Four Facets of Privacy and Intellectual Freedom in Licensing Contracts for Electronic Journals

Alan Rubel and Mei Zhang

This is a study of the treatment of library patron privacy in licenses for electronic journals in academic libraries. We begin by distinguishing four facets of privacy and intellectual freedom based on the LIS and philosophical literature. Next, we perform a content analysis of 42 license agreements for electronic journals, focusing on terms for enforcing authorized use and collection and sharing of user data. We compare our findings to model licenses, to recommendations proposed in a recent treatise on licenses, and to our account of the four facets of intellectual freedom. We find important conflicts with each.

In July 2011, the U.S. Attorney for the Eastern District of Massachusetts issued an indictment for Aaron Swartz, an Internet activist and fellow at the Harvard Safra Center. The indictment alleged that Swartz had used a guest account to log in to servers at the Massachusetts Institute of Technology (MIT) and run a script to automatically download thousands of articles from JSTOR. After the activity was detected and his access blocked, Swartz took various measures to continue the downloading, including masking his IP and MAC addresses and, later, entering a server closet and directly connecting his computer to MIT servers. Based on these allegations, Swartz was charged with multiple federal crimes, among them violating the Computer Fraud and Abuse Act (CFAA).[1]

The case received substantial media attention at the time of the indictment, again in the wake of Swartz's suicide in January 2013, and once again in July 2013 when MIT released a report on the case and MIT's involvement.[2] Two features of the indictment are of particular interest for this paper. First is the basis for the charges. Swartz's activity violated the terms of service for guest accounts at MIT and violated the terms of MIT's license contract with JSTOR, which among other things prohibited automated downloading and downloading all articles in single issues of journals. Such unauthorized access (or exceeding of authorized access) to "protected computers" (in other words, computers used in interstate commerce) provides grounds for charges under the CFAA.[3] Hence, the case illustrates the potential reach and importance of license terms.[4] Second, the actions leading up to the indictment involved monitoring of MIT's Internet traffic to identify Swartz and understand his actions, illustrating that library

Alan Rubel is Assistant Professor in the School of Library and Information Studies, Program in Legal Studies, at the University of Wisconsin, Madison; e-mail: arubel@wisc.edu. Mei Zhang is a PhD student in the School of Library and Information Studies at the University of Wisconsin, Madison; e-mail: mzhang48@wisc.edu.

doi:10.5860/crl.76.4.427 crl14-599

patron access to licensed electronic resources carries different privacy implications than (for example) print.

Although the Swartz case is an extreme example of unauthorized access to licensed materials, the case illustrates the centrality and scope of licenses for electronic journals in academic libraries. And in light of the case, it is worth asking just how licenses affect library patron privacy, whether any of those effects are problematic, and, if so, why?

This paper addresses each of these questions. We begin by examining whether, and why, we should care about library patron privacy in the first place. The Library and Information Studies (LIS) literature links privacy to the concept of *intellectual freedom*. But precisely what intellectual freedom is, and how it relates to privacy, require some clarifying. Drawing on the LIS literature on privacy and intellectual freedom and the philosophical literature on freedom, we distinguish four facets of intellectual freedom. We then turn to the question of how licenses treat patron privacy by performing a content analysis of 42 license contracts between publishers and academic libraries.

We analyze our findings by, first, comparing the language of the licenses to model licenses proffered within the library profession. We find that the licenses in our data set differ in important ways from the model licenses. We then compare the terms of the licenses to the recommendations set forth in Tom Lipinski's recent treatise on library licenses for electronic resources.[5] Here, too, we find important differences between the licenses in the data set and the recommendations. Next, we examine whether license provisions implicate the four facets of intellectual freedom we describe. We find that the licenses conflict with only certain aspects of intellectual freedom. We conclude with some recommendations for changes to licenses.

The paper expands the existing scholarship on licensing in several ways. First, it explicitly links license terms to two central library values—privacy and intellectual freedom. Second, it links those core library values to the philosophical literature on privacy and freedom. Third, it shows a disconnect between the model licenses proffered by professionals and licenses that actually govern electronic journal access.

Privacy and Four Facets of Intellectual Freedom

ALA, Privacy, and Intellectual Freedom

Privacy is a core library value, and it has been for some time. The first iteration of the American Library Association's (ALA) *Code of Ethics*, adopted in 1939, posited that librarians have an "obligation to treat as confidential any private information obtained through contact with library patrons."[6] The most recent version of the code is similar, stating that librarians aspire to "protect each library user's right to privacy and confidentiality with respect to information sought or received and resources consulted, borrowed, acquired or transmitted."[7] Although the *Library Bill of Rights* does not explicitly list rights to privacy and confidentiality, the ALA's interpretations of the Bill are unequivocal that it serves as a basis for privacy protections. In "Privacy: An Interpretation of the *Library Bill of Rights*," the ALA maintains that privacy and confidentiality are "integral" to the mission of libraries, and that rights to privacy and confidentiality are "implicit" in the guarantee of free access to all users in the *Library Bill of Rights*. It further explains that privacy rights are rights "to open inquiry without having the subject of one's interest examined or scrutinized by others," and that confidentiality demands that libraries keep "personally identifiable information about users … private on their behalf."[8] A separate interpretation specifically covering academic libraries places even greater emphasis on privacy: "The privacy of library users is and must be *inviolable*."[9]

These statements of general principle are further specified to provide guidance in concrete cases. For example, the ALA explains that libraries have the responsibility

to protect not only patrons' personal information, but also "database search records, circulation records, and other materials that identify a person's use of library materials, activities, or facilities."[10] Moreover, the ALA interprets personally identifiable information expansively to include "any data that can link choices of taste, interest, or research with a specific individual."[11] Protecting privacy reflects a "long-standing commitment to an ethic of facilitating, *not monitoring*, access to information."[12] The commitment to privacy demands more than simply restricting others' access to patron information. It also requires that users be informed of "what policies and procedures govern the amount and retention of personally identifiable information, why that information is necessary for the library, and what the user can do to maintain his or her privacy."[13]

The ALA's commitment to protecting user privacy parallels language in other professional association codes of ethics. For example, the Canadian Library Association (CLA) states that member libraries have the responsibility to "protect the privacy and dignity of library users and staff."[14] The International Federation of Library Associations and Institutions (IFLA) *Code of Ethics* affirms that librarians should respect and protect users' privacy, by requiring librarians to "take appropriate measures to ensure that user data is not shared beyond the original transaction."[15] In her analysis of the library association codes of ethics from 28 countries, Pnina Shachaf found that privacy is one of the "most global and common principles in the LIS profession." Specifically, she determined that 85 percent of the codes of ethics in her study contain privacy protection.[16]

The ALA and other professional organizations justify privacy protections by appealing to intellectual freedom.[17] In the ALA's words, intellectual freedom requires that one have the ability to "seek and receive information from all points of view without restriction" and "true freedom of inquiry" does not exist where "users recognize or fear that their privacy or confidentiality is compromised."[18] The *IFLA Statement on Libraries and Intellectual Freedom* also claims that one principle of freedom of expression is to "not disclose the identity of users or the materials they use to a third party."[19] A number of other library associations understand privacy protection to be a key component of intellectual freedom, including those from Croatia, Japan, New Zealand, Turkey, and United Kingdom.[20]

Our discussion here focuses on privacy as it relates to intellectual freedom, following the primary justification for privacy protections set out in professional codes of conduct in the United States (where the licenses in our data set come from) and in much of the LIS literature. There are, however, a number of other possible moral foundations for privacy protections explored in the literature. For example, privacy may be instrumentally valuable as a part of personal relationships, mental health, professional transactions, and avoiding fraud; privacy may facilitate democratic processes; privacy may be a key component of human dignity; and privacy may be crucial in ensuring that important decisions about one's life and projects be made for one's own reasons and values.[21] Many of these justifications for privacy protections overlap in important ways with intellectual freedom.

Moreover, a number of commentators have argued that at least some aspects of privacy are of *negative* value. Richard Posner, for one, argues that privacy protections for individuals serve primarily to hide discrediting information that others could put to economically beneficial use.[22] Others have pointed out that privacy protections have historically been deployed to the detriment of women and other minority groups by shielding oppressive actions in "private" spheres (such as the home or private commercial enterprises), calling into question either the value of privacy per se or traditional interpretations of the nature and proper scope of privacy.[23] Further, regardless of the underlying reasons for privacy's value (if any), there is always a question of the degree

to which other goods warrant diminishing privacy. In recent years, for example, there has been significant debate about privacy and its relation to security,[24] and there is a perennial question in U.S. law about the proper scope of privacy protections in the enforcement of criminal law. Addressing the full scope of the privacy literature is well beyond what we can do here. Nonetheless, because the connection between privacy and intellectual freedom is drawn so explicitly in the LIS literature, and because many accounts closely link privacy to freedom (either intellectual freedom or freedom more broadly), that is the focus of this project.

Four Facets

What, though, is intellectual freedom, and how does privacy affect it? While there appears to be some consensus that intellectual freedom includes freedom with respect to thought, belief, access to information, and exposure to ideas, there is not a settled view regarding what *freedom* means in this context.[25] Positing that freedom just means lack of restriction does not get us very far, as that would leave open the question of what counts as a restriction in the relevant sense. Perhaps the lack of a settled view shouldn't surprise us, as freedom is a deeply contested concept with a large philosophical literature. However, it is crucial that we make sense of it in order to address novel issues of intellectual freedom such as those under consideration here. Comparing the statements linking privacy to intellectual freedom in the professional literature to accounts of freedom in the philosophical literature, we discern four facets of intellectual freedom that are implicated by privacy and privacy loss, each grounded in the LIS literature and each corresponding to a philosophical conception of freedom itself.[26] These will serve as a basis for our discussion of license content.

Negative Freedom

The starting point for understanding the four facets of intellectual freedom is to look at the most easily understood sense of freedom in the philosophical literature—negative freedom. Negative freedom is the freedom from *external constraints*. External constraints are restrictions on, costs to, or harms resulting from activities that are imposed by others and that limit a person's ability to act in certain ways—in this case engaging in intellectual pursuits.[27] For example, a law forbidding one from selling food without complying with health code regulations is a limitation on one's negative freedom insofar as it imposes a constraint on a person's ability to sell food that is imposed by others and independent of facts about the person who might sell food without compliance. Such a restriction is justified by health and safety outcomes; nonetheless, it is a restriction of freedom that demands some kind of justification. In the context of libraries and intellectual freedom, we can see that censorship and Internet filtering diminish negative intellectual freedom insofar as they are restrictions imposed by others on persons' abilities to access information.

Loss of privacy does not by itself limit this facet of intellectual freedom. Rather, privacy may bear upon negative intellectual freedom instrumentally, as when others use information gleaned about a person's intellectual pursuits to that person's detriment. For example, if Francis's employer learned of his research into a serious medical condition, the employer might infer that Francis is prone to that condition and deny him some opportunity or position of responsibility.

Positive Freedom

Not all of the concerns raised regarding privacy and intellectual freedom can be explained by the negative conception. For example, the ALA maintains that "true freedom of inquiry" does not exist where "users *recognize or fear* that their privacy

or confidentiality is compromised."[28] Likewise, a number of scholars have pointed to the potential for privacy loss to "chill" inquiry, regardless of whether persons who are monitored suffer any externally imposed, negative consequences from their loss of privacy.[29] But chilling effects based on fear of privacy loss are not constraints imposed by others. Rather, they are based on a person's psychology, which is *internal* to the person. To account for such effects, we can appeal to a different philosophical conception of freedom: positive freedom. Positive freedom refers to persons' abilities to assert control over their lives regardless of whether there are externally imposed constraints. Often this is understood in terms of persons' abilities to act according to their "higher" or more stable desires or the values of their more rational selves.[30] So, an addict may be unable to control his actions when he needs a cigarette, despite his actual desire to not smoke and to do something other than go looking for a cigarette, and would to that degree lack positive freedom.

Apprehension of privacy loss and chilling effects based on beliefs regarding privacy loss are limitations on internal, positive intellectual freedom insofar as they do not rely on others actually imposing limitations on a person's actions. Suppose, for example, that Jo believes that her library browsing and borrowing records are monitored. She worries that her coworkers, employer, or community would think ill of her if they knew that she reads controversial materials or (more prosaically) that she reads maudlin fiction. As a result, she does not browse or borrow such materials, even though she would like to. In that case, her activity is chilled not by constraints imposed by others, but by her own psychology. Hence, it is a limitation on positive intellectual freedom. Note that if her fears are the result of harms to her interests based on others actually thinking ill of her, the constraint would be external and hence a limitation on negative freedom.

Freedom and Autonomy

Other features of intellectual freedom cannot be explained by appeal to external, negative freedom or internal, positive freedom. Consider the ALA's insistence that library patrons should be "informed what policies and procedures govern the amount and retention of personally identifiable information, why that information is necessary for the library, and what the user can do to maintain his or her privacy."[31] Merely informing patrons of how their information is collected does not protect their privacy per se. At most, it places the onus on patrons to take steps to protect their privacy, if they are willing to incur the opportunity costs of doing so. Further, it does not protect patrons from external constraints should they fail to maintain their information privacy. Nor does it protect internal positive freedom. Informing patrons of how their information may be collected or used could actually help *cause* chilling effects by making them "recognize or fear that their privacy or confidentiality is compromised."[32]

Moreover, appeals to the first two facets of intellectual freedom cannot explain two seminal conflicts regarding library patron privacy. One is the U.S. Federal Bureau of Investigation's Library Awareness Program. In the late 1980s, the public learned that the FBI had a program in which agents would inquire at libraries about "suspicious" persons' uses of materials and services and request circulation records. In response to the program, the ALA's Intellectual Freedom Committee adopted formal policies regarding patron confidentiality and sought (though never received) in-depth information about the extent of the program.[33] The other is the USA Patriot Act's expansion of "business records requests" pursuant to the Foreign Intelligence Surveillance Act. Section 215 of the Patriot Act allows federal agents to not only make records requests at libraries, but prohibits recipients from disclosing that such requests have happened.[34]

In both of these cases, a key issue for the library profession is the *secrecy* of the surveillance. But secrecy does not necessarily undermine negative freedom; it does

not make it any more likely that persons will suffer external constraints based on their library records. Moreover, the fact that surveillance is secret may actually mitigate any chilling effects of surveillance by making people less likely to believe that they are being surveilled.[35] Here it is worth distinguishing secrecy of particular acts of surveillance, secrecy of surveillance programs, and secrecy of legal authorities for surveillance programs. Knowledge of any of these could affect people's behaviors, and the mere fact that, say, particular acts of surveillance are secret may not mitigate the effects of a known surveillance *program*. The key point here, though, is that there does seem to be an affront to freedom where persons surveilled do not believe (and hence do not know) that they are being surveilled.

To explain how informing patrons about retention and use of personally identifiable information and government powers to conduct surreptitious surveillance bear upon intellectual freedom, we need to articulate a third facet of intellectual freedom. John Christman has advanced the view that freedom is not simply a matter of absence of constraints, but a *quality of agency*.[36] On this view, a person is free only if she acts autonomously. Precisely what autonomous action demands is itself deeply contested, but it at least demands that persons be able to act according to their values as they see fit. Internal and external constraints will matter to whether persons act autonomously—coercion, threats of harm, and inability to act on one's more rational desires are all ways by which persons are limited in their ability to exercise autonomy.

However, autonomy demands more than simply being free of coercion and threats and having the wherewithal to act on one's higher-level desires. For one, it requires information important in determining how to steer one's choices to comport with one's values, and it demands information important in making sense of how one is being treated.[37] Consider, for example, a person who wishes to eat only foods that are not derived from genetically engineered (GE) organisms. This person is not forced into eating such foods, would be able to resist foods with ingredients derived from GE organisms, and has sufficient resources to purchase and prepare foods that are free of such ingredients. Suppose, though, that foods with GE ingredients are not labeled as such and that labels claiming foods do not contain GE ingredients are prohibited. There is an important sense in which one is not free to avoid GE foods because one does not have sufficient information to act on her desires as she sees fit. This is because she cannot exercise her autonomy in this regard.[38]

Interpreting intellectual freedom as requiring that persons be able to act autonomously can explain the importance of revealing to patrons the policies governing collection and use of their information. Likewise, it can account for why secrecy surrounding Patriot Act business records requests and the FBI library surveillance program impinge on intellectual freedom, even if patrons never learn that their records have been accessed. That is because information about the possibility of one's information being gathered is important to many people, and having relevant information is an important facet of autonomy.

Republican Freedom

The fourth facet of freedom relevant to our analysis is freedom from arbitrary exercise of power, sometimes referred to as *republican* freedom. The key idea is that, while one may be free in the sense of not having external constraints, being able to act in accord with one's higher-order desires, and exercising autonomy, one might still lack an important facet of freedom where one could be subjected to arbitrary constraints or harms to one's interests on the basis of another's whims.[39]

Consider the example of a person living in an area that is largely controlled by organized crime. Legitimate authority is ineffective, businesses open and close at the

discretion of the criminal organization, people are dependent on the organization for money, and so forth. Now suppose that one is able to navigate life pretty well because one is in the good favor of the mob, can predict the behavior of the organization, and so forth, and that as a result one can do largely what she wishes. Hence, one would appear to be free on each of the conceptions described so far. Nonetheless, one could at any moment be subjected to harms based on the arbitrary whims of the criminal organization. A key insight of the republican conception is that such potential subjection to arbitrary domination is antithetical to freedom.

This fourth facet of freedom finds support from the ALA, which quotes Bruce Schneier's view that monitoring creates the potential that "patterns we leave behind will be brought back to implicate us, by whatever authority has now become focused upon our once-private and innocent acts."[40] In other words, privacy loss need not chill inquiry to be problematic; rather, it is problematic where our intellectual inquiry—our "innocent acts"—can be used to our detriment, even if they are not so used.

Frank Lovett's prominent account posits that republican freedom is undermined where one party wields an imbalance of power over another, the party with less power would incur costs from exiting the relationship with the other party, and there is an absence of known rules or conventions governing use of power in that relationship, such that one party may wield that power arbitrarily.[41]

Precisely what constitutes arbitrary power sufficient to undermine this facet of freedom is the subject of debate. However, one threat to republican freedom is broad and vague law that others can deploy to a person's detriment. As Braithwaite and Pettit state in their seminal book on republican freedom and the criminal law, "[i]f the criminal justice authorities are not bound by *precise criminal laws*, then their power is relatively unchecked and there is a threat to the subjective component of dominion."[42] This relatively unchecked power is closely related to Lovett's absence of rules condition. That is, precise and tailored criminal laws constitute rules governing use of prosecutor discretion. As we will explain later, the possibility of liability under the CFAA for exceeding authorized access of electronic resources appears to suffice for subjecting persons to arbitrary power.

Authentication and Authorized Use

Although the focus of this paper is on privacy and intellectual freedom, it is set against a backdrop in which licensors seek to limit access to copyrighted works to a group of authorized users and for a range of authorized uses and in which various technologies are deployed to control use. It is therefore worth pausing to clarify several aspects of authorized use and authentication.

To begin, the line between authorized and unauthorized uses, especially for users who are unfamiliar with licenses, is not always clear. Despite provisions defining authorized user and permitted/prohibited uses explicitly, there will inevitably be uses that are either unclear or are routine but unauthorized. For example, a person might be an authorized user in one capacity but not another, as when a student or faculty member works for a private company (including her own venture) "after hours" and her access advances that work.

In other cases, uses that are widely accepted in scholarly communication might be prohibited by certain license provisions, which could easily lead to unauthorized uses. For example, fair use principles plausibly allow authorized users to share copyrighted work with research partners at other institutions (who are not authorized users); however, it is not uncommon for publishers to restrict sharing licensed materials with unauthorized users,[43] since contract provisions "often take away fair-use or other rights which would otherwise exist under the copyright law."[44] As Eschenfelder et al state,

licenses in some cases "forbid activities that many end users would consider morally unproblematic," such as occasional sharing with others (scholars, friends, family) not authorized under a particular license.[45] Hence, where an authorized user is unaware of such a provision, she can easily breach the licenses inadvertently, and thus the user might have to face punishments from publishers or even charges of federal crimes based on the CFAA. While one major prerequisite of such charges is publishers' ability to identify a certain user who conducts unauthorized uses, the authentication technologies libraries use may open an effective channel for publishers to obtain users' personal information.

Consider the two authentication technologies most commonly used by libraries today: IP filtering and proxy servers. Specifically, consider the types of user information collected and maintained by libraries, by publishers, and by other relevant parties and how the technologies may affect user privacy.

IP filtering is often used for on-campus users, where publishers compare the IP address of incoming requests with the IP addresses authorized under licenses and allow access only from those within an approved range.[46] This allows publishers to capture a certain campus IP address and activities (for instance, databases accessed and articles downloaded) associated with that IP address. Campus IP addresses can be either static or dynamic. A computer with a static, or fixed, IP address (for example, in a faculty office or in a lab) is easy for a campus IT department to identify. A computer with a dynamic IP address, such as a student's laptop connected to a campus's wireless network, is randomly assigned to some available IP address each time it connects to the network. Although its IP addresses might vary from time to time, a campus IT department can generally track down a particular computer with some investigation. For instance, it can match activity from a particular session with either the username/password used to log into the wireless network, or it can associate session information with a particular computer's MAC address.

Proxy servers envelop users' requests within approved IP addresses. Thus, where proxy servers are used, the IP address communicated to a publisher is a university IP address, and not the IP address from which a user request originates (such as from home). Proxy servers are often deployed for off-campus users, but some libraries (including the authors' home institution) also use proxy servers for their on-campus users. For off-campus uses, the proxy server requires users to provide a valid campus ID and password for the purpose of authentication.[47] Thus, libraries' proxy server log files collect information about users' IDs/passwords, activities, and the mapped proxy IP address sent to publishers. For on-campus uses, libraries' proxy servers keep users' static/dynamic IP addresses and their activities in the log files, and a mapping between the users' IP addresses and the IP addresses received by publishers' server. In both cases, publishers only see that the IP address comes from the acceptable range without knowing the "real" IP address of a certain activity and thus cannot identify individual users based on the proxy IP address.

So, even where IP filters and proxies are in use, information may be collected by libraries, publishers, and campus IT departments. Publishers may obtain such information from universities to the extent that they have the ability to require libraries and campus IT departments to provide user information. That can happen either because licenses provide that ability or via court order. To reduce such privacy risks caused by authentication technologies, many libraries purge their log files on a regular basis.[48] Washing server log files on the campus IT side would also be necessary to prevent publishers from identifying a certain user through IP addresses.

With this framework as a background, let's turn to the licenses.

License Analysis

Methodology

Data Set and Sampling

To evaluate licensing contracts with respect to privacy provisions and intellectual freedom, we conducted a content analysis of a set of licenses collected by Bergstrom, Courant, and McAfee for a study of journal pricing.[49] Using state open records laws, Bergstrom et al. sent records requests to a large number of state university libraries throughout the United States. They received licenses from 38 universities and 8 consortia from 28 states, involving 11 different publishers. Bergstrom shared 216 licenses on CD-ROM with Eschenfelder for a study of downloading, scholarly sharing, interlibrary loan, and e-reserves.[50] We chose to analyze a sample of the full Bergstrom set because it is common in social science research to employ a random sample to represent a population that one does not have the resources to query. Moreover, Eschenfelder et al. found substantial repetition within the 216 licenses, which suggests that a subsample will capture most variances. Thus, we analyzed a subset of the 216 licenses. To better reflect current practices, we sampled licenses from 2007 to 2009.[51]

Following Eschenfelder et al., we used final licenses rather than standard licenses in this study. Standard publisher licenses were drafted by publishers and presented to potential licensees as a basis for final licenses, and final licenses were those licenses negotiated between licensees and licensors.[52] We selected 42 licenses from the data set using stratified sampling based on publisher to ensure capture of the variations in licenses across different publishers. We applied simple random sampling to draw a sample of 4 licenses from each of 10 publishers. We had only two licenses from the eleventh publisher (Wiley-Blackwell) and thus included both licenses in our data set. The demographic information about the sample is illustrated in table 1.

TABLE 1
Licenses by Publishers and Year

Commercial Publisher N=30				
	Total Licenses	**2007**	**2008**	**2009**
Wiley (WLY)	4	—	2	2
Blackwell (BLW)	4	3	—	1
Wiley-Blackwell (WBL)	2	—	1	1
Elsevier (ELV)	4	—	2	2
Emerald (EMR)	4	1	2	1
Sage (SGE)	4	3	—	1
Springer (SPR)	4	—	3	1
Taylor & Francis (T&F)	4	2	1	1
Non-commercial Publisher N=12				
American Chemical Society (ACS)	4	—	3	1
Oxford University Press (OUP)	4	2	1	1
Cambridge University Press (CUP)	4	3	—	1
Total	**42**	**14**	**15**	**13**

Codebook Development

To conduct a reliable content analysis of the sample licenses, we developed a codebook based on our literature review and a review of privacy clauses found in 3 standard e-journal licenses available on the Internet. We revised the codebook based on 10 rounds of testing licenses selected from Bergstrom's data set that were not part of the official study sample. During each test, we independently coded licenses and calculated the inter-coder reliability (ICR) of our results. In the last round of coding, we obtained 98 percent ICR overall, higher than the targeted 90 percent. The minimum ICR at the question level was 75 percent.

The final codebook included 38 variables in five categories: monitoring authorized use/users, data collected by publishers, data shared with third parties, data sent to licensees, and personalized services. For most of the variables (27), we marked "1" if the statement was true, and marked "0" if it was not true. For the rest of the variables, we filled in free text to capture more complex information about the licenses.

Data Collection and Analysis

Each coder coded 25 licenses in total, among which 8 licenses were in common to calculate the ICR for the final coding. All the licenses were printed out and coded on paper. License coding ended in February 2013, and ICR for the 8 common licenses was calculated as 96 percent. The minimum ICR at the question level was 87.5 percent. All data collected in the coding were then imported to Excel for further analysis.

In addition, we compared the data to two model licenses developed by library professionals: the LIBLICENSE model license and the license agreement checklist from the California Digital Library (CDL).[53] We discuss differences between the data set and the model licenses in section 4.1.

Findings

Lipinski and Harris distinguish two primary areas in which privacy may be of concern in licensing agreements: (1) enforcement of authorized use and (2) collection of personal information by licensors and sharing that information with third parties.[54] In addition, Magi addresses licensors offering personalized services that may in turn collect information about individual users.[55] Our findings reflect this split, with privacy-affecting provisions clustering around provisions for enforcing authorized use and for licensors' collecting and sharing user data. We found only 2 licenses addressing personalized services.

Monitoring Authorized Use/Users

Most licenses authenticate users by IP address; therefore, most licenses specify that licensors will obtain IP address information from licensees. This authentication method in some cases provides a mechanism for enforcing license terms, as 16.7 percent of licenses state that licensors may suspend access of the IP address(es) from which unauthorized use occurs. Another 9.5 percent of licenses stipulate that the licensor may suspend the access of any authorized user violating the terms of use, without specifying the mechanism. In addition, several contracts have provisions either allowing (2.4%, 1 license) or even *requiring* (9.5%) that libraries suspend authorized user access upon request from the licensor. Ten licenses (23.8%), all from commercial publishers, that allow the licensor to suspend access also require (in at least some cases) that the licensor provide the licensee notice and time before suspending access, potentially allowing the library to remedy unauthorized use.

A substantial number of licenses (38.1%) require that the licensee *monitor* for unauthorized use of licensed materials.[56] Further, 42.9 percent of licenses require that

libraries *take disciplinary action* when they become aware of unauthorized use. Just what constitutes "disciplinary action" is not spelled out in the licenses. Finally, the large majority of licenses (81%) oblige libraries to notify publishers when they become aware of unauthorized use. See table 2 for summary.

TABLE 2
Terms Enforcing Authorized Use

	Total	
	N=42	Total %
Licensor can suspend authorized user access based on violation of license*†	4	9.5%
Licensor can suspend access based on IP address*†	7	16.7%
Licensee *may* suspend unauthorized users†	1	2.4%
Licensee *shall* suspend unauthorized user based on request from licensor*	4	9.5%
Some suspensions require licensor to provide notice and time†	10	23.8%
Licensee shall notify publisher of unauthorized use	34	81.0%
Licensee shall monitor for unauthorized use	16	38.1%
Licensee shall take disciplinary action when aware of unauthorized use	18	42.9%
*Provision is included in the CDL model license. †Provision is included in the LIBLICENSE model license.		

Other Enforcement Provisions

Along with the results of our content analysis, close reading of the contracts in the data set reveals a number of privacy-affecting terms that are relevant here.

Several licenses have provisions regarding monitoring of use that warrant additional attention. Two licenses from Wiley/Blackwell require that "licensee shall *use all reasonable endeavors* to monitor compliance" and report any unauthorized use or breach. That would appear to increase the licensee's responsibility beyond the requirement of merely doing some sort of monitoring. Other licenses, including two licenses from Springer and one from ACS, reserve the licensor's right to monitor access "to detect misuse of [publisher's] content." The ACS license states explicitly that it will engage in routine monitoring of each IP address authorized to access its materials: "[Publisher] will monitor the volume of searching and answer downloading activity associated with each Authorized IP Address on a routine bases, for the purpose of: (1) benchmarking what 'average' use is among Authorized IP Addresses, and (2) noting any significant variance in patterns of usage for particular Authorized IP Address(es)." This implies a degree of information gathering and specificity by the publishers greater than a provision simply stating that some monitoring will occur. Interestingly, an addendum to one license from Oxford University Press makes explicit that the library will *not* monitor use, stating that "The Licensee does not have the ability to monitor or to control actual uses by authorized users of the information from the Licensed Materials or to notify authorized users of all the restrictions on the use of information in this Agreement." We found similar language in an addendum to an Emerald license from the same institution.

Thirteen licenses require that licensees maintain and share records of authorized users *and their access details*. The licenses do not specify precisely what details about user access are required; one possibility is that it includes user logs that can be correlated with the IP addresses requesting resources. Perhaps even more important for our purposes here, five licenses, including all of the four Emerald licenses and one from Wiley in 2009, require libraries to share information about users and their activities with licensors. As the Emerald licenses stipulate: "full and up-to-date records of all Authorized Users and their access details…shall be provided to [the publisher] upon request."[57]

In addition to provisions regarding information gathering, monitoring, and sharing to enforce license terms, six licenses (one Blackwell 2007 license, one ACS 2009 license, and all four Wiley licenses) require that libraries "cooperate" or "cooperate fully" with publishers' investigations of copyright infringement and unauthorized use.

Data Collection and Sharing

A second area that can implicate privacy interests is licensors' data collection, analysis, and sharing information with third parties. As Lipinski notes, information is generated from database searches and downloads of licensed content and that information may be connected to particular patrons or particular computers.[58] While similar information collected by libraries (for example, searches of an OPAC) would likely be protected from disclosure under many state library privacy statutes, such information is generally *not* protected by statute when collected by licensors (who are generally not covered in state library statutes).[59] Hence, licensors may have the ability to collect, analyze, and even share information gleaned from searching and using licensed content.[60]

As shown in table 3, we found that 66.7 percent of licenses state that publishers collect non-IP data, including (for example) *usage* data. We found that 26.7 percent of commercial licenses expressly allow publishers to share data with third parties. All but one of the licenses from noncommercial publishers are silent regarding data sharing with third parties—neither specifically allowing nor prohibiting data sharing. Perhaps most important is that a number of licenses explicitly limit the form or type of data that publishers may share with third parties. Table 4 shows that eight licenses (19.0%) specify that data shared with third parties is usage data. Five of these licenses expressly state that usage data shared with third parties should be in anonymous and aggregated form. Moreover, 31.0 percent of licenses specify that publishers may not share with third parties either raw usage data or data that can identify individual users.

A total of 9.5 percent of the licenses allowing publishers to share data with third parties specify the *types* of third parties with whom publishers can share data. For example, an Elsevier license states that it may provide data to "vendors or other third-parties retained by the subscriber," and a Wiley license states that it may provide data to third parties "where necessary in connection with services provided by appropriate

TABLE 3
Terms Specifying Data Collecting

	Total	
	N=42	**Total %**
Non-IP data collected by publisher*†	28	66.7%
Specifying reasons for data collection	9	21.4%
Including deletion of collected data	1	2.4%

*Provision is included in the CDL model license.
†Provision is included in the LIBLICENSE model license.

TABLE 4
Terms Specifying Data Sharing with Third Parties

	Total	
	N=42	Total %
Contains terms about publisher sharing data with third party* †	14	33.3%
Specifies that publisher may provide data to third party*†	9	21.4%
Specifies the type of data that publisher may provide to third party*†	8	19.0%
Specifies the type of third party to which publisher will provide data*†	4	9.5%
Specifies that some data will NOT be provided to third party*†	13	31.0%
Specifies that publisher will disclose data to third party if required by law	1	2.4%
Requires third party to comply with the confidentiality provisions of license	1	2.4%

*Provision is included in the CDL model license.
†Provision is included in the LIBLICENSE model license.

intermediaries." Several licenses (21.4%) specify the reasons that non-IP address information is collected. These reasons include 1) assisting both licensor and participating library to understand the impact of this license; 2) improving the services provided by the publisher; and 3) internal use for the publisher and the licensee. It is unclear what the lifecycle of such data is, as just one license mentions the possibility of deleting data collected, stating that the deletion will occur "when [such data are] no longer needed."

Data Sent to Licensee

Data about usage of electronic journals are extremely useful for libraries to calculate cost-effectiveness of specific products and to allocate resources appropriately. Many licenses (including both model licenses) oblige licensors to provide licensees with usage data. One important question is whether the data contain personally identifiable information. Contracts can ensure that usage data do not contain personally identifiable information by expressly stating that usage data will be provided in aggregated form; likewise, contracts can state that usage data will comply with the Counting Online Usage of Networked Electronic Resources (COUNTER) Codes of Practice, which includes some privacy protection.[61]

As illustrated in table 5, we found that 38.1 percent of licenses require that usage data be in an aggregated form (either explicitly or by requiring COUNTER compliance). We also found that 64.3 percent of licenses provide usage data to licensees, and 21.4 percent of the licenses include COUNTER-compliance provisions. Most of the publishers in this study are registered with COUNTER as providing COUNTER

TABLE 5
Terms Specifying Data Sent to Licensee

	Total	
	N=42	Total %
Providing usage data to licensee*†	27	64.3%
COUNTER-compliance†	9	21.4%
Aggregated usage data*†	16	38.1%

*Provision is included in the CDL model license.
†Provision is included in the LIBLICENSE model license.

compliant usage reports. Thus, their practices may comport with COUNTER practices regardless of whether such practices are made explicit in licenses. Also note that, even if licenses require the licensor to provide licensees with aggregated usage data, we cannot determine from the licenses whether or not licensors actually are able to collect usage data about individual users.

Discussion: Is There Reason for Concern?

The next question motivating the paper is whether any of the privacy-affecting provisions in licensing contracts are problematic. To address this, we compare our findings: first, to two model licenses established by library professionals; second, to recommendations outlined by Lipinski (2013); and last, to the facets of intellectual freedom described above.

Model Licenses

As noted, we recorded which provisions from our analysis of the data set were in the LIBLICENSE and CDL model licenses. Because the LIBLICENSE and CDL model licenses were developed by library professionals, they reflect considered professional judgments that presumably take into account both professional library values and the need to secure access to electronic journals through licenses. Hence, comparing them to the data set should reflect the extent to which professional values are reflected in actual license contracts.

The major difference between the model licenses and the actual licenses concerns libraries' obligations when unauthorized use occurs. The actual licenses impose greater obligations on libraries. In particular, neither of the model licenses requires libraries to monitor for or notify publishers about unauthorized use. Nonetheless, we found 38.1 percent of licenses have monitoring provisions, and 81 percent have notification provisions. Moreover, nearly half of the licenses in the data set (42.9%) oblige libraries to take disciplinary action when they are aware of unauthorized use, though neither model license contains such an obligation.

These discrepancies between actual licenses and model licenses reflect different understandings regarding libraries' role in addressing unauthorized uses. Publishers, on one hand, have an interest in having libraries act in the publishers' interest in preventing unauthorized use as much as possible. By contrast, at least some librarians are hesitant about incurring obligations to prevent unauthorized uses. Duranceau et al. argue that libraries should focus on user education rather than "policing of license breaches" and suggest libraries exclude license terms requiring "taking a specific disciplinary action" to avoid creating obligations to police for unauthorized use.[62] These concerns are reflected in the model licenses developed by library communities, which do not contain any provisions requiring monitoring, notification, or disciplinary actions.

Another important difference between model licenses and actual licenses is the requirements on data sharing. Both model licenses have clearly specified provisions to limit publishers' ability to share usage data with third parties. Both specify that publishers may only share aggregated usage data with third parties and that raw usage data that can be used to identify individual users cannot be shared with third parties. Only a small percentage of the licenses in our data set specify the data type that can be shared with third parties, and most of the licenses are silent with respect to data sharing (hence not prohibiting it).

Conflicts with Lipinski (2013)

Lipinski's discussion of privacy implications of common licensing provisions provides a further point of comparison for the data set. He maintains that requirements

for the "reporting of infringers raises issues of privacy" and suggests that a "kinder, gentler" approach is a more general affirmation that licensee assistance in halting abuses would "be appreciated."[63] He also suggests that, instead of an obligation to notify the licensor of all infractions, libraries should be able to address the infraction without having to contact the licensor, and hence have the option for a "teachable moment."[64] Most licenses (81%) in our data set, however, treat licensee's notification of unauthorized uses as an obligation, without an option to address infractions themselves. Lipinski further recommends, in cases where licensors have the right to terminate authorized user access, that they provide the licensee with notice and opportunity to remedy violations before terminating use. Only a few licenses (23.8%) in our set have such provisions.

Lipinski points out that clauses requiring that libraries assist in the investigation and enforcement of licensing provisions may run afoul of state library privacy protection laws, for providing information would generally involve divulging patron information. He suggests inserting a clause in contracts that indicates that libraries will not divulge protected information in enforcing contract terms. As noted earlier, we found several licenses requiring libraries "cooperate" or "cooperate fully" with publishers' investigations of copyright infringement and unauthorized use. We did not capture clauses in any contracts indicating that libraries would not divulge protected information, and we are not aware of any such provisions from our close reading of the licenses. In addition, Lipinski recommends that licenses require licensees to give users reasonable notice of terms and enforce license terms at levels consistent with other institutional policies.[65] Likewise, we did not capture such terms in our coding and are not aware of any such provisions from our close reading.

Four Facets of Intellectual Freedom

So there are ways in which the licenses in our data set are in at least some tension with statements of professional library values. But so what? To analyze whether the potential privacy issues in licensing contracts are indeed a problem, we have to analyze them in light of the underlying justifications for privacy protections in the library context—intellectual freedom. The picture here is more complicated. As we have established, we can distinguish four facets of intellectual freedom. Licensing provisions implicate only some of these, to varying degrees.

Negative Freedom

Consider first negative freedom, or freedom from external constraints. Recall that privacy may bear upon negative intellectual freedom insofar as others may use information gleaned about a person's intellectual pursuits to that person's detriment. It does not appear that any of the contract terms outlined impinge upon this facet of intellectual freedom. Nothing in the monitoring provisions suggests that publishers or third parties will harm individual users of electronic resources based on the content of their research, and it is difficult to see why there would be any incentive to do so. One might argue that monitoring and reporting will indeed impose costs on some persons insofar as those who violate authorized use terms may lose access privileges. That, however, would not be a restriction on the content of one's intellectual inquiries, but on the form they take. Put another way, whatever external constraints are applied to persons based on their research activities would be content neutral in the same way that late fees might place external constraints on persons' research activities, but not in a way that is plausibly an impingement of intellectual freedom. Intellectual freedom does not include the ability to conduct inquiry in *any manner* that one wishes. Constraints on time, place, and manner (at least where those are not onerous and are justifiable

as means of ensuring overall broader and easier access) are compatible with negative freedom. Indeed, the ALA considers them compatible with intellectual freedom.[66]

There is some concern, though, insofar as provisions for monitoring could be used as sources for investigation. For example, requirements that libraries keep usage logs and provide information to publishers seeking to enforce authorized use terms will make such retained information available for, for instance, subpoenas. This type of concern is nothing new, though, as any library record is subject to such information requests.

Positive Freedom

Consider next positive freedom, or the "true freedom of inquiry" that is imperiled, where "users *recognize or fear* that their privacy or confidentiality is compromised."[67] Some types of license provisions potentially conflict with this facet of intellectual freedom. One concerns policing for unauthorized use. Libraries are in many cases required to monitor for, report to publishers, and discipline unauthorized use, and publishers may maintain records of IP addresses of persons accessing licensed materials, which are in some cases able to be linked to particular computers. A few licenses require that libraries maintain and share records of authorized users and their access details, and at least one stipulates that "full and up-to-date records of all Authorized Users and their access details…shall be provided to [the publisher] upon request." Such use records can potentially be correlated with unauthorized activity. Recall the license explicitly stating that it will engage in routine monitoring of each IP address authorized to access its materials: "[Publisher] will monitor the volume of searching and downloading activity associated with *each Authorized IP Address on a routine bases*, for the purpose of: (1) benchmarking what 'average' use is among Authorized IP Addresses, and (2) noting any significant variance in patterns of usage for particular Authorized IP Address(es)."

These license terms could constrain internal, positive freedom insofar as they could cause patrons to believe that they lack privacy, which may in turn lead them to limit their inquiries. That is, the degree of required and potential monitoring, *if known*, could serve to chill use. But the potential limits on internal positive freedoms are merely speculative. The extent to which privacy loss affects persons' behavior is difficult to measure, and it is plausible that people become so accustomed to privacy loss that remote monitoring does not actually affect their behavior. Moreover, it is likely that most users of electronic journals in academic libraries do not know whether, or the degree to which, their use is actually monitored. After all, license provisions are not widely publicized (the licensing contracts in our data set, for example, were disclosed pursuant to open records requests), and the licenses only reveal that some monitoring takes place. The intensity and precision of that monitoring remains unclear and is likely to vary by publisher and by library. Thus, patrons may not know enough that monitoring actually affects their behavior.

That is a fairly powerful argument that monitoring provisions in licensing contracts do not undermine internal, positive intellectual freedom. If it is true that no one actually alters their inquiries based on monitoring because they are unaware of monitoring, then there would not appear to be any limitation on persons' internal, positive intellectual freedom. For actions to be chilled requires that persons be aware of monitoring and for that awareness to have some effect upon their behaviors. If monitoring is well hidden or if persons are simply unaware of it, then there will be correspondingly less concern about chilling effects and internal positive freedom. Indeed, if chilling effects were the sole concern of intellectual freedom, there would be a strong case against *disclosing* monitoring activities.

Freedom and Autonomy

Because it is unlikely that patrons have a clear idea of how much of their activity is actually subject to monitoring, the license provisions implicate the third facet of intellectual freedom, or quality of agency. Recall that this facet is at root about individual autonomy, or the ability to act according to one's own reasons as one sees fit. Autonomy in turn requires information important in determining how to steer one's choices to comport with one's values, and it demands information important in making sense of how one is being treated. To the extent that licensing provisions require or allow monitoring and data collection, and to the extent that such actions matter to patrons, this facet of intellectual freedom demands disclosure of those provisions. Whether users are actually aware of the degree to which their uses may be monitored and information collected is unclear, but license terms are obscured in a number of ways. They are in the relatively difficult language of contracts; they may require open records requests to see (and if they are not licenses involving public institutions, such requests are not an option); they differ across publisher and license; interpreting them demands knowledge of the underlying technologies and institutional practices of libraries and publishers.

The upshot is that surveillance will implicate a different facet of intellectual freedom depending on whether it is overt (which will implicate internal, positive freedom) or covert (which will implicate freedom-as-agency). Thus, the fact (if it is a fact) that persons using electronic journals are unaware of the degree to which their uses are monitored is not sufficient to conclude that their intellectual freedom is unimpinged.

Republican Freedom

A different potential concern bridges privacy and authorized use. In the separate section of the paper, we describe a number of ways in which uses may conflict with the terms of licensing agreements and a number of uses that appear both routine and unauthorized. This is significant in two ways. One is that licensing terms provide bounds of criminal conduct.[68] Under the U.S. Computer Fraud and Abuse Act (CFAA), any person who "intentionally accesses a computer without authorization or *exceeds authorized access*, and thereby obtains…information from any protected computer" may be subject to criminal and civil sanction.[69] This is one of the statutes under which Aaron Swartz was charged. While Swartz's actions involved large-scale downloading and attempts to mask his identity, at root those actions are relevant because they are instances of exceeding authorized use. That is, Swartz's actions fall within the purview of the CFAA based upon licensing terms. Moreover, license terms both set the bar for unauthorized access well below what happened in the Swartz case, are vague about what precisely exceeds authorized access, and in some cases seem to prohibit mundane types of access. Hence, there is a legally plausible case for criminal conduct and civil liability for activities that are either within, or not far outside, norms of scholarly activity.

A second reason this is significant is that licenses generally provide that users may lose authorization to access materials based on violations of license terms. Further, publishers may monitor the activities of individual users and may retain that information over time. Licenses almost all stipulate that they retain the right to enforce those provisions, even if at times they fail to enforce license provisions. This combination exposes users to restrictions on use (and both criminal and civil liability) in the future based on data collected in the past. The potential for enforcement, either by restriction of use in the undefined future or for criminal charges, implicates the fourth facet of freedom—freedom from arbitrary power.

As discussed, Lovett outlines three conditions under which republican freedom is limited, such that one may be subject to arbitrary power: an imbalance of power, a "dependency" relation, such that costs would be incurred by the party with less

power by exiting the relationship, and an absence of known rules or conventions governing use of power.[70] In the case of users and publishers and prosecutors, there does seem to be a difference in power; individuals have less power than publishers and prosecutors. And users of university libraries would incur substantial costs by not using electronic resources for their research. These conditions by themselves are unproblematic, though. It is only with the further presence of the third condition that republican freedom is limited.

Is there, then, an absence of known rules or conventions governing use of power? The Swartz case does suggest an absence of known rules or conventions. Prior to the Swartz case, there do not appear to have been indictments under the CFAA for unauthorized access, or exceeding authorized access, of materials licensed by libraries, and the indictment under the CFAA was surprising to many people. Further, the principal parties did not express support for the prosecution in the first place. JSTOR indicated that, once its documents had been secured, it "had no interest in [the case] becoming an ongoing legal matter," and MIT expressed no statement regarding the merits of prosecution at all.[71] Hence, there does not appear to be a rule or convention that excessive downloading would be subject to criminal charges. Now, a number of commentators have argued that the charges were *legally* cognizable under the language of the CFAA, and one might argue that the law itself provides the kind of rule governing the use of power that is necessary for republican freedom. But, even if we assume that is the case, whether the charges were supported by the law is a separate question from whether the exercise of prosecutor discretion to bring the charges was itself within rules or conventions surrounding unauthorized use, or exceeding authorized use, of licensed materials. After all, both the law itself and its application may be so broad as to be arbitrary. Moreover, to say that rules or conventions exist surrounding discretionary prosecution just in virtue of the fact that the law supports prosecution would entail (implausibly) that laws or the use of laws is *by definition* nonarbitrary. Moreover, it would assume the answer to the very question we are asking in this section, that is to say, *whether* it is an arbitrary exercise of power to use the CFAA to bring charges for unauthorized use or exceeding authorized use of licensed resources.

The facts of the Swartz case, however, are extreme—large-scale, automated downloading that Swartz likely knew to conflict with license terms. That looks nothing like the kind of mundane violations we have described, and one might argue that such low-level offenses would never be prosecuted in the way Swartz was. On the republican conception of freedom, however, it is not *actual* interference that limits liberty, but the *ability* of others to interfere arbitrarily (even if they don't). That is the force of the mob example: the mere fact that one happens to steer clear of mob interference does not entail that one is free of the mob, for without effective legal authority controlling the mob, it has the ability to exercise its power over others arbitrarily. Regardless of whether it is justifiable to prosecute actions such as Swartz's, the language of the CFAA is broad enough that comparatively minor cases could also be subject to criminal sanctions or civil liability. Up to this point, they haven't been prosecuted, but the *ability* to do so is what limits republican freedom. Moreover, even if there were explicit rules adopted by the U.S. Department of Justice not to prosecute these sorts of unauthorized uses, the CFAA still imposes civil liability. Hence, there would remain the possibility of licensors suing library patrons for unauthorized access or exceeding authorized access, which also would appear to meet the three conditions for being subject to arbitrary power. It is certainly true that the extent to which arbitrary power can be exercised is less in the civil liability case than in the criminal case. That, however, only tells us that this fourth, republican, facet of freedom admits of degree.[72]

Now, one might argue that this is not about privacy, but instead about authorized access provisions and the scope of the U.S. criminal code. It is correct that the primary concern is that license terms underwrite criminal and civil liability. However, because of the requirements that libraries monitor and notify publishers about unauthorized use, and in some cases are required to maintain relevant information and cooperate with investigations, the provisions affecting privacy are mechanisms by which this fourth facet of intellectual freedom is affected.

Summary and Recommendations

To sum up, our data indicate that licenses conflict with at least some language advanced by the ALA, with some aspects of two model contracts developed by library professionals, and with recommendations set out by Lipinski. Whether they conflict with the underlying value justifying privacy protections—intellectual freedom—depends on the facet of intellectual freedom. It is difficult to see how they conflict with negative intellectual freedom or internal, positive freedom. However, the likelihood that patrons are unaware of privacy-affecting provisions, and the difficulty of figuring out just how privacy is affected by use of electronic resources, creates an important conflict with the quality-of-agency facet of intellectual freedom. Moreover, the relationship between unauthorized access, or exceeding authorized access, and criminal and civil liability, the possibility of losing access to licensed material, publishers' ability to gather information about unauthorized access, and libraries' obligations (in some cases) to monitor, notify about, and collect information about such use, create an important conflict with the fourth facet of intellectual freedom, or freedom from arbitrary power.

There remains, however, an important question about whether such tensions between licenses and the values of the library profession are sufficient to justify the risk of being unable to reach agreement on license terms and thus risking patron access to licensed resources. Likewise, it is unclear whether it would be worth paying more for licenses that provide greater privacy protections. After all, failing to provide such resources may *also* limit intellectual freedom.[73] Nonetheless, there are several things that are worth considering.

First, libraries should ensure that contract terms pertaining to data collection, data sharing with third parties, and monitoring and disciplinary actions for unauthorized use are transparent to users. Although most users are unlikely to peruse the terms of licenses, making the terms readily available is a necessary (though not sufficient) condition for respecting the autonomy of patrons and thus for supporting the quality-of-agency facet of intellectual freedom. Related, libraries should seek to make privacy-affecting provisions similar across licenses. The fact that contracts are so different from publisher to publisher makes it all the more difficult for users to know what provisions are in effect. Next, libraries should resist license terms obliging them to monitor for unauthorized use, and they should make clear in licenses where they do not have the ability to identify individual users.

In fact, libraries would go a long way to securing intellectual freedom (in all its facets) by implementing several of Lipinski's suggestions. For example, rather than notifications requirements, libraries could push for provisions allowing them to correct unauthorized access issues "in-house." Related, libraries could push for provisions such that licensors can terminate authorized access only were they provide licensees with notice and the opportunity to remedy violations. Libraries might also specify in contracts that they will not provide individual user information in any case, regardless of any language suggesting that they will "cooperate" or "cooperate fully" with publishers' investigations of copyright infringement and unauthorized use. And they could specify that any information shared with third parties be in anonymous and aggregate form.

Last but not least, libraries can reduce the privacy risks associated with investigating unauthorized uses in multiple ways. The first step is to adopt privacy-preserving authentication technologies such as Shibboleth (as many libraries already do) to reduce the user information directly obtained by publishers. Then both libraries and campus IT departments should regularly purge the records kept in their servers to minimize the amount of user information that could be requested by publishers to a great extent.

We would like to thank Kristin Eschenfelder, Dorothea Salo, Sue Dentinger, Bryce Newell, Faye Jones, and Anne Klinefelter, and the anonymous reviewers for College & Research Libraries *for comments on earlier versions of this paper*.

Notes

1. Indictment, *U.S. v. Swartz* No. 1:11-cr-10260-NMG (D. Mass. July 14, 2011); Superseding Indictment, *U.S. v. Swartz* (Sept. 12, 2012).
2. Massachusetts Institute of Technology (MIT), *Report to the President: MIT and the Prosecution of Aaron Swartz*, available online at http://swartz-report.mit.edu/docs/report-to-the-president.pdf [accessed 1 November 2013].
3. 18 U.S.C. § 1030(a). "Protected computer" is interpreted to mean any computer used in interstate commerce, or "effectively all computers with Internet access." *U.S. v. Nosal*, 676 F.3d 854, 859 (9th Cir. 2012).
4. See Nancy Sims, "Library Licensing and Criminal Law: The Aaron Swartz Case," *College & Research Libraries News* 72, no. 9 (Oct. 1, 2011): 534–37.
5. Tomas A. Lipinski, *The Librarian's Legal Companion for Licensing Information Resources and Services* (Chicago: Neal-Schuman, 2013).
6. American Library Association, *Intellectual Freedom Manual*, 8th ed. (Chicago: American Library Association, 2010), 312.
7. Ibid., 304.
8. Ibid., 177.
9. Ibid., 150. The emphasis on privacy is reflected in professional attitudes. A survey in 2012 found that 95 percent of librarians in the study agreed that users should be able to control who could look at their personal information; and more than 95 percent of surveyed librarians believed that businesses and government agencies shouldn't share personal information with third parties unless they obtain users' permission and only for specific purposes. *ALA News*, "CIPR Survey Confirms Librarians' Commitment to Protecting Privacy Rights," May 1, 2012, available online at http://cipr.uwm.edu/?p=81 [accessed 7 April 2015].
10. American Library Association, *Intellectual Freedom Manual*, 41. (citing American Library Association, "Privacy: An Interpretation of the Library Bill of Rights," 2002, available online at www.ala.org/Template.cfm?Section=interpretations&Template=/ContentManagement/ContentDisplay.cfm&ContentID=132904 [accessed 7 April 2015].)
11. American Library Association, *Intellectual Freedom Manual*, 179.
12. Ibid., 178.
13. Ibid.
14. Canadian Library Association, *Canadian Library Association Code of Ethics* (June 1976), available online at www.cla.ca/Content/NavigationMenu/Resources/PositionStatements/Code_of_Ethics.htm [accessed 7 April 2015].
15. Loida Garcia-Febo et al., *IFLA Code of Ethics for Librarians and Other Information Workers* (The Hague International Federation of Library Associations and Institutions, August 2012), 3, available online at www.ifla.org/files/assets/faife/publications/IFLA%20Code%20of%20Ethics%20-%20Long_0.pdf [accessed 7 April 2015].
16. Pnina Shachaf, "A Global Perspective on Library Association Codes of Ethics," *Library & Information Science Research* 27, no. 4 (2005): 513–33.
17. American Library Association, Intellectual Freedom Manual, 305–319; Stacey L. Bowers, "Privacy and Library Records," *Journal of Academic Librarianship* 32, no. 4 (July 2006): 377–83, doi:10.1016/j.acalib.2006.03.005.
18. American Library Association, *Intellectual Freedom Manual*, 178.
19. International Federation of Library Associations and Institutions (IFLA), *IFLA Statement on Libraries and Intellectual Freedom* (2013), available online at www.ifla.org/publications/ifla-statement-on-libraries-and-intellectual-freedom [accessed 7 April 2015].
20. Croatian Library Association, *Statement on Free Access to Information* (Sept. 2000), available

online at http://www.ifla.org/files/assets/faife/statements/hkdstat.pdf [accessed 7 April 2015]; Japan Library Association, *A Statement on Intellectual Freedom in Libraries* (May 30, 1979), available online at http://www.ifla.org/files/assets/faife/statements/jlastat.pdf [accessed 7 April 2015]; Library & Information Association New Zealand Aotearoa (LIANZA), *Access to Information* (May 11, 1978), available online at http://www.ifla.org/files/assets/faife/statements/nzacc.pdf [accessed 7 April 2015]; Turkish Librarians' Association, *Intellectual Freedom Statement* (Feb. 22, 2008), available online at http://www.ifla.org/files/assets/faife/statements/tlastat.pdf [accessed 7 April 2015]; Chartered Institute of Library and Information Professionals (CILIP), *User Privacy in Libraries: Guidelines for the Reflective Practitioner* (2011), available online at www.cilip.org.uk/sites/default/files/documents/Privacy_June_AW.pdf [accessed 7 April 2015].

21. See, for example, Charles Fried, *An Anatomy of Values; Problems of Personal and Social Choice* (Cambridge: Harvard University Press, 1970); Julie C. Inness, *Privacy, Intimacy, and Isolation* (New York: Oxford University Press, 1992); and James Rachels, "Why Privacy Is Important," *Philosophy and Public Affairs* 4, no. 4 (Summer 1975): 323–33 for views on privacy's value in terms of personal and professional relationships; Ruth Gavison, "Privacy and the Limits of Law," *Yale Law Journal* 89, no. 3 (Jan. 1, 1980): 421–71; Anita L. Allen, *Uneasy Access: Privacy for Women in a Free Society* (Totowa, N.J.: Rowman & Littlefield, 1988); Judith Wagner DeCew, *In Pursuit of Privacy: Law, Ethics, and the Rise of Technology* (Ithaca, N.Y.: Cornell University Press, 1997) for views relating privacy to important decisions and democratic processes; and Edward Bloustein, "Privacy as an Aspect of Human Dignity: An Answer to Dean Prosser," *New York University Law Review* 39 (1964): 962–1007 for a view explicitly linking privacy to human dignity.

22. Richard A Posner, "The Right of Privacy," *Georgia Law Review* 12 (1977): 393.

23. Catharine A. MacKinnon, *Toward a Feminist Theory of the State* (Cambridge, Mass.: Harvard University Press, 1989); Anita L. Allen, *Uneasy Access: Privacy for Women in a Free Society* (Totowa, N.J.: Rowman & Littlefield, 1988).

24. Jeremy Waldron, "Security and Liberty: The Image of Balance*," *Journal of Political Philosophy* 11, no. 2 (2003): 191–210, doi:10.1111/1467-9760.00174; Richard A. Posner, "Privacy, Surveillance, and Law," *University of Chicago Law Review* 75 (2008): 245.

25. Eliza T. Dresang, "Intellectual Freedom and Libraries: Complexity and Change in the Twenty-First-Century Digital Environment," *The Library Quarterly* 76, no. 2 (Apr. 1, 2006): 169–92; Alan Rubel, "Libraries, Electronic Resources, and Privacy: The Case for Positive Intellectual Freedom," *The Library Quarterly* 84, no. 2: 183–208, at 188–89

26. This theoretical framework has been developed more fully developed in Rubel, "Libraries, Electronic Resources, and Privacy" and Alan Rubel "Privacy and Positive Intellectual Freedom," *Journal of Social Philosophy*) 45(3) (Fall 2014): 390–407.

27. Isaiah Berlin, *Four Essays on Liberty* (London: Oxford University Press, 1969); Ian Carter, "Positive and Negative Liberty," in *The Stanford Encyclopedia of Philosophy*, ed. Edward N. Zalta, (Fall 2008), available online at http://plato.stanford.edu/archives/fall2008/entries/liberty-positive-negative/ [accessed 7 April 2015].

28. American Library Association, *Intellectual Freedom Manual*, 178.

29. Julie E. Cohen, "Information Rights and Intellectual Freedom," in *Ethics and the Internet*, ed. Anton Vedder (Antwerp: Intersentia, 2001): 11–32; Neil M. Richards, "Intellectual Privacy," *Texas Law Review* 87 (2008): 387; Marc Jonathan Blitz, "Constitutional Safeguards for Silent Experiments in Living: Libraries, the Right to Read, and a First Amendment Theory for an Unaccompanied Right to Receive Information," *UMKC Law Review* 74 (2006): 799.

30. Carter, "Positive and Negative Liberty"; Berlin, *Four Essays on Liberty*; John Christman, "Saving Positive Freedom," *Political Theory* 33, no. 1 (Feb. 1, 2005): 79–88.

31. American Library Association, *Intellectual Freedom Manual*, 178.

32. Ibid.

33. Linda Greenhouse, "F.B.I. Defends Library Monitoring Program," *New York Times* (July 14, 1988), available online at www.nytimes.com/1988/07/14/us/fbi-defends-library-monitoring-program.html [accessed 7 April 2015]; Herbert N. Foerstel, *Surveillance in the Stacks: The FBI's Library Awareness Program* (New York: Greenwood Press, 1991).

34. Public Law 107–56 (2001).

35. Of course, secret surveillance often gets revealed, which could in turn result in chilling effects and other negative consequences. Nonetheless, if it is *only* such effects that matter for intellectual freedom, one could maintain that it is the revelation, rather than the surveillance, that is the primary threat. Surely, though, the surveillance itself is the limitation on freedom rather than awareness of it.

36. John Christman, "Liberalism and Individual Positive Freedom," *Ethics* 101, no. 2 (Jan. 1, 1991): 343–59; Christman, "Saving Positive Freedom."

37. Thomas E. Hill, Jr., "Autonomy and Benevolent Lies," *Journal of Value Inquiry* 18 (1984): 251–97. Alan Rubel, "Privacy and the USA Patriot Act: Rights, the Value of Rights, and Autonomy,"

Law & Philosophy 26 no. 2 (2007) 119–59.

38. See Robert Streiffer and Alan Rubel, "Democratic Principles and Mandatory Labeling of Genetically Engineered Foods," *Public Affairs Quarterly* 18, no. 3 (2004): 223–48

39. Philip Pettit, "Keeping Republican Freedom Simple: On a Difference with Quentin Skinner," *Political Theory* 30, no. 3 (June 1, 2002): 339–56; Philip Pettit, "Freedom as Antipower," *Ethics* 106, no. 3 (Apr. 1, 1996): 576–604.

40. American Library Association, *Intellectual Freedom Manual*, 42.

41. Francis N. Lovett, "Domination: A Preliminary Analysis," *Monist* 84, no. 1 (Jan. 2001): 98–112.

42. John Braithwaite, *Not Just Deserts: A Republican Theory of Criminal Justice* (Oxford [England], New York: Clarendon Press; Oxford University Press, 1990), 95. (emphasis added)

43. Lesley Ellen Harris, *Licensing Digital Content: A Practical Guide for Librarians*, 2nd ed. (Chicago: American Library Association, 2009).

44. Lipinski, *The Librarian's Legal Companion for Licensing Information Resources and Services*, 476.

45. Kristin R. Eschenfelder et al., "How Institutionalized Are Model License Use Terms? An Analysis of E-Journal License Use Rights Clauses from 2000 to 2009," *College & Research Libraries* 74, no. 4 (2013): 351.

46. Xiaohua Zhu and Kristin R. Eschenfelder, "Social Construction of Authorized Users in the Digital Age," *College & Research Libraries* 71, no. 6 (Nov. 1, 2010): 548–68.

47. Ibid.

48. Karen A. Coombs, "Protecting USER PRIVACY in the Age of DIGITAL LIBRARIES," *Computers in Libraries* 25, no. 6 (June 2005): 16–20; Peter E. Murray and Association of Research Libraries, *Library Patron Privacy: SPEC Kit*, SPEC Kit 278 (Washington, D.C.: Association of Research Libraries, Office of Leadership and Management Services, 2003).

49. Ted Bergstrom, Paul Courant, and R. Preston McAfee, "Big Deal Contract Project," 2009, available online at www.econ.ucsb.edu/~tedb/Journals/BundleContracts.html [accessed 7 April 2015].

50. Eschenfelder et al., "How Institutionalized Are Model License Use Terms," 327.

51. Because our data set included licenses through 2009, we contacted the libraries whose licenses were in our data set to determine whether the licenses were the latest agreements between the libraries and publishers. We were unable to obtain information from all libraries with respect to all publishers. However, most of the licenses about which we were able to obtain information were still in effect through 2013. As part of this process, we also collected seven new licenses. Close reading of these licenses indicate few changes in the privacy-affecting provisions examined in this paper.

52. Eschenfelder et al., "How Institutionalized Are Model License Use Terms."

53. "Liblicense Model License Agreement & Commentary" (2008), available online at http://liblicense.crl.edu/wp-content/uploads/2011/09/licenseagreements/standlicagree.pdf [accessed 7 April 2015]; "Standard License Agreement: Publisher and the Regents of the University of California," *California Digital Library* (2011), available online at www.cdlib.org/gateways/vendors/docs/Model_License_LATEST_Revised_10-09.docx [accessed 7 April 2015].

54. Lipinski, *The Librarian's Legal Companion for Licensing Information Resources and Services*; Harris, *Licensing Digital Content*.

55. Trina J. Magi, "A Content Analysis of Library Vendor Privacy Policies: Do They Meet Our Standards?" *College & Research Libraries* 71, no. 3 (May 2010): 254–72.

56. Requirements that libraries monitor patron activity have relatively recent precedent. When photocopy services first appeared in libraries, they were provided by library staff. To limit liability for copyright infringement, most libraries required patrons to fill out a request form affirming that the copies would be used only for the requestors' private study. Betsy A. Bernfeld, "Free to Photocopy?" *Legal Reference Services Quarterly* 25, no. 2/3 (2006): 1–49, doi:10.1300/J113v25n02_01. When libraries began providing self-service copiers, there was substantial discussion of libraries' responsibility to monitor patron copying. LaHood and Sullivan, for example, were concerned about some librarians' "laissez-faire attitude" toward users' copying, and they contended that librarians should monitor in the self-service copying. Charles G. LaHood and Robert C. Sullivan, *Reprographic Services in Libraries: Organization and Administration*, LTP Publication; No. 19 (Chicago: Library Technology Program, American Library Association, 1975). In contrast, Treece argued that Congress chose not to impose a requirement that libraries supervise photocopying, since the Copyright Act of 1976 stated that libraries were not to be held liable for copyright infringement "for the unsupervised use of reproducing equipment on its premises: Provided, that such equipment displays a notice that the making of a copy may be subject to the copyright law." James M. Treece, "Library Photocopying," *UCLA Law Review* 24 (1977): 1025–69.

57. "The licensee shall keep full and up-to-date records of all authorized users and their access

details, and if appropriate provide the Publisher with periodic lists of additions, deletions or other alterations to such records as agreed between the parties from time to time." And "Licensee shall keep full and up-to-date records of all Authorized Users and their access details, which shall be provided to Emerald upon request."

58. Lipinski, *The Librarian's Legal Companion for Licensing Information Resources and Services*, 452.

59. See Theresa Chmara, *Privacy and Confidentiality Issues: A Guide for Libraries and Their Lawyers*, 1st ed. (American Library Association Editions, Chicago, 2009) for a compendium of state library privacy laws.

60. Lipinski, *The Librarian's Legal Companion for Licensing Information Resources and Services*, 531–32.

61. Launched in 2002, the COUNTER codes set standards for libraries, publishers and intermediaries, to regulate the recording and reporting of online usage statistics "in a credible, consistent, and compatible way." Although COUNTER does not use the exact words "in aggregated form" to set restrictions on the usage report provided by publishers, it specifies in the "customer confidentiality" section of the COUNTER Codes (which is in turn based on International Coalition of Library Consortia (ICOLC) Guidelines) that publishers must obtain the "the permission of that individual user, the consortium, and its member institutions" before they release or sell "statistical reports or data that reveal information about individual users." "The COUNTER Code of Practice for E-Resources: Release 4" (Apr. 2012), 1, 29, available online at www.projectcounter.org/r4/COPR4.pdf [accessed 7 April 2015].

62. Ellen Finnie Duranceau et al., "After the License Is Signed," *The Serials Librarian* 48, no. 3/4 (2005): 340, doi:10.1300/J123v48n03_18.

63. Lipinski, *The Librarian's Legal Companion for Licensing Information Resources and Services*, 448.

64. Ibid., 451.

65. Ibid., 448.

66. American Library Association, *Intellectual Freedom Manual*, 178.

67. Ibid.

68. Sims, "Library Licensing and Criminal Law."

69. 18 U.S.C. §§ 1030(a)(2)(C), 1030(g). There is some dispute regarding the proper interpretation of "unauthorized" and "exceeds authorized access" in the CFAA. The Ninth Circuit Court of Appeals, for example, has determined that using information one has authorization to access for unauthorized purposes is *not* unauthorized access or exceeding of authorized access. *U.S. v. Nosal*, 676 F.3d 854 (9th Cir., 2012); see also *U.S. v. Drew*, 259 F.R.D. 449 (C.D. Cal., 2009) (holding that providing false information to a website, in violation of terms of service does not constitute unauthorized or exceeding of authorized access). However, these cases do not appear to provide grounds for avoiding criminal charges in cases like Swartz's. Lee Goldman distinguishes several interpretations of unauthorized access and exceeding unauthorized access: an "agency" approach, such that acting contrary to one's employer's interest violates the law; a "contract" approach, such that violating an employment or terms of service contract violates the law; and a "plain meaning" approach such that only access, and not use, of information violates the law. "Interpreting the Computer Fraud and Abuse Act," *Pittsburgh Journal of Technology Law & Policy* 13, no. 1 (2012). Under any of these interpretations, the CFAA would appear to give rise to liability for unauthorized use or exceeding authorized use of licensed library materials.

70. Lovett, "Domination."

71. MIT, *Report to the President*, 84–85.

72. Lovett, "Domination," 101. A related issue of privacy and arbitrary power in electronic resources concerns potential loss of attorney-client privilege. See Anne Klinefelter, "When to Research is to Reveal: The Growing Threat to Attorney and Client Confidentiality from Online Tracking," *Virginia Journal of Law & Technology* 16, no. 1 (Spring 2011): 22–30.

73. Rubel, "Libraries, Electronic Resources, and Privacy," 200.