# Will An Institutional Repository Hurt

## Calming the faculty fear

By James M. Donovan and Carol A. Watson

Librarians have every reason to support the creation of an institutional digital repository (IR). An IR preserves the output of the intellectual life of the school, enables anyone with internet access to enjoy the benefits of the new knowledge, and promotes the institution and scholar by bringing to the foreground their intellectual achievements.

Plans for a new IR project within the law school, however, can quickly find such worthy motives swept aside as faculty members invariably voice some version of the following comments: "Won't posting my articles elsewhere steal downloads away from SSRN? That would lower my rankings in SSRN and perhaps reduce my professional stature."

One can regret that law academics today reflexively cower at the thought of appearing to perform poorly on any new ranking system that crosses their path, no matter how dubious. Even so, there can be no denying that SSRN, or the Social Science Research Network, has earned a respectable cachet among the professoriate. This is a tool they believe they understand and with which they've grown comfortable. The proper response, then, is not—however tempting it may be—to point out that ranking by downloads is an easily gamed and essentially meaningless metric. Rather, the more successful strategy appeals to the fact that such fears are based upon a flawed appreciation of how readers connect with scholarship of interest.

### What's the Difference?

The question assumes a fundamentally zero-sum view of readers. In this model, a fixed number of readers exists for any given posted article. If the piece is available in only one place, such as SSRN, then all these readers will access the file from SSRN. By concentrating that limited readership in one place, the article and author enjoy their maximum ranking. Should another version of the article become available, as in an IR, that limited audience becomes split, divided between SSRN and the IR. Every download in the repository signifies a lost download by SSRN and vice-versa.

Such folk sociology can be remarkably resistant to correction, not least because it *could* be true. There is nothing obviously false in the view that multiple versions divvy up a limited audience and consequently that the effect of an IR, aside from all the larger virtues it promises, will be to lower the status of any individual author in the SSRN rankings from what it otherwise would have been. This possibility reasonably motivates faculty to jealously shield their SSRN download statistics from potential dilution by a competing website.

But just as the argument is not obviously false, neither is it necessarily true. While the total number of readers of any given work is certainly finite, this fact can lead to the mistaken conclusion that it is therefore also bounded. In other words, if the SSRN and IR copies both get 100 downloads, we needn't leap to the conclusion that without the IR copy the SSRN downloads would have been 200. There is at least as good an argument that the 100 IR downloads represent *new* readers who would otherwise not have found the piece at all, yielding a net increase in the audience.

Choosing between these competing scenarios cannot be based on mere rhetoric but instead must be based on the facts. It can be shown, we believe, that the zero-sum fear is unwarranted. SSRN and IRs more likely draw from different readerships, meaning that downloads recorded for the repository copy represent not diverted SSRN readers but a new audience for the content. SSRN and IRs do not fight for the same eyeballs, but instead target different populations defined by how readers find their way to the desired content.

SSRN, through use of subscription lists and institution-created paper series, intends to saturate the small but keenly interested audience of legal scholars. Through these services, legal scholars who have an ongoing interest in a given topic will become quickly apprised of new content, allowing SSRN to fulfill its primary function as a current awareness tool and distribution point for drafts and preprints. An IR, by contrast, excels at calling its content to the attention of those with an acute need for specific information tied to a particular project. These users typically identify a paper not through subscribing to paper series but by doing keyword searches in web browsers like Google.

We do recognize that these differences are more of degree than of kind. Still, they follow reasonably from how the different platforms view their own strengths and where they put the majority of their development resources. SSRN, whose content is also discoverable through Google, earns its profits by subscribing schools and journals to papers series to be pushed to subscribers. More recently, it has sought to generate additional revenue by selling bound copies of the deposited articles. Presumably, SSRN works hard to make those features efficient and useful to both the content creators and end users. IRs, on the other hand, derive their primary benefit from visibility on the web, and, consequently, the more successful of these repositories, even when they allow subscription to RSS feeds and other alert tools, invest much expertise to make the content discoverable by web crawlers.

All told, then, the target SSRN audience is comparatively small and stable while that for the IR is considerably larger but amorphous, in constant flux. We would expect this difference to be somehow reflected in the download patterns within each resource.

### Download Comparison

In the main, one would anticipate SSRN downloads to experience a quick burst of activity triggered by appearance of the paper abstract in one or more of its paper series, followed by a plateau after saturation of the target audience. IR downloads would display continual increases as the content is discovered by an ever-changing audience of short-term users.

The data available to compare downloads between SSRN and IRs are simple snapshots, usually showing that IRs enjoy more downloads than does SSRN. We can pause a moment to consider this finding, which becomes a true puzzle in the zero-sum worldview. If both copies are dividing a limited audience, one would expect that SSRN should be as likely to surpass the IR in downloads. Perhaps more likely is that SSRN should typically receive a greater share of downloads since it "pushes" the item to the most interested audience. That the reverse occurs lends presumptive credibility to an alternative, non-zero-sum understanding of readership.

Considered alone, however, synchronic comparisons can offer no conclusive support for either zero-sum or non-zero-sum perspectives. As we have described the problem, the signs to distinguish between the two must be

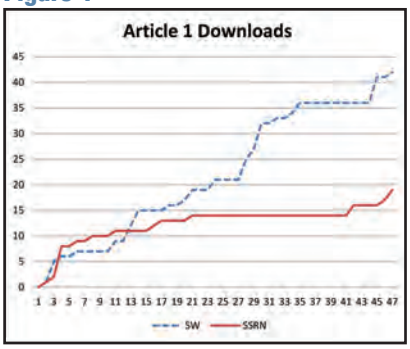


© 2012 James M. Donovan and Carol A. Watson

read not in final download statistics but rather in the patterns of downloads over time. It is to the latter, not the former, that we must look to choose between the models.

Toward that end, we collected illustrative data on two articles. Article 1 was deposited in both SSRN and Selected Works (SW), which is the personal module of bepress's institutional repository platform, Digital Commons. The primary distinction between the two is that the SW page follows the faculty member if he or she leaves to join a different law school. Each Monday, for 47 weeks total, downloads of each paper were recorded and mapped in Figure 1.

**Figure 1**



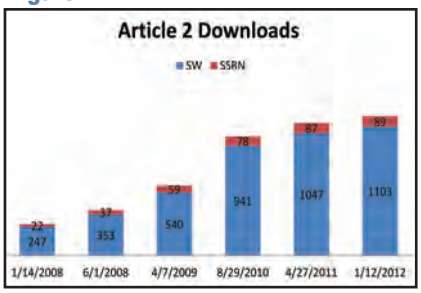

**SW versus SSRN Downloads**

*(Weekly comparison from November 30, 2009, to October 18, 2010)*

The results appear to conform to the predictions of the nonzero-sum model: SSRN downloads initially outpace those from SW with the biggest jump following its announcement in SSRN subject matter journals in week four. But these hits soon experience an extended plateau. The SW copy, after a slower start, evidenced a consistent increase in downloads. This is the pattern one would expect when one version is initially "pushed" to a small audience that is quickly saturated while the other receives consistent preference in web browser results.

The same story is told by approaching the question through a different methodology. Article 2 was uploaded into SW in early 2007 and into SSRN a few months later (in March and June, respectively). From 2008 to 2012, five periodic download totals were recorded (Figure 2). SW was clearly more successful at finding readers for this article than was SSRN, but that fact does not dispel the zero-sum worry. Perhaps all the SW readers *would* have downloaded from SSRN had the IR copy not been available.

**Figure 2**

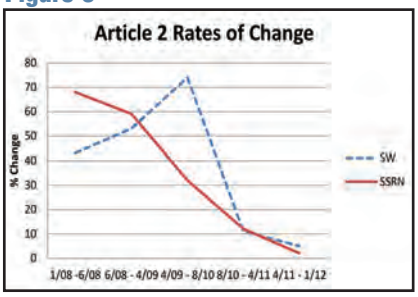



**SW versus SSRN Downloads**

*(Periodic comparisons from 2008-2012)*

To address this more specific question, we looked at the rates of change among the six snapshots (Figure 3). The results again contradict the zero-sum expectations, fitting more in line with the contrary view. After a greater rate of change, SSRN downloads precipitously fall while those for SW rise significantly, falling to SSRN levels only much later.

**Figure 3**




**SW versus SSRN Downloads**

*(Percent change from 2008-2012)*

These two case studies are, of course, insufficient to conclusively settle the argument between these two mutually exclusive models of how readers connect with articles of interest. The zero-sum position expects either that SSRN would be the preferred source for this legal scholarship or that readers who find the content through keyword searching in web browsers would be equally shared between SSRN and the IR alternative. Neither of our case studies support these expectations: contrary to the expected greater SSRN success, in both tests the articles found fewer readers in SSRN than through the IR. The evidence further suggests that over time readers preferentially access the non-SSRN version of the article, contradicting the second prediction of the zero-sum model. These results instead support the non-zero-sum model, in which IR downloads represent penetration of the content into new audiences outside that achieved by SSRN alone.

If the data favor the non-zero-sum model, we can give some thought as to the actual mechanism by which this is achieved. Our own experience suggests that much depends on how SSRN and the IR interact with Google. Commercial IR products, such as bepress's Digital Commons and Selected Works, appear to be more transparent to Google than is SSRN. For example, common keyword searches that apply to both the case study articles routinely return Google results with the IR version as the first entry, or at least on the first results page, while the SSRN version, which contains exactly the same content, is not listed until much later. Another measure of the greater Google-compatibility by IRs than by SSRN is that when tracked, we have found simultaneously deposited IR content appears in Google searches a week or more before the SSRN copy.

Transparency to web browsers offers a reasonable mechanism accounting for these data. The upshot is that many of the hits experienced by IRs will be "new" downloads, not diverted SSRN downloads. These users often would not have found the SSRN version, especially as studies show most users don't look past the first page or two of Google results.

## Conclusion: Use Both!

Faculty members should not view the proposed IR as a drain on their SSRN rankings. While SSRN excels at delivering their work to the cadre of legal specialists, IRs typically do a better job of presenting it to a broader readership. This expanded exposure should be judged a

(continued on page 16)

 continued from page 13

positive benefit of participation in the IR, helping to mitigate criticisms of law faculty as sequestered, insular, and writing only for themselves. Anyone interested in giving their ideas the widest possible hearing should deposit their intellectual work in as many venues as possible. For law professors, this means they should have both SSRN and the IR working for them. ■

***James M. Donovan,*** *J.D., Ph.D., (james.donovan@uky.edu) is director and associate professor of law at the University of Kentucky College of Law Library in Lexington, and* ***Carol A. Watson*** *(cwatson@uga.edu) is director of the law library at the University of Georgia Alexander Campbell King Law Library in Athens.*