitutional mandate. The Constitution is about four handwritten pages when you look at it, but they found space for the Copyright Clause. It's about creating markets for works, and that's because the founding fathers knew that markets were the best way to incentivize creation of all sorts of things. And in markets, it's not just the seller that benefits; the theory of markets is that it's a two-way transaction and both benefit. When you buy some fruit down at the farmers' market at Union Square, you benefit as a consumer just as the farmer benefits. And you benefit because the farmer then takes the time to go to that market and make those things available so people can have them. And so we want to make sure that those markets are still functioning because it's no accident that there are twelve million volumes in the University of Michigan library worth archiving and collecting. It's the result of a careful, planned out process that set up a system so that authors and publishers could benefit, and there would be markets created. Those works got out there and they are incredibly valuable. We have to make sure, as we go forward into this information age, that when information professionals—which include authors who go to libraries and look at all the information they can get—go into those libraries, they have the materials they need. And we have to make sure that they can digest, reinterpret, make relevant again all this information, and have professional editors who can help edit this material and publishers who can help get the material out there to a public that wants them. All those components have to be in place if we are going to continue to have a robust, literary market in the information age.

One thing I think is clear about the information age: we are awash in information. What is missing in many, many cases is that we don't have enough people working to make sense of the information. And it's not just authors and publishers; I'm thoroughly aware that this is librarians as well: people who collect and archive and say, "This is the stuff that's the most interesting, that's the most relevant to what you are doing"—we need that more than ever. We have to make sure that as part of this ecosystem we still have lots and lots of physical libraries and communities across this country and the world, so that there is a place for people to go where there are books around them or for people who are interested in books, and that is part of the conversation. Because the way independent bookstores are going, the way chain bookstores are going, it could be that libraries will remain the only relevant place in lots of communities that are book-centered places. And we need that more than ever. So there are my opening remarks.

Temple Claggett: Thank you, Paul. Janice?

Janice Pilch: Good afternoon. First I'd like to thank the organizers of this

symposium, the U.S. Copyright Office and in particular June Besek, for the opportunity to be a part of this important event. I'm here as a representative of Rutgers University libraries, but the views I present are my own. My views will benefit from past participation in the section 108 public discussions in 2006 and 2007 as a librarian. I will point out that I am not a lawyer. I'm pleased to serve on this panel with representatives of authors and creators of literary and artistic works. The interest that librarians share in preserving and fostering works of the mind is a testament to those who created them.

We are pleased that the Copyright Office is reopening the discussion of section 108 in the context of mass digitization. The safe harbor in section 108 still plays an important role in enabling libraries to meet immediate, practical needs and in ensuring continued recognition of libraries and archives by Congress as trusted institutions that preserve the cultural record of societies, of future generations, and make it accessible to people. We are interested in seeing that the system used for centuries to preserve intellectual and cultural heritage—libraries, archives and museums—succeeds in the twenty-first century within the landscape of changing technology, evolving practices and the law. As a matter of public policy, this needs to happen, and it should be viewed as a strategic, national priority. We also have a collective responsibility to establish a global infrastructure for library collections. Advances in the European Union and in other countries reinforce the obligation to agree on a legislative approach to mass digitization and, within that, the orphan works issue. Successful section 108 reform would result in a provision that provides reasonable certainty for the goals of digitization and that is flexible enough to accommodate developments in technology and society into the coming decades. It would also strengthen the role of U.S. libraries within the international framework of institutions designated as guardians of the public trust that are working to preserve and make available the intellectual and creative record within the larger vision of a digital future. I look forward to an interesting discussion.

Temple Claggett: Thank you, Janice. And Gloria?

Gloria Phares: I am in private practice here in New York. I represent libraries, museums, publishers, authors, software developers, software users, etc., but these are my views. It is not clear to me that section 108 reform leads inevitably to mass digitization, if that means digitizing entire collections generally. digitization"—the very words on their face seem to me, and with no qualification, incompatible with the notion of copyright law as we know it. On the other hand, we all can think of examples, at least a few, where digitization of large numbers of works might facilitate worthy societal goals. If we accept comprehensive violation of copyright owners' rights—and that's what that would be, to my mind—then the tradeoff should be a legislative scheme under the auspices of some part of the government that provides supervision and control with respect to four factors. First of all, articulated purposes. One example that seems to me possible, and that would not probably work any harm to copyright owners, is search vehicles used for various reasons. Second, the institutions that should be permitted to make such copies, and whereas this is an old fashioned word, there should be the notion of perhaps "trusted institutions." I think that people have often conceived virtues of

digital collections, large scale, perhaps not in private hands but in certain kinds of public institutions. And that leads to three: where there is considerable attention to security, which I mean both in terms of preservation and protection of access, etc. And four, the involvement, if any, that profit-making entities should have in that process, beyond of course being paid for services.

And a topic that I still think, and especially after this morning, needs a little bit of addressing: even though reliance on section 108 does not preclude 107 reliance as I think the statute says pretty clearly, 6 I see no reason why we shouldn't have an updated section 108 that provides reliability for libraries by expressly authorizing activities rather than relying on the extremely fact-bound aspects of fair use analysis, which people often misunderstand, or relying on district court decisions, which as I said earlier, are all on appeal.

Temple Claggett: Thank you Gloria. You all really set up well some of the issues that we're going to discuss in more detail throughout the panel. So the first question I'll direct to the library representatives as a kind of broad question that we can address throughout our discussion: can and should section 108 be expanded to allow mass digitization, at least for purposes of preservation? As we heard earlier today from Lolly, currently only unpublished works are allowed to be copied for preservation, while published works are allowed to be copied for replacement. Should libraries and archives be allowed to digitize published works for preservation as well, and should they do so through section 108 or should other avenues be explored such as licensing or reliance on section 107 as has been alluded to as well? And so I'll start with Eric.

Harbeson: So the way I see this, the problem with the distinction between published and unpublished works is twofold. One, the definition can get blurry. It can get muddled and that already gives us a little bit of unclarity. But even if you accept the existing definition of published versus unpublished works, it fails to get at the real crux of the problem. The way sections 108(b) and (c) appear is as trying to draw a distinction between those works that are irreplaceable and those which are not. Now, it's true that in general the unpublished works are going to be the ones that are extremely rare and difficult to replace. There are also, of course many published works which are ubiquitous and easy to replace, but that is not always the case. I'll take an example from my institution, which has hundreds of thousands of copyrighted sheet music in its collections from the 1920s and 1930s, much of which is impossible to find anymore. It's published, it's under copyright but we could not make section 108(c) copies the way that it's currently written. So I think that that's the problem with the published-unpublished distinction. I think that Congress identified a good goal in trying to identify preservation as a very, very important mission. It is one of the few places where I actually think that a mass digitization in the sense of a collection-level decision might be appropriate. In general I think that Congress needs to be very careful if it starts talking about collection-level rather than item-level decision-making, but this is one place where I think that would be appropriate.

Temple Claggett: Mass digitization for purposes of preservation?

Harbeson: For purposes of preservation. I would actually go so far as to say that for purposes of preservation the laws should be sufficiently open—that pretty much anything can be digitized for preservation. I do not think that this would cause a problem for the markets of the rights holders because what would happen is the preservation copy itself would be subject to item-level access restrictions. Rather than having the restrictions made at the point of preservation, it would be made at the point where someone actually wants access to it. That would allow us to take care of the preservation problems which are often more urgent than the time to make copyright decisions allows, and yet it would still protect the rights holders' interests in protecting their markets.

Temple Claggett: Thanks Eric. Janice, did you have anything to add in terms of whether it would be helpful for libraries to be able to preserve both unpublished and published works under section 108?

Pilch: Yes. I see two issues here. The first is not a new one. It's the issue of the separate treatment of unpublished and published works in 108(b) and (c). Because 108(c) requires replacement as the first option for preserving published works, it allows preservation only if replacement is not possible. I think, and I repeat, this could be stated as such: a single preservation provision entitled, "Preservation," with replacement being the first required option for preserving published works. This would provide clarity. It would be beneficial. And this is a relatively small point but it leads to a larger point: the issue of whether libraries, archives and museums should be allowed to digitize published works for preservation without the replacement requirement. Here the Section 108 Study Group recommendation for a preservation-only exception for at-risk published or publicly disseminated works might be helpful to start, to build on. It was a recommendation to allow digitization prior to detectable deterioration under four conditions.⁷ It might be the basis for an exception permitting digitization for preservation without the replacement requirement just for the preservation, as Eric described. Any further ruling in the HathiTrust case might clarify further whether libraries should be permitted to do this. There are compelling reasons, as we know, for allowing libraries to preserve works before they go out of print and before they become damaged or deteriorated, become lost or stolen, or end up in an obsolete format. We know this. Consider the alternative: libraries wait for works to deteriorate, get damaged, etc., preserving them only on a case-by-case basis, perhaps not in time. We have an obligation to preserve works. This is what libraries do, and this is a realistic way to move into the twenty-first century. And although libraries have already embarked on large-scale digitization efforts, the benefit of creating an exception now would be for the clarity and certainty it provides into the future.

Although access is not the subject of this panel, to bring this issue to the logical conclusion, the upfront preserved works should be made accessible under lawful conditions only: as in-copyright works fall into the public domain, as U.S. pre-

^{7.} SECTION 108 STUDY GRP., supra note 4, at vi-vii.

1972 sound recordings are brought into the federal regime in 2067 and fall into the public domain—or earlier perhaps; as works are determined to be in the public domain, as right holders of in-copyright works opt in and give permission, as termination of transfer provisions allows authors to regain their copyright and they opt in, as right holders of suspected orphan works are identified and either of those two things happens; and as works are established with reasonable certainty to be orphans, perhaps after a reasonably diligent search, and they can be made available, etc. But I think, and I'll be very quick about this final point, I think the key to mass digitization will be in the national coordination. It might even include some more centralization, some kind of more central authority than we're accustomed to thinking about. I'm not against the idea of some libraries, quoting Maria Pallante's comment this morning, of some libraries being allowed to do things that other libraries can't do for the purpose of mass digitization. When it comes to mass digitizing, maintaining and eventually making publicly available digital copies of published materials—because I view mass digitization as systematically scanning works with the ultimate objective of making them available electronically, possibly, publicly available on the Internet—we need a national plan to get the digitization done, to avoid duplication, to eventually make material publicly available lawfully. There is no need for every library to handle mass digitization on its own. So when we say: "Should libraries be allowed?," I say libraries as libraries coordinated in a national effort underpinned by legislation and also underpinned by a plan, a clear framework as a cultural imperative. And I view this as an octopus with one brain. I think they have one brain.

Temple Claggett: Thanks, Janice. And to give both Paul and Gloria an opportunity to respond: I think that your answers might be different in terms of preservation as opposed to access, but on the preservation point, are there concerns from rights holders in terms of allowing preservation to be one of the goals of section 108 reform?

Aiken: I'd first like to respond to part of what Janice had to say, and that's to say that there's a lot I agree with there, that there should be a national plan on how to deal with this. It shouldn't be ad hoc, see how hard you can push the boundaries of 107, 108 or whatever. This is a critical matter for the future of literature, and we have to figure it out, we have to get it right.

And it's really very appropriate we're meeting here in February of 2013 because this month is the fiftieth anniversary of the first debate that I know of over whether or not books should be digitized as a matter of fair use. It was February 20, 1963 and there was Abe Kaminstein presiding from the Copyright Office. The subject of the debate was a proposal for a new copyright law; of course, this led up to the 1976 Act. At that time the preliminary proposal was that there would be a copyright for information storage and retrieval use in the right of the rights holder, which would be to control information storage and retrieval, putting it into computers. Before the discussion even began a patent attorney from Los Angeles named Reed Luller sent a telegram to Kaminstein with a proposal, and so

Kaminstein read this into the record: "Since we're going to have this information storage and retrieval, propose that in any event reproduction of a copyrighted work in machine-readable form for use in the analysis, citation and reasonable quotation of the work by means of an information storage and retrieval system shall be considered a fair use." Well, people had lots to say about that. A Yale Law School professor discussed whether this would go beyond note-taking because everyone thought of library uses as being some proxy for note-taking: that when you did a photo-duplication of some sort, it was just substituting for the normal research purposes of note-taking. And one lawyer said, "With the way these things seem to be going, there's a good possibility, within the lifetime of this statute"that's where we are right now, in the lifetime of this statute—"they're going to eliminate printed books for most purposes. If we take Professor Brown of Yale's view that you can put material into these machines as a matter of note taking, you may find that for practical purposes, you have eliminated the market for the book entirely." ASCAP was around back then, of course, and a lawyer for ASCAP of course was there, and he said:

Well when we talk about information retrieval systems at another point why do we say 'information retrieval'? Suppose that it merely retrieves entertainment? I wonder if at the beginning, right at the introductory sentence about what rights a rights holder has, and this is a matter of drafting, we couldn't say, 'The rights granted under copyright shall include the right to authorize any of the following with respect to the copyrighted work.' 10

And that made it into the next draft of the formally introduced bill in copyright. That's where it came from, that concern about losing control as it goes into the machine. And then Goldman from the Copyright Office later says, "Are you suggesting that the need is to control putting the work into the machine?" And ASCAP attorney: "Yes." Goldman: "And then you don't have to worry about the use by taking it out of the machine?" And Rothenberg: "Then it would be merely by contract. Whatever arrangement the copyright owner wishes to make with Remington Rand"—that was a computer company—"or whoever the company is, or the program." Irwin Karp, of the predecessor to Author's Guild, Author's League, was there: "I don't think you can solve it in any other way than by controlling the right to put it in. I think you have to control both the right to put it in and the right to take it out." "

Phares: Sounds like Irwin.

Aiken: So that was fifty years ago this month. Information storage and retrieval: you would tend to think this is all new, but this was being discussed actively when TVs were black and white and there were three networks. This was something that was active—there were demonstrations. In the early 1960s, American library associates and others had a demonstration of these coming machines in San Francisco and invited everyone to look at the libraries of the

^{9.} Id.

^{10.} *Id*.

^{11.} Id.

future. There was a report in 1965. Special Libraries Journal reprinted something from the New Scientist in London, and here is what the New Scientist in London had: it was an article called "Computers in 1984." I don't think that was a random date, so it was 1965 predicting what computers in 1984 would be able to do: "Connection to a central location will be very necessary to perform another function, which will, by then, be delegated to the omnipresent computer. I refer to information retrieval. The entire contents of large central files"—here's the 1984 part—"or at least the portion the government elects to make available, will be readily retrievable at a moment's notice. One will be able to browse through the fiction section of the central library, enjoy an evening's light entertainment, viewing any movie that has ever been produced, for a suitable fee, of course, since Hollywood will still be commercial," the movie guys always have it, "or inquire as to the previous day's production figures for tin in Bolivia all for the asking fee for one's remote terminal. Libraries for books will have ceased to exist in the more advanced countries, except for a few which will be preserved at museums." And some of you familiar with the main branch of the New York Public Library may think this resonates a bit: "Perhaps it would be more correct to say that all the world's recorded knowledge will be in some form since programming computers will be able to read not just printed material and input them in but also handwritten stuff by that time," this guy thinks. 13 One other short selection from this article on computers in 1984 from a prescient 1965 writer:

Computers will have largely taken over the task of composing and arranging music, at least for the popular entertainment, and many people will vie with each other in regard to the quality of mood music which their personal computer or personal programmer can produce. As far as literature is concerned, the computer will still be a neophyte, although who-done-its will be turned out by the million. ¹⁴

Temple Claggett: Now do I take that to mean, for preservation purposes, you think that the rights holders should still have to grant the authority for mass digitization projects, or whether an exception expanding section 108 would be appropriate?

Aiken: I think for fifty years we have recognized there is a tremendous risk being taken when things are digitized, and we have to do it carefully. This is not something to be done ad hoc. It's not something to be done by fiat. It's not something to be done because you happen to have a very well-funded tech company behind you who is willing to push the limits of the law and what it views as fair use. This is a national discussion of national importance and should not be left to those who want to push the boundaries of the law because we're playing with fire. We could digitize everything. Who's protecting it, and what are their incentives to protect it and keep it from getting out? Do they have sovereign immunity so they don't have to worry? They don't have skin in the game. This should be something that's carefully thought through, so we don't wind up having

^{12.} A. Samuel, Computers in 1984, 56 SPECIAL LIBRARIES 58 (1965).

^{13.} *Id*.

^{14.} Id.

literature Napsterized. It's vital; it's not a minor matter that should be decided by those who think they can push the limits of what's allowed.

Temple Claggett: Thanks, Paul. Gloria, did you have anything to add?

Phares: Actually I think what's more interesting about Paul's example is the tremendous mistrust of technology—and maybe partly from people who don't understand it as well as they might. But the fact is that we all know how easy it is for copies to go viral and to lose control of materials. And personally, Eric's notion of preservation, as he says, at a collection level with item level access restrictions is appealing in many ways. I happen, however, to be a book collector; I still think paper survives very well. But at the same time, I think the notion that I was referring to earlier of trusted institutions—and those institutions don't always necessarily have to be brick and mortar or even governmental, but they probably would not be commercial—is the notion that these are not things that have to be done everywhere. In fact, the proliferation of it is partly what I think causes even more anxiety on the part of copyright owners, so that we have to think about the balance of trying to achieve goals that we think further our national goals, but also protect the rights of the people who create that national culture.

Temple Claggett: Thanks, and Eric, did you have something you wanted to add?

Harbeson: I did. I just wanted to say that there's a lot in Paul's example that is not inconsistent with the idea of libraries engaging in mass digitization for preservation. Now, we can argue whether digitization is the right way to preserve an item. Preservation in the library and information science community is a vibrant and well-researched field of study. Our researchers are doing a very good job of determining what makes sense from a preservation point of view. When the people who are determining what the best way to preserve something are saying one thing and the law is saying that you can't do that, that's where we have a problem. Now, I want to be clear: I am not advocating that we necessarily should put everything that we preserve into a dark archive, but by the same token, I'm not necessarily arguing that we should put everything on the open Web either. That would be where we would have the item-level decision-making as far as access decisions. But there are two real points that I wanted to make. The first is, we're not talking about fair use. We're not talking about people who are trying to push the limits. We are talking about people who are looking for limits that allow them to do what they need to do. Section 108 is for people who want to stay within the limits. If we wanted to talk about the Google Book Project—depending on which version of the Google Book Project—that is not necessarily something that I would advocate a section 108 safe harbor for. But that's not what's being discussed when we're talking about preservation. I know a lot of librarians who would argue it's not really preservation.

Temple Claggett: Well thank you, Eric. Giving Paul an opportunity to respond to that: if we're still talking just about preservation—would that address some of your concerns if it was limited to preservation? And if it was also limited in terms of who could do the preservation—so, for example, if it was limited to just libraries and archives and potentially museums. Would that address some of your concerns

about security and restrictions in terms of how the material would be protected?

Aiken: Well, is digitization necessary for preservation? In the book world, and I know there's a lot of other worlds out there, but just looking at books, unless you are talking audio books, it's hard to see why digitization is a necessary component of preservation. Why do we have to change? And certainly for physical books, why do we have to change that to a new medium that is inherently at risk and have that outside the rights holders' control for the purposes of preservation? We have been able to preserve books for decades and decades without resorting to digital copying. So it's hard to see the need as far as physical books are concerned.

Harbeson: May I respond?

Temple Claggett: I was going to say either Janice or Eric.

Phares: I could actually give you an example, but I would also say that I don't know that the digitization would be required for the preservation. Suppose you have a fine press book of 250 copies which have long since been sold and are now, if they are to be obtained, could only be obtained at \$35,000 apiece. I mean I could understand why someone might want to have one in their collection and maybe make a copy, maybe not even a digital copy but a paper copy, just so that people could handle it at that level and preserve the physical copy, that one single copy of an edition of 250.

Temple Claggett: And Eric, did you have something to add in terms of why digitization might be a necessary factor or potentially necessary for preservation?

Harbeson: I do. I actually was going to point out that this is exactly the debate that people who are researching best practices for preservation are discussing in the library and the information science community. We're very good at determining whether something should be preserved and how it should be preserved. This is something that we are very good at and something that Congress is not very good at. And when you say I don't know that digitization is appropriate for this book, I think that most librarians would probably agree with you because preservation is expensive. Digital preservation is very expensive. So librarians, when left to their own devices, may or may not decide to make the choice to digitally preserve something. But if they make that decision to expend those resources when they take into account actual costs and opportunity costs and the like, they should be able to make that decision and know that the law is behind them. If we want our materials preserved for the future, the law has to be unequivocal in its support for library preservation.

Temple Claggett: Now, Janice, did you want to add anything?

Pilch: Two quick comments. Just to say that digital media are the way that works are going to be communicated in the future, number one. And number two: the publishers don't preserve them. And so if the future is going to be digital and publishers aren't preserving them, we have to build the digital future. That's what libraries do.

Temple Claggett: And sticking with preservation for one second, in terms of the types of works that might need to be preserved, are we talking just about books, or do we need to make sure that preservation would include, for example, sound recordings and other types of categories of works as well?

Harbeson: Every work that has been created is potentially in need of preservation.

Temple Claggett: Turning to a slightly broader topic: outside of preservation, is there a need for mass digitization of copyrighted works for purposes other than preservation, and if so, what would the need be? Gloria?

Phares: Well, I mean it's not that someone chose it but as you know, there certainly have been uses of, as I said, large databases of text material for search. I know of programs that have been used for using the same kind of materials for language study—just because they provide a humongous database of human expression and how it changes and so forth. There are objectives that I think are worthy of that kind of use, but in my mind, that doesn't justify 100,000 libraries making a digital database. Or even one thousand maybe.

Temple Claggett: Janice, did you have something to add?

Pilch: My reply on this is very different from Gloria's. I can see three reasons why libraries should be permitted to digitize for purposes other than preservation. The first is getting back to the issue of 108(d) and (e): the issue of copies for users, direct copies and copies for interlibrary loan is an unresolved issue from the Section 108 Study. The Study Group concluded in principle that electronic delivery could be permitted under (d) and (e) if libraries and archives were to take additional adequate measures to ensure access was provided only to specific users and to deter unauthorized reproduction. ¹⁵ 108(d) and (e) currently don't prohibit the use of digital copies, but they don't expressly allow it. ¹⁶ It's possible that copies made through a large-scale mass digitization effort might be used under 108(d) and (e). That said—

Phares: Do you really have to have mass digitization in order to have a digital copy for interlibrary loan?

Pilch: You don't need it. What I'm really saying is that large-scale digitization is not being done primarily for that purpose. We're not doing it primarily for that purpose, and we wouldn't do it primarily for that purpose. That would never be the reason that we mass digitize. But just to say that mass digitized copies could be used and that would be another reason why it would be desirable. Again, these works would be used within a very defined shell for purposes of private study, scholarship or research. The idea would not be to increase access to the works in any way, just to provide more efficient access to these works for private study, scholarship or research as a safe harbor. So that might be one reason.

The second reason would be that at least some categories of libraries should be permitted as authorized entities to digitize published works to create accessible copies for blind and print-disabled persons, to supply them, upon request, to those users, and to loan or provide such copies to other libraries or archives to supply to a reading-disabled user, upon request. This might not exactly be a 108 issue, but it would be beneficial to clarify the goals of library digitization in relationship to the exception in section 121 as the text of a proposed WIPO instrument on limitations

SECTION 108 STUDY GRP., supra note 4, at x.

^{16.} See 17 U.S.C. §§ 107(d), (e) (2012).

and exceptions for visually impaired persons or persons with print disabilities moves to a diplomatic conference in June of this year. ¹⁷ Any further ruling in the *HathiTrust* case will also be relevant to this issue.

And the third reason why we might digitize would simply be that libraries and archives should be permitted to digitize and make available orphan works, if agreement can be reached on an appropriate framework for making orphan works more broadly available. The Study Group concluded that because orphan works issues arise broadly across many different uses, in addition to those of libraries and archives, orphan works legislation and not the Section 108 Study Group would be the appropriate place to address them. However, if an orphan works provision were to be adopted that would apply only to libraries and archives, 108 would be a good place for it. So I can think of those three purposes.

Temple Claggett: Well I think that you raise a good point. I think in discussing mass digitization, orphan works is something of an elephant in the room in the sense that you can't really discuss mass digitization without addressing some of the orphan works issues. And I know, Paul, you might have something to add on that particular point.

Aiken: We have a well known interest in orphan works. So I will once again go way back in history, to 1959 this time, when you see the first ad from UMI about print-on-demand books; ¹⁹ the first ad that I was able to find. The preliminary drafts from the Copyright Office of the new 1976 Act are being circulated and commented on, and the first report suggested that entire books should be available for copying, in their entirety, by a library for its users. ²⁰ If a user requests it, you can photocopy an entire book and make it available to users. There was a bit of an uproar about this from publishers and authors saying, "That doesn't really protect our market. Can we do something else? This is supposed to be about note-taking, not about reproducing entire books." So a revised version was then circulated and this version said entire books, but there has to be a good faith effort first to see whether or not the book was available on the market.²¹ So basically out-of-print books could have been copied in their entirety under that proposal. Well, Irwin Karp didn't take that too kindly, and there was a lot of discussion about that. And one of the things that was discussed was that, look, we've got print on demand technology; it wouldn't be surprising if the costs of these reproductions came down lower and lower, and you could be talking about something that could actually supplant a market that should develop for these currently out-of-print books. And what happened, when a version was finally introduced, was that there was no provision allowing for copying of entire books.²² This was around 1964 or 1965.

^{17.} See generally World Intellectual Prop. Org. [WIPO], Draft Text of an International Instrument/Treaty on Limitations and Exceptions for Visually Impaired Persons/Persons with Print Disabilities, VIP/DC/3 (Feb. 5, 2013).

^{18.} See SECTION 108 STUDY GRP., supra note 4, at 24–25.

^{19.} Source on file with speaker.

^{20.} Source on file with speaker.

^{21.} *Id*.

^{22.} Id.

The 1965 Kaminstein Report is what people should be looking to, not CONTU, to find out what was going on in these early meetings, how the law was being adjusted to address these sorts of concerns. The May 1965 Kaminstein Report is the critical document.

But while this is going on, in the early sixties you start to see ads saying that from UMI, Bell and Howell and 3M; they start competing to see who can make the most works available.²³ UMI, of course based in Ann Arbor, had Michigan's library that it could use, and it ran ads such as this one here: "Why should you look for things when we've already found them?" UMI says—and by this time UMI is owned by Xerox, the black-and-white TV version of Google—"If you don't have what you're looking for, we can get it for you. We'll find the book in Timbuktu if we have to. We'll clear the rights, pay royalties to the authors. We have tens of thousands of old books that, through these methods, we've already microfilmed."24 This is UMI. Bell and Howell focused on another set of books; I think they did especially Russian books—there was a lot of Cold War interest in Russian books and 3M did most of its work with the New York Public Library. And there was a race; you can see they're advertising how many rights they've cleared, how many out-of-print books are now available through this print-on-demand process. They're finding rights holders, which is the best solution to orphan works: find the rights holders. They're finding rights holders, clearing rights.

By 1977, UMI could boast that it had 84,000 books available through print on demand. And we now have proud owners of catalogues of those 84,000 books available. They look like phone books. If you browse through those pages, you find interesting things like seven of the works that were purported orphan works, that in the first round of works were "unfindable" by HathiTrust—there they are, they found them without Google, before 1977. The orphan works problem is a problem that can largely be solved by markets if we let the market solve it. It's especially easier now because we have Google, that can help find all sorts of things; most of the books that have been published and that are in copyright were published in 1993 or 1994 or more recently. That's why you can find rights holders. Most of the books that are in copyright have been published since 1993 or 1994, and a lot of the older ones you can find because UMI already found them for us in 1977. The orphan works problem is grossly overstated. It can be solved. We have to allow markets to solve it.

Temple Claggett: Thank you, Paul. **Aiken:** That's my other speech.

Temple Claggett: Eric, you actually look like you might want to add something in response to that?

Harbeson: There are a couple things I have. First of all, I love the history here. This is a lot of stuff that I had not heard.

Aiken: I think you're not the only one.

23. Source on file with speaker.

24. *Id*

Source on file with speaker.

Harbeson: I think that that's probably true. You know, the story from 1959 was interesting to me partly because despite the fact that we were talking about this back in 1959, I know of some libraries that are engaging in print on demand, but I don't know of any that are trying to engage in full scale print on demand without clearing rights. Maybe there are some. There may be some. I don't really think that the orphan works problem is a section 108 problem, that it is what this boils down to. I think that we get close to orphan works in section 108(h).²⁶ I think that we could play with 108(h) to make something that comes close to the orphan works problem, but if you can find orphan works—possibly through Google, maybe you can find them through Google because Google already scanned them—but if you can find the rights holders, then great, they're not orphan works. But the problem with orphan works is that they are in fact orphans, and there are works where you cannot find the rights holders. We can argue about orphan works, I don't really think this is something where we're looking for the kind of safe harbor that 108 is providing. In some places we may be able to do it through 107.27 In other places we may be able to do it through damage remission or other creative strategies, but I don't think that this is a 108 problem.

Temple Claggett: And to put in a plug for the Copyright Office, obviously this is something that we are looking at right now. As I think was mentioned earlier today, our orphan works comments period just closed, and so we're going to be engaging in this discussion with the public more fully. Before we turn it over to the audience, and I know we wanted to make sure that we have an opportunity to get your questions, we alluded to market-based solutions possibly as being one approach to section 108 reform. Is the market a solution in this area in terms of allowing for direct licensing to create situations where libraries can preserve works, on a mass digitization scale, or do you really need to have specific exceptions and limitations in this area?

Harbeson: So are we talking about nonpreservation issues?

Temple Claggett: I would start with preservation because I think that's a little bit easier, but then expand in terms of the other purposes that folks on the panel raised in terms of potential uses for digitized works.

Harbeson: I think that licensing is a solution for some problems, but it is not a solution for everything. There are a number of agencies that are doing good work in making licenses available. That's very, very good, but it is not the solution for all of the problems that we're facing. With preservation, I really don't think that a licensing arrangement is the right solution because, as I said, preservation is not an inexpensive process. It requires significant skill on the part of very highly trained, very well paid preservation experts. It requires significantly expensive tools to do so. We may be talking about digitization, or we may be talking about other kinds of preservation, but preservation is expensive. If libraries have to go through the whole process of not only searching for but also paying for licenses to preserve their materials, they're unlikely to do it. I think that this is a case where—

^{26.} See 17 U.S.C. § 108(h) (2012).

^{27.} See id. § 107.

especially if we're talking about just the preservation, we're not talking about access yet—if we let libraries do what libraries do best, if we let libraries do the preservation the way they see fit and then we legislate on how we can provide access, and we have the start for language with the existing 108(c) through (e), I think that that's a better strategy. I admit what I'm saying is I think we should be able to preserve everything with abandon. I'm saying that we could be allowed, in theory, to preserve *Harry Potter and the Sorcerer's Stone*. If the regulations are put together carefully, I think that we could come to a place where if a library for some reason decided to digitize *Harry Potter and the Sorcerer's Stone*, they would be able to do nothing with it except put it into a dark archive, because it's obviously ubiquitous, it would obviously pose a threat to a rights holder. As the rights holder's interest in exploiting her material decreases, the library's interest in providing access might increase. But that's a different place.

Temple Claggett: Does anyone else on the panel have anything to add on that? **Phares:** Well, the only thing I was going to say, Eric, is that I sometimes think that, I mean, what I hear you say over and over again is that the libraries know best how to preserve things, and I think, "OK," and then you sort of hesitated there about access. And I think that, in fact, what sometimes makes people nervous about the license that they're willing to give to the preservation side is the fear that the desire to deliver will not be suitably controlled in the interests of the copyright owners. And I think ultimately these are issues of trust. And also because of this notion of mass digitization and how many people manage to screw it up. I mean, it's hard for me to believe that the HathiTrust managed to put as many works that were in copyright into its first round of materials. I think everybody has to understand that each side has legitimate concerns and that everybody is going to have to—and I think it's what drives this concern that you also say, "Well, we make access," and I'm thinking, "Well, who's 'we'"? That's a huge decision: who are going to be the deciders of the access, item by item?

Harbeson: This is where I am proposing that we actually do have legislation to regulate. I guess I could have seen this coming, but I'm hearing a lot about Google Books and HathiTrust. I don't actually regard these as a model that is useful when thinking about mass digitization from the point of view of 108.²⁸ We can argue—and I would love to argue with you some time—over whether it's fair use, but that's not the question we are having here. So I don't really think that HathiTrust is the right model to look at for what might happen.

Temple Claggett: With respect to section 108.

Phares: Well alright, but I have to just tell you that I sat in a Columbia room, a different one, a few years ago and heard the director of the University of Michigan Library say, "If I'm paid enough money I will make my library available for copying." So, it is those kinds of comments—it may be said in jest, maybe not—that make people anxious when it comes to sitting down at the table and deciding who is going to control what.

Temple Claggett: Who would be able to do it?

Phares: And whether or not it falls directly into the 108 basket or they overlap. I can see why orphan works don't fall precisely into this area, but if we are talking about mass digitization, it's inevitably going to come up. I think people can resolve these issues, but everybody has to be listening very hard to the other side.

Harbeson: I agree with that.

584

Temple Claggett: And I want to make sure that we have the opportunity, as I said, for the audience to be able to ask questions. Paul, did you have something you wanted to ask?

Aiken: Yes, very quickly. Eric, what I am hearing though is that preservation is a case-by-case thing and has to be carefully thought through. And that sounds right. This is a matter that requires some expertise. And I'll say that preservation is vitally important to authors. A lot of the things become subjects of books because they've been carefully preserved and can be used for research. But that case-by-case analysis suggests that there are two things wrong with mass digitization for preservation: the mass part and the digitization part. Do you really need to digitize everything to preserve it? We really have to look at when it makes sense and when it does not make sense to put things at digital risk.

QUESTIONS AND ANSWERS

Temple Claggett: Thanks, Paul. I do want to open it up to the audience because I don't think that we have a lot of time left, but I think that your provocative last statement will generate a number of questions.

Aiken: It wasn't that provocative. I can do better than that.

Temple Claggett: I see Professor Ginsburg in the back.

Jane Ginsburg: This clearly is inspired by both Gloria and Eric, and I'm not sure how it fits into 108 or if it is actually [inaudible], but when we talk about preservation, are we talking about preservation of particular libraries' collections or are we talking about preservation of the work?

Phares: I actually began to think about that myself as we were talking.

Ginsburg: Because I think you can get very different answers depending on where you pose that question.

Temple Claggett: Thank you. That's a great question. So the question was: when you are talking about preservation, were you talking about preservation in the sense of the library's collection or were you talking about preservation in the general sense of the work itself?

Phares: In other words physical versus the intangible property work. I guess I was thinking of it in terms of the physical.

Temple Claggett: I think it's in terms of, are you going to preserve the library's collection itself or are you going to preserve just the work generally for purposes of the public, but not necessarily limit it to the library's collection? Is that right, Professor Ginsburg?

Ginsburg: That's right. If what you're concerned about is having an instantiation of the work somewhere from which it will eventually be accessible under whatever conditions are determined, it doesn't necessarily follow that every

library may have an instantiation of its entire collection. I can see that libraries which have particular collections that are not available elsewhere are excellent candidates. But I guess this is the redundancy test that Gloria was asking about. In terms of public policy, what is our objective? Is our objective to preserve the collection? Through digital as opposed to [inaudible]?

Temple Claggett: And I think this was alluded to in the last panel discussion as well in terms of, could you possibly just have one library such as the Library of Congress preserve one copy of the work, so that you have the works preserved individually as opposed to necessarily having all libraries be able to preserve their entire collections in every way?

Harbeson: The great existential problem, of course, is to make sure that the work survives. So, in that sense, it's true that if one library preserves it, and preserves it perfectly for the rest of time, that that work will be preserved. Let me say, too—and this is something I was dying to say during the last panel—libraries ask these kinds of questions of themselves all the time. Some of you, many of you, most of you, are familiar with what started out as, I believe, the Ohio College Library Consortium. It had nothing to do with the Library of Congress. This was a bunch of people that got together and said why are we not consolidating our catalogues and making a great big catalogue of all of our works? And then they start to think in terms of, well, why do we all have to catalogue this if we are using the same standards? One of us can catalogue it and the rest of us use it. The thing about the OCLC model that worked so well was that it was distributed. Now, because of that, I am reluctant to say that Congress should be involved in picking who or which library should engage in any preservation activities, because I think that libraries when left to their own devices will recognize when there is a need for redundancy—redundancy can be itself a preservation measure—and when there isn't, and there will be cooperation to reduce costs. So, yes, we are trying to preserve the work. We are also trying to preserve our library collections, but I think that it's the wrong question to ask when we say who should be allowed to preserve it and who shouldn't because I think that libraries are pretty good at selfregulating on that front.

Temple Claggett: Question right here.

Nancy Kopans: Nancy Kopans from Ithaka. I think that some of the challenge here is staying rooted in existing models of institutions. There are other models that are developing, developed as well, as to who preserves material. Some of this could be going forward and some of this could be retroactive, but I am thinking of *Portico*, ²⁹ I am thinking of *CLOCKSS*. Where, with *Portico*, it's the creator's decision to put content in an archive for born digital content, with certain trigger events making it available. So, I say that to introduce that as part of the conversation: it's not always simply about will libraries be preserving; libraries and content creators might lean on other resources that are brokering these arrangements in new kinds of ways.

^{29.} PORTICO, http://www.portico.org (last visited Aug. 1, 2013).

^{30.} CLOCKSS, http://www.clockss.org (last visited Aug. 1, 2013).

COLUMBIA JOURNAL OF LAW & THE ARTS

[36:4

Temple Claggett: Right, and that was actually one of the questions we kind of didn't get to in terms of the panel. Should libraries be able to partner, for example, with for-profit entities or others in order to preserve their collections?

Kopans: These are non-profit, just to clarify, for what it's worth.

Temple Claggett: Jonathan.

Jonathan Band: Jonathan Band. Paul, didn't you actually make my argument for me?

Aiken: Not on purpose.

586

Band: I mean you basically demonstrated that we've been having these debates for the past fifty years. So, if we were waiting for legislation to resolve them we would never do anything. Not only would we not have search indices, we wouldn't have DVRs, we wouldn't have VCRs, we wouldn't have personal computers, because the affected stakeholders would never agree on any of these devices that rely on what courts have found to be fair uses.

Aiken: Actually, we have all those things, and we also have section 108. We did come to very important decisions about limiting the uses that could be made. We have got to preserve books, but we also have got to preserve markets here. And we have to carefully craft the law so that we do that. Otherwise, the valuable stuff we all want to preserve and make available just won't be created in as much abundance in the future. We'll have disagreements, but over time we have been able to make changes. We were able to get 107 and 108, and were able to get 108 amended once upon a time to include certain digital uses and it wasn't that long ago. So I think there is room. In the heat of the legislative moment, it may seem impossible, but I think that things do progress.

Temple Claggett: We don't want to go into the other panel too much so I think we have time for one last question.

Stephanie Gross: Stephanie Gross from Yeshiva University. I've heard Jim Neal talk about fair use, and he's talked about alternatives issues. And I, as an academic librarian, am more intrigued less by maximizing use than not just simply what survives. In other words, you have "just in time" versus "just in case," as well as issues of collection development and issues of what librarians should be planning for the future, and not retroactively disputing issues of damage control, print going to digital, and so forth. This is something I feel is a major issue that seems to fall through the cracks. I think that we as librarians should be sensitive to the issues that are brought up through the copyright laws, but then try to build other structures such as institutional repositories and so forth that are global and universal in their concept, so that we are not niggling over individual titles or institutions and instead are trying to rephrase and define better channels for scholarship and dissemination of ideas.

Temple Claggett: Thank you, and I think that was more of a comment than a question. We went a little bit over. Thank you again to all of the panelists. This was a very good conversation.