

Preservación digital y derechos legales

Laura Campbell

Subdirectora de Biblioteca para Iniciativas estratégicas y Jefe de los Servicios de Información, Biblioteca del Congreso (EEUU)

Ha habido muchos retos en la adquisición, preservación y acceso al contenido digital durante nuestro trabajo bajo el Programa Nacional de Preservación e Infraestructura de Información Digital (NDIIPP). Se trata de una iniciativa nacional para crear una colección digital de contenidos creados originalmente en formato digital para la preservación a largo plazo trabajando con una red de organizaciones asociadas. Como parte del Programa Nacional de Preservación e Infraestructura de Información Digital emprendemos, junto con la Oficina de Derechos de Autor de Estados Unidos, un proceso formal de revisiones recomendadas para la sección de la ley de derechos de propiedad intelectual de Estados Unidos que trata lo que las bibliotecas y archivos pueden hacer para preservar el contenido digital. También descubrimos desde un punto de vista práctico las decisiones a las que las bibliotecas y archivos se enfrentan en la gestión de los contenidos digitales con respecto a los derechos y restricciones legales. Este breve artículo y la charla de la mesa redonda se centrará en la experiencia del Programa Nacional de Preservación e Infraestructura de Información Digital y planteará algunas cuestiones a considerar.

Introduction

Thank you. I am delighted to be with you today to talk about digital content—about creating it, preserving it, and making it accessible. This is one of my favorite topics because there is so much good news to report. Extraordinary technical innovations have made it possible to access information from almost anywhere on the planet, whether we are sitting in front of a computer, looking at the dashboard of our cars, or reading, writing, watching movies, listening to music, sharing photos. This access to information worldwide has enabled rapid advances in science, technology, and medicine, and greater access to education and culture, especially for those far removed from the wealthy cultural capitals of the West.

Libraries in your country and mine have taken the lead in putting high-quality cultural content online. I remember how inspiring it was to see the earliest high-quality, high-resolution scans of primary source documents—records about Columbus's voyages—scanned by the Archives of the Indies in Sevilla. This was long before most other institutions were putting documents online. I am proud that the Library of Congress has been a leader in this area, committed to making our cultural treasures as accessible as possible to citizens in America and of course throughout the world. For we are a nation of immigrants. Many of our most significant treasures are part of the global heritage of our citizens, who hail from so many different countries. We have the first book printed and published in the

New World, *Doctrine Breve* by Juan de Zumarraga, Christopher Columbus' *Book of Privileges*, and the first edition of *Don Quixote*.

For the past two decades we have been building a digital library at the Library of Congress. We started in 1992 with digitizing the unique treasures of the Library, photographs, manuscripts, diaries, maps, rare books, music and movies that tell the story of our cultural heritage. In the late 90's, we became concerned about capturing content that only exists in the digital form, and if not collected and saved would be lost to future generations. With funding from Congress, we embarked upon a national digital preservation program, called the National Digital Information Infrastructure and Preservation Program (NDIIPP). Today, we have saved over 300 terabytes of "at risk" digital content working with a network of 170 partners. It is through this work that I can share my views on an area where we have not made much progress: intellectual property.

Despite the progress made in technology, our intellectual property operating system was built for the information world of 200 years ago and has never been upgraded. Even though information is critical to government, business, technology, and culture worldwide, our intellectual property laws remain stranded in the world of 1810. They are unfit for today's world of 2010. Our obsolete copyright regimes create the greatest single challenge libraries face in fulfilling their historic mission—to collect, curate, and preserve knowledge and make it accessible to this generation and future generations.

Because this is our core business, libraries have a unique perspective on intellectual property. And we have unique responsibilities—not only to copyright owners, but also to those for whom intellectual property regulations were designed to benefit—the users.

For the rest of my time with you, I will talk about the *three dimensions* of this problem that I think are most important and hope that I can stimulate your thinking and this will lead to action.

First: The problem is urgent.

Second: The problem is difficult.

Finally: We can do something about the problem.

In fact, I will argue that not only *can* libraries do something about the problem. *We must* do something.

First, as I said: **the problem is urgent**. We are losing far too much valuable digital information because intellectual property regulations do not allow libraries to collect and preserve at-risk material for the public. Or—just as bad—libraries do not collect and preserve material out of fear and confusion about what is legal to do versus what is right.

For commercial materials that are born digital, such as motion pictures, popular music, news, photography—these are all potentially valuable at-risk digital content. Libraries do not collect and preserve these materials because the right to do so belongs exclusively to the copyright owner. In the United States, for example, when we license materials from content vendors, we lose our right to preserve that content, no matter how valuable. That means that the owners must preserve the material or it will not survive. But major news corporations, film studios, music producers—they are in the business of creating content and disseminating it. They are not in the preservation business and most of them do not want to be in the preservation business.

Even more challenging than commercially created cultural material is the content on the open Web, such as blogs, YouTube, social networking sites like Facebook and Tuenti. This is all content with

potentially important historical significance. It has been created collectively by many, many people. This is a new historical phenomenon—collective creation at this scale. If for no other reason than that, it cries out to be preserved. The problem is we think we can't, because there is no clear ownership. Libraries shy away from collecting this material—again, out of fear and confusion.

Take the example of Twitter, which the Library of Congress has agreed to archive for future users. We will work with Twitter to ensure this material will be available to researchers. We are not sure what that means yet, and to be honest, that is precisely why we are doing it. We can only understand what it means to archive this material by doing it. But no sooner had we announced our partnership with Twitter, then legions of people using twitter—let's call them the *twitterati*—asserted ownership of their own tweets!

The fun has just begun. Even though this is a complicated space, it is our judgment that it is better to act in time than regret the loss of this information at leisure. We hope this is an experiment that all libraries will be able to learn from.

While this open Web content may strike some of you as an extreme case, our work with the Blue Ribbon Task Force on the Economics of Digital Preservation and Access tells us otherwise. This task force brought together leading librarians, archivists, and economists to look at issues, such as IP, that affect the economics of archiving, a topic, incidentally, that has received scandalously little attention and study. It was the opinion of this expert group that collectively created content is not an exception to the rule of digital content. In fact, increasingly content will be collectively created. It is perhaps *the most* at-risk material because there are no clear paths of ownership and responsibility. For libraries this is a critical issue, and an urgent one. For if we are unable to provide stewardship for contemporary culture, no matter how ephemeral, we will have failed in our obligations to the next generation.

The second challenge is that the problem is difficult. What does that mean? It is *not* difficult because it is a technical problem. It is *not* difficult because libraries cannot develop the right standards for metadata. It is a difficult problem because it is not a problem that we, as libraries, control. Intellectual property regulation is public policy, determined by laws and regulations that libraries have little influence over. To address the problem, we have to step outside of our cozy world of day-to-day concerns about digital repositories, metadata schema, markup language—all things librarians like to work on.

It means that we must engage people outside of our community who are involved in public policy. We must become advocates, and we must make a compelling argument for the value that we create. And we must make that case until we are blue in the face.

That is difficult. Frankly, librarians take for granted that people understand our value and will support us, even in the toughest of times. But I believe that the era of unconditional public support—at least in the financial sense—is behind us. We cannot afford to take robust funding for granted anymore. We cannot take for granted that the upcoming generation of digital natives even know what libraries are, let alone care about what happens to them.

We must learn to navigate the often turbulent waters of public policy, and we must expect set-backs from time to time without giving up. Very early on in NDIIPP, our national digital preservation initiative, we identified intellectual property as one of the key barriers to digital preservation. So

imagine our delight when the Register of Copyrights, my colleague Marybeth Peters, now retiring, agreed that current copyright law needed to be changed for the digital era! I believe the way she put it to me at the time went something like this: “When people have to break the law to do their job, the problem is with the law, not with the people.” She worked tirelessly with us for two years to address one small part of the United States copyright code—Section 108—that deals with preservation exemptions for libraries. With the Register of Copyrights behind our effort, we thought we could make real progress, and we worked very hard. Our group agreed on set of very strong and fundamental recommendations. They were accepted by the Register and the Librarian of Congress, and put in the hands of our legislators.

And there it has rested for over a year. Because there are more pressing problems that are urgent: economic recovery is urgent, national security is urgent, disaster preparedness and recovery is urgent. Where does digital preservation fit in here?

You can see why this is a difficult problem.

So why am I confident that **we can do something about the problem?**

The reason is simple: *because we have to*. Our job is to ensure the transmission of knowledge on behalf of society. Unlike politicians or business people, we actually are in the business of thinking long-term. All throughout history libraries have been challenged by wars, famines, neglect, poverty, even political regimes that try to destroy a people by destroying their culture. But somehow we have prevailed. We have not saved everything that we’ve tried to save. But we have saved a very great deal.

But long-term means something very different today than it did even 25 years ago. This is another key lesson from NDIIPP: we must think long-term in five-year increments. Five years is a long time for digital content. What if we had built the perfect 100-year digital preservation solution in 2001, when NDIIPP started? Within 3 short years we would have failed—because of Web 2.0!

No, we cannot build for the ages and wait for things to slow down so we can get it perfect. That time is over. We must **do and learn**. We must get things good enough to try, and be ever ready to adapt to the next change. That is why the Library of Congress entered into its historic partnership with Twitter—*not* because we know how to archive this content, but because *we do not*. That is what it means to “think long-term” now.

We must change the way we think about preservation in another way: We must move from thinking passively about receiving content, to acting aggressively to get content and rescue it when need be. To do that, we must go where the content is, not wait for it to be offered to us as if we picking from a vendor’s list of book titles and checking off what we want to order. That is why we need copyright laws that *encourage* us to harvest from the Web and not wait for permission.

This also means that we must think of preservation as a *community effort*, not something each library does on its own. We need to build a shared infrastructure for digital preservation that leverages economies of scale worldwide and provides real value for every investment that we make in preservation. This means cooperating in a way that libraries have only paid lip service to in the past.

As you know, Americans tend to be competitive. This includes libraries. For too long research libraries at American universities have competed with each other like rival football teams. They compete to have the biggest collections with the biggest buildings. Today we must commit ourselves to

building a *shared collection*. The only way to *large-scale, sustainable* collections in the digital world is through collaboration, not competition.

If we can *cooperate to leverage each other's strengths*, there is almost no limit to what we can do. We can build distributed collections that allow each participant to focus on what is their area of expertise—geospatial content, music, social science data sets, whatever they specialize in—knowing that they can rely on other experts for the best digital collections in other fields. We can put together a shared service that does what the old-style copyright clearing house did in the world of books.

We should encourage our users to *assert fair use* when appropriate. As librarians we tend to avoid risk. And certainly at the Library of Congress, we have made sure that, like Caesar's wife, we have been above reproach in all matters having to do with copyright. This is appropriate for a national library. But we should encourage other libraries to act more boldly than we are able to. One of our NDIIPP partners, the Internet Archive, makes it a practice to archive what they think is of high value, believing that it is better to ask forgiveness than permission. Those of you who work in research libraries should advise patrons to assert fair use for educational purposes and avoid asking permission of the overly cautious general counsels that we find in higher education.

We can *partner with commercial companies* to extend the reach of stewardship by working with them to develop best practices, as the Library of Congress is doing with many creative industries in Hollywood. We can also partner with companies or groups that own content to ensure that when they no longer wish to provide stewardship of the content, it comes directly into a library that will assume responsibility for it. What would those partnerships look like? They would vary according to each situation, but they would start much like traditional library relationships with donors of content. Our librarians and archivists can easily identify what current materials have potential historical value. We can approach owners of these materials and begin working with them as equal partners. This is the approach we've taken with Twitter. The Library also has such relationships with some public television stations, and we are working on others.

Finally, we must *think globally and act globally*. Digital content does not respect national boundaries. No country can solve its own intellectual property problems in isolation from others. Working with other nations means learning about how other cultures view access and respecting cultural traditions. It means building trust among partners from different parts of the world, something that libraries have a centuries-old tradition of doing. Working across national boundaries does not mean that we must accept all limitations a sovereign nation puts on information without question. It means realizing that the decisions that we make for our local users have impact throughout the world. Our users are not just the people who live in our neighborhood, in our state or province, or in our country. Together libraries must meet the information needs of *global citizens*.

Closing thoughts

I have tried to give you a sense of why I think intellectual property regulation for digital preservation is an urgent problem. And also why it is a difficult problem—though I imagine you know a lot about that already! What I hope more than anything is that I have persuaded you that this is a problem that we can do something about. Libraries are uniquely positioned to make a difference, and libraries have

uniquely important *responsibilities*. If we *do not* address these problems, nobody else will. If we *do* address these problems, many people will join us. Libraries have always provided leadership in the public sphere. Now is not the time to pull back.

My biggest fear is that we will look at the problem of intellectual property in the digital world and be daunted by it, that we will think it is too big for us to tackle, that we will shrug our shoulders and say we have no power to change this.

I would like to end this talk by waving my hand and producing a PowerPoint with a list of five bullets telling us we need to do to solve this problem. But I can't. We all know that it takes more than just a formula. It takes our committed engagement. It takes our passionate advocacy. And it takes patience. Remember that the system we are building today is for the benefit of future generations, just as our predecessors labored to build the present system from which we reap benefits. I hope the conversation about how we can upgrade our intellectual property operating system from the 1810 version to the 2010 version can begin today, here, now.