Addressing the Copyright Law Barrier in Higher Education – Access-to-Clean-Content Technology in the 21st Century

by
Dr. Roland Vogl, Esq., CodeX - Stanford Center for Legal Informatics
Franny Lee, CodeX - Stanford Center for Legal Informatics
Martha G Russell, Media X at Stanford University
Prof. Michael Genesereth, CodeX - Stanford Center for Legal Informatics
The Problems

Imagine a university professor reads an article in a scholarly journal, which is directly relevant to that day’s class lesson. The professor uploads the article to the class’ e-learning management system. The reading stimulates a lively class discussion, and one student decides to use the article on a website he built as a class project. The article can now be accessed by anyone with Internet access, the publisher’s ability to sell the article is impacted, and the university is potentially liable for having contributed to copyright infringement.

TEACHERS’ AND STUDENTS’ COPYRIGHT PROBLEMS
A professor or student who wants to respect copyrights and follow proper procedures is often unable to because the paths are obscured and prohibitive. How could this professor have shared the article with his students without violating the publisher’s copyright interests? Similar questions apply to the student’s posting of the article on his class project website.

Many universities have subscription agreements with publishers that provide for access to scholarly content via various databases. The professor could check whether the article is available through university subscriptions. However, the agreements that govern subscription access are generally complex legal documents with differing terms and conditions of use, so the ability to access an article through a university-licensed database might not necessarily guarantee the professor can legally share the article with his students. Even if it is possible to share with students, the license conditions might not permit sharing with the greater online community.

Alternatively, he can buy a license to distribute the content to his students from the copyright owner or the Copyright Clearance Center (CCC), an agent for many publishers. It is a difficult process and he might not have a budget to pay for permissions for his entire class. Although highly impractical, he could try to collect the cost of an individual permission from each student to cover the amount. Moreover, although purchasing a license might provide some legal peace of mind, the question of whether he even had to pay in the first place (because of available university subscriptions) is still open.

To further complicate matters, if he only needs to use a certain portion of the article for his class, he might not need to request permission at all under the fair use exception in copyright law, or other legal reasons. How much can he use without violating publishers rights? Or, is permission not necessary because the article is public domain content with no restrictions on use?

PUBLISHERS’ COPYRIGHT PROBLEMS
For a publisher, the unauthorized distribution of content typically means lost revenue. An immediate goal of a publisher facing unauthorized distribution of his content may be to stop the unauthorized uses and prevent future unauthorized distribution, and it is easier to target universities as a whole rather than individual students. A further larger problem is that publishers are losing track of where and how their content is being used, so they have little or no data to rely on to properly price content, or know how or where to best create reliable revenue streams.

UNIVERSITIES’ COPYRIGHT PROBLEMS
From the viewpoint of a university – especially the university librarians, IT professionals, and university counsel - the fact that unauthorized distribution of content is facilitated by university employees and students on university-managed learning management...
systems is problematic and a potential liability. Furthermore, academia tries to share and grow knowledge. A system in which professors might refrain from sharing beneficial content with their students because of copyright risks is undesirable, particularly if such restraint also involves under-utilizing subscription resources a university has already purchased and having university communities members pay. Equally undesirable is the possibility that professors or students might pay again for content that is already licensed by the university.

The Copyright Permissions Maze

Some believe that “the current clearance process for educational uses of content can resemble a permissions maze calculated to prevent [new educational] uses.” Given the frustration and difficulty involved to get clear and authorized access to relevant content, educators frequently resort to using content without seeking permission. Sometimes such use of content is based on interpretation of the fair use exception for educational use, which given the overwhelming ambiguity involved, could be legitimate or overly broad. At other times, unauthorized use of educational content represents a deliberate disregard of copyright laws. In all probability, a good proportion of unauthorized use is due to common problems that arise too often, and create excessive amounts of friction that doom even good faith efforts to seek legitimate use of content to failure. Current processes force users to navigate through notoriously difficult and ambiguous legal questions. Additionally, ownership can be hard to identify and track down, or rightsholders might not have the resources to respond to and transact licensing requests, leaving educators in the position of having no express permission to share an article with the class. In cases where the professor might be successful in request efforts, for reasons of time and cost the process might still wind up prohibiting the sharing of current copyrighted materials with the class.

The Changing Uses of Content in Higher Education

The educational experience is no longer merely a combination of textbooks, lectures and exams. Digital formats and technological innovations are allowing educators, researchers and students to explore interactive, immersive environments that promote learning and creativity. Today, most universities use learning management systems, such as Sakai, Blackboard or Moodle to manage courses online and distribute content and class assignments to students. While digital learning environments have been rapidly evolving and improving, the ability to obtain permissions for the copyrighted content needed as class materials on these e-learning environments has not kept pace. Particularly so in these digital platforms, the path for how an educator might get the necessary permissions to distribute content is unclear or impractical. At Stanford, for example, the learning management system (Coursework) tries to direct professors to request permissions from copyright owners and braces them for a 4- to 6-week process, or to the CCC.

Figure A: Traditional copyright clearance practices in higher education
As future waves of educational tools further enable educators to distribute, mix together and create interactive multimedia content that enhances the educational experience, these innovations will have to be matched with ways to address and resolve copyright law barriers.

BREAKING THROUGH THE TRADITIONAL BOUNDARIES OF EDUCATIONAL INSTITUTIONS

In their 2006 whitepaper “The Digital Learning Challenge: Obstacles to Educational Uses of Copyrighted Material in the Digital Age,” Terry Fisher and William McGeeveran focus on new learning opportunities that could be enabled by digital technology, which can break through the traditional boundaries of educational institutions, such as high schools or universities. In 2006 when the article was published, some early efforts of sharing university classes beyond the campus were under way – most notably MIT’s OpenCourseWare project. Now, in 2012, other universities have also embarked on using the Internet to share with the rest of the world knowledge conventionally reserved for an exclusive university community. Stanford, for example, opened several popular computer science courses to the public in Fall 2011. Many courses support several tens of thousands of students – in one case over 160,000 students. Making educational content available to people unaffiliated with traditional educational institutions also raises new difficult copyright issues. Problems arise in particular when third party owned content is included because traditional subscription licenses between universities and publishers cover only university affiliates.

Due to copyright concerns, online learning platforms such as Coursera.com or Udacity.com currently refrain from distributing third party course materials along with the video lectures. Consequently, students typically do not get access to supplementary content to enhance their learning, and publishers miss out on the chance to sell content to an unprecedented number of customers.

The Current Copyright Clearance Process in Higher Education

THE BASIC PERMISSIONS PROCESS

In its simplest representation, the permissions process might be broken down as depicted below.

**Figure B: Traditional copyright clearance practices in higher education**
For example, to obtain permission for the simple traditional practice of making copies of an article for a class, the professor can seek permission from either the current copyright owner or authorized agent (such as the CCC):

i. by identifying the owner or confirming the agent’s authority to represent the content,
ii. locating the contact information,
iii. requesting permission to make the copies,
iv. being advised of any applicable royalty price,
v. collecting the royalty payments from the class or paying out-of-pocket, and
vi. delivering payment to the owner.

This process must be repeated for each article the professor wishes to use for his class, as well as for any different type of use of the same article the professor wishes to make. Typically, if the course is repeated in the future, permission has to be sought again. Substantial time and costs resources are involved even in cases where all steps were immediately successful.

It can be nearly impossible sometimes to identify whom to ask for permission to use an existing work, or locate the contact information for an identified copyright owner, or determine whether one even needs to ask permission at all. While going through an agent, such as the CCC, can facilitate the permissions process, it is still time-consuming and does not always lead to success. Some copyright owners have manual submission practices, such as faxing forms back and forth and 6-8 week minimum processing times. Copyright clearance time delays frequently force educators to use the same content over and over again, instead of updating for more current or relevant materials they may later encounter.

Also, copyright ownership can be complex in that a single article might be owned by numerous parties, such as each joint author, the publisher, and/or subsequent transferees. With no formal registration required before copyright protections come into existence, records from government copyright offices are usually incomplete or outdated. Similarly, there are no formal registration requirements when copyrights are shared by several authors or publishers and/or transferred to subsequent owners, and there are no authoritative centralized comprehensive databases available that track all of these details. Ownership data is therefore often obscure, unreliable and highly decentralized.

If ownership is indeterminable or there is no response to permissions requests, then the professor has no express permission to share the article with the class. In cases in which the professor is successful in his efforts, the permission may still wind up being prohibitive due to time and cost. Finally, fair use is another layer of complexity that envelops this process. Often, if the content taken is small enough or used for certain purposes, no permission is even required. The four factors in qualifying for fair use must be judged on a case-by-case basis, and unfortunately for professors and universities, even legal courts admit that the test is so difficult that litigation and a judge are often necessary to know if fair use applies. The recent trial decision involving Georgia State highlights the expense and risk to universities if professors rely on fair use incorrectly.

Given the high level of complexity, much of current copyright clearance practices are manual and labor intensive. Not surprisingly, copyright clearance intermediary services charge significant fees for their services, which carry forward to the final cost paid by students.

Figure C below demonstrates the rough breakdown of what a student pays when purchasing a traditional hardcopy coursepack.
Lost Revenues for Publishers

Educators increasingly use learning management intranets and online platforms to distribute materials to students, but the path for users on these systems to ensure legal access to content remains unclear. Royalties that might otherwise be paid are lost because the current process is too de-centralized and onerous, or perhaps because the user does not even know how to begin.

In addition, publishers lose revenue by not unbundling content. Publishers have traditionally relied on subscription models with university libraries. However, with library budgets under significant pressure, subscriptions accessed by a small subset of users are being dropped. Although some publishers might increase the price of more popular publications to make up for dropped subscriptions, the core problem is that many publishers are not positioned to monetize institutional per-article transactions. Furthermore, when considering the unbundled sales potential of online courses which boast student enrollment in the tens of thousands, the opportunity loss of unbundled revenues for publishers is significant.

The Duplicate Payment Problem

A core problem that exists for copyright transactions in academic institutions is the difficulty in identifying the content, and the uses of that content, that a school has already licensed for use by its community. As discussed above, negative results can occur: under-utilization of content, or alternatively, unnecessary duplicate payment of copyright royalty fees for the same subscribed content. In the university environment, access to content and authorization for various uses is typically granted through different avenues. For example, blanket licenses, with publishers, distributors, or copyright agents might be in place for the university community, each with a different set of rights and conditions. Many of Stanford University’s Libraries and Academic Information Resources’ (SULAIR’s) existing electronic access agreements also allow Stanford professors and students to make coursepacks with the same content. In addition, individual professors and departments might purchase permissions for their students through their department libraries. Furthermore, public domain and royalty-free licensed content (such as Creative Commons licensed content) is available to educators, researchers and students.

Even if a user is able to access the agreements underlying access to certain content, they still face the question of whether the license covers the type of use they intend to undertake. The information is often ambiguous or too legally complex for non-lawyers to evaluate. The problem is compounded when commercial services, with no access to these agreements and no direct incentive to spend extra time checking for pre-existing rights, are used to create
course materials, such as print or digital coursepacks. Consequently, students end up paying again to purchase copyright licenses for content and rights their university has already licensed on their behalf.

University legal departments and librarians engage in efforts to educate faculty about copyright and highlight the existence of purchased subscription access to content, to attempt to mitigate the double payment problem. SULAIR, for example, launches regular efforts to educate faculty on copyright issues and risks, and encourage use of content already licensed for the community.

“Because copyright is so deeply entwined with what we do here at Stanford, it is important that everyone in our community have a working understanding of the laws and regulations under which we all must operate. This memo provides guidance on those laws, and includes a listing of resources to consult for additional assistance in making use of copyrighted materials.

One example of an everyday campus activity with significant copyright implications, highlighted in the Copyright Reminder, is the creation and use of coursepacks. Stanford's University and Coordinated Libraries license considerable quantities of e-journals, e-books, and other networked academic information resources, and we urge you to take advantage of those resources through linking rather than licensing materials on your own.”

–Stanford University Head Librarian Michael A. Keller Email to Faculty of November 17, 2010

Lawsuits against Universities
As a result of educators frequently uploading copyrighted materials into learning management systems without purchasing permissions from copyright owners, tensions between publishers and their university clients have risen, escalating in some cases to formal and expensive litigation. The recent U.S. lawsuit Cambridge University Press v. Patton exposes a university's vulnerabilities regarding fair use and e-reserves practices. In that case, the defendant Georgia State University's policies and training of faculty regarding copyright issues did not protect it from liability. The far-reaching impact of the decision should have a persuasive effect on university practices generally, as the trial court reinforced state schools' liability for professors' copyright activities; confirmed that the fair use exception does not automatically apply for teaching or research activities by non-profit educational institutions; and established that prospective determinations on fair use (in advance of expensive litigation) are likely “futile”, because fair use is “notoriously difficult to apply” and too ambiguous – even very detailed policies on fair use cannot provide enough guidance.

Still, universities actively seek ways to help mitigate risk. Through its online learning management system (Coursework), Stanford provides some guidelines for where university-licensed content might be found, links to information about fair use, and links to information about requesting permissions from the CCC. Figure D shows the Coursework file upload pages, where the faculty member is asked to confirm the copyright status of the file to be uploaded. Figure E shows how faculty members have the ability to cause a Copyright Alert to appear when the file upload is accessed by others.
Unfortunately, none of the traditional tools available to faculty members provides a convenient and complete mechanism to assemble materials and ensure permissions compliance by downstream users.
The Solution – The Stanford Intellectual Property Exchange

The Stanford Intellectual Property Exchange (SIPX) is a copyright registry, marketplace and clearing engine that facilitates legal access to copyrighted works. SIPX has been deployed at Stanford University since April 2011. It first connected with a new print-on-demand system (PrintGroove) for coursepacks and is now also used through Coursework, Stanford’s online learning management system. SIPX’s real-time copyright clearance technology yields impressive cost-savings for students and a previously unknown level of convenience and "copyright clarity" for professors.

Background

SIPX is a research collaboration between Media X at Stanford University and CodeX – The Stanford Center for Legal Informatics, a research center jointly operated by Stanford Law School and the Stanford Department of Computer Science. Through years of rigorous research, SIPX was developed by CodeX’s interdisciplinary team of computer science, legal and business researchers. SIPX is based on:

- a rights repository that allows users to register content they own,
- an exchange of rights that allows users to buy and sell copyrights, and
- rights verification that assists users in knowing what rights they have to use content.

Figure F: Graphic representation of SIPX system within course material setting

SIPX USE CASE: Distribution of Course Materials

Stanford IP Exchange (SIPX) makes it clear, quick and cost-effective to distribute copyrighted materials the right way.

Minimize legal liability while respecting the rights of content creators.
SIPX has been used on the Stanford campus since Spring Quarter 2011 to facilitate educational electronic and print materials. It enables professors to easily handle copyright clearance for materials included in both electronically-distributed educational materials as well as hard-copy coursepacks. The first SIPX pilot connection was with a new print on demand production system developed by Media X industry partner Konica Minolta, that allowed students to print single customized copies of their coursepacks in book quality. During this initial deployment, participating professors assembled their course materials online through Konica Minolta’s PrintGroove platform. Students customized and purchased their course materials online through PrintGroove also, which procured permissions in real-time through a SIPX API, and linked to Konica Minolta’s BizHub production system installed on campus. Hard copy coursepacks are now distributed through the Stanford Bookstore operated by Follett Corporation. In this most recent Spring Quarter 2012 deployment, SIPX supported course materials for a variety of different classes totaling over one thousand Stanford students, including classes using the Stanford’s e-learning management system Coursework.

**License filtering**

The SIPX system is the result of research focused on understanding the requirements to develop an efficient mechanism for translating complex legal provisions into computer-processable data (in other words, ‘codifying’ copyright and contract law). SIPX’s design is based on an approach coined “license filtering.” License filtering means that when a request for copyright clearance is sent to SIPX, SIPX first checks its copyright database to see whether the user already has the rights needed to use the requested content and activity. In this sense, SIPX is a user-centric system that dynamically filters for pre-existing rights from which a user may already benefit. If no royalty-free authorization is identified after SIPX sifts through these preliminary filter layers, SIPX then communicates to the user the pricing and licensing conditions under which the copyright owner is willing to grant the necessary permissions. SIPX also works in close collaboration with SULAIR, particularly to design and refine the Stanford subscription filter layer.

First, SIPX checks for authorizations granted by law, such as public domain materials, in which the statutory term of copyright protection for the content has expired thus making the content generally free for the public to use without restriction. If the material is not public domain, SIPX checks for royalty-free conditional licenses, such as whether the content is available under Open Access or Creative Commons licenses. Then, SIPX checks for any pre-existing contractual arrangements which may cover the user and desired uses, such as licenses purchased earlier by the user or subscriptions purchased by university libraries on the user’s behalf. Finally, if SIPX cannot identify a subscription benefitting the user, it enables the user’s royalty purchase transaction.

The effects of license filtering have been remarkable. Two results are most notable:

1) License filtering eliminated duplicate royalty payments and drastically reduced student cost for course materials, and

2) License filtering replaced elements of manual clearance with real-time automation and drastically reduced the time required for educators to assemble course materials.

In the four SIPX deployments on Stanford campus since Spring 2011, these results were empirically measured. Of the copyright royalty cost component in student coursepacks, SIPX reduced the amount by 4%-78% for participating pilot courses. On average, each pilot class was able to offer its coursepack for approximately $30 (35%) less.
License filtering not only yielded significant cost-savings but also significant time-savings. The measurements above were calculated using the number of “ownership contacts” (the content chosen for a course is owned by different publishers and rightsholders, each referred to as an “ownership contact”.) As explained above, the basic process for identifying, locating and requesting permissions from an ownership contact can entail a significant amount of work and frustration. Because SIPX takes pre-existing rights into account, the number of ownership contacts that require manual copyright clearance efforts is reduced. Figure H presents a comparison summary between the manual ‘ownership contacts’ required under traditional methods of copyright clearance and the SIPX method of copyright clearance. The number of manual ownership contacts under the SIPX method decreased between 17%-100%, because it identifies pre-existing rights that eliminate the need for a manual contact.22

By fully leveraging existing purchased library subscriptions and eliminating double royalty payments, the cost savings for educational materials (at research universities of similar resources and size as Stanford, ~15,000 students) could be upward of $720,000/year.24 In addition, SIPX provides cost-savings in the form of reduced staff time necessary to handle copyright permissions.

For university libraries and IT managers, SIPX provides a frictionless way to avoid duplicate payments for content paid for by libraries and an easy way to make non-subscribed content available a la carte. Finally, SIPX empowers professors with copyright clarity and reduces litigation risk for universities.

---

**Figure G:** Summary comparison of royalties paid to copyright owners in coursepacks (Spring 2011 pilot)

<table>
<thead>
<tr>
<th></th>
<th>Traditional method</th>
<th>SIPX method</th>
<th>Total cost savings to student in royalty cost component ($)</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot 1 (Econ)</td>
<td>$44.87</td>
<td>$0</td>
<td>$44.87</td>
<td>100%</td>
</tr>
<tr>
<td>Pilot 2 (Psych)</td>
<td>$42.34</td>
<td>$36.55</td>
<td>$5.79</td>
<td>14%</td>
</tr>
<tr>
<td>Pilot 3 (Physics)</td>
<td>$29.63</td>
<td>$2.88</td>
<td>$26.75</td>
<td>90%</td>
</tr>
</tbody>
</table>

**Figure H:** Summary comparison of ownership contacts in coursepacks (Spring 2011 pilot)21

<table>
<thead>
<tr>
<th></th>
<th># of manual ownership contact efforts required under traditional method</th>
<th># of manual ownership contact efforts required under SIPX method</th>
<th>Ownership contacts eliminated by license filtering ($)</th>
<th>(%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pilot 1 (Econ)</td>
<td>13</td>
<td>4</td>
<td>9</td>
<td>69%</td>
</tr>
<tr>
<td>Pilot 2 (Psych)</td>
<td>14</td>
<td>11</td>
<td>3</td>
<td>21%</td>
</tr>
<tr>
<td>Pilot 3 (Physics)</td>
<td>3</td>
<td>1</td>
<td>2</td>
<td>67%</td>
</tr>
</tbody>
</table>
Available Content

SIPX provides greatest value to users when it can easily facilitate access to content and seamlessly transact the necessary permission for any request. In other words, more content means more SIPX benefits for users. SIPX is building a critical mass of content to provide access to hundreds of millions of individual articles and works. SIPX enables this experience in the following way:

- Providing deep computable copyright knowledge layer on top of existing university subscriptions: SIPX leverages the entire catalogue of a university’s subscription agreements and adds clarity about authorized uses under the subscriptions (i.e. verifying that an article can be included in a printed coursepack and/or distributed through the e-learning management system).
- Content Owner and Aggregator Partnerships: SIPX has partnerships and API connections with publishers, copyright agents, such as the CCC, and content-hosting intermediaries. This content can be easily accessed through SIPX, and if SIPX finds there are no pre-existing rights and a purchase is necessary, each student/user can pay individually or charge the transaction to their own tuition account.
- User Generated Content: Creators and authors can register their own content and make it available under their preferred licensing terms and conditions.
- Public Domain and Open Access Content: SIPX builds partnerships with public organizations to harvest quality content freely available for users.

How to Bring SIPX to your Campus?

SIPX is a sophisticated but simple and easy-to-use system that integrates with existing learning management systems. It provides for tremendous cost savings and reduces legal risks. In Spring 2012, with the encouragement of Stanford University, SIPX incorporated as a company to extend its technology and functionality to other universities.

FOR MORE INFORMATION, PLEASE VISIT HTTP://SIPX.STANFORD.EDU OR EMAIL SIPXINFO@GMAIL.COM.
Endnotes


2. See supra, note 1.


4. Some forward-looking publishers recognize this opportunity. For example, Morgan Claypool Publishers offers content at discounted rates to participants in Stanford’s online courses and has had much success in sales numbers (May 17, 2012 interview with Michael Morgan, President).

5. Interviews were conducted by CodeX with Stanford professors in Fall 2010 as part of the Stanford Publish on Demand research theme.

6. Currently, several registries of bibliographic information that cover large areas of content exist; however, each is conditioned upon subscription fee or restrictions on use. Companies like Bowker (http://www.bowker.com/), for example, provide comprehensive bibliographic databases of published books that can be accessed for a subscription fee. The ISSN Institute provides a registry of bibliographic information about published journals, also available for a subscription fee.


10. Stanford’s Sakai platform Coursework supports over 1400 courses.

11. See supra note 9.

12. These agreements can exist independently or they can be the result of an institutions membership of a larger library consortium, such as the NERL (NorthEast Research Libraries) consortium for example; see http://www.library.yale.edu/NERLpublic/ (last visited May 30, 2012).

13. See supra note 9.

14. http://creativecommons.org/licenses/

15. Such as AcademicPub.com, Xanedu.com, UniversityReaders.com, or University Custom Publishing.


17. Supra note 16.


20. Supra note 8, at p. 9.

21. This pilot coursepack was offered through the SIPX copyright clearance system only. Royalty cost comparison data is hypothetical and based on quotes obtained from CCC.

22. This range is based on all courses that have participated in SIPX since Spring 2011.

23. Supra note 8, at p. 8.

24. In the ten different Stanford courses SIPX has supported to date, the average cost savings were $30. There are about 24,000 hardcopy coursepacks produced at Stanford every year. The hypothetical total savings accomplished by license filtering for hardcopy coursepacks alone would therefore be $720,000. This figure does not take into account the significant additional savings accomplished from courses using e-reserve and learning management systems to distribute their class materials.
The SIPX technology is based on many years of research conducted by CodeX - the Stanford Center for Legal Informatics, and supported through strategic industry partnerships of Media X at Stanford University. This research continues to explore human and technology issues related to advanced media technologies, new content and copyright challenges in education, business and commerce.

For information about CodeX - Stanford Center for Legal Informatics, please contact Dr. Roland Vogl, Esq., rvogl@law.stanford.edu; or Prof. Michael Genesereth, genesereth@stanford.edu.

Media X seeks strategic partners and research collaborators for requirements definition and use case projects; contact Martha Russell, Martha.Russell@stanford.edu.

SIPX seeks to share its technology with educational platform and content partners and make the SIPX technology available to other campuses; contact Franny Lee, FSLEE@stanford.edu.

For information about Konica Minolta Business Solutions that use SIPX, contact:
Chris Bilello, cbilello@kmbs.konicaminolta.us.