When The Licensor Calls: Some Thoughts On Systematic Downloading
by Andrew Waller

Systematic downloading rears its head when someone downloads a large amount of licensed electronic content, beyond what would be taken during "normal" use, in a very short period of time. As one of the main contacts with publishers and vendors for the library in which I work (University of Calgary, Alberta, Canada), I am often the first to be informed about a systematic downloading incident and I am part of the team that ameliorates the situation. I share a few thoughts on the problem here.

Systematic downloading (also known by a few other terms) is very much frowned upon by licensors. Clauses and wording dealing with systematic downloading appear in virtually every electronic product license, especially in e-journal licenses. Licenses that are skimpy on other issues will often still cover systematic downloading. A good example of a fairly fulsome systematic downloading clause can be seen in the MIT Press Institutional License Agreement, which states "…that bulk downloading is expressly forbidden, as is serial or cumulative downloading that results in the user reproducing on his or her computer, in effect, all or any substantial subset of any issue of any journal" (http://www.mitpressjournals.org/userimages/ContentEditor/1144768735931/ILA1.doc).

As systematic downloading is of high concern to publishers and vendors, libraries are quick to hear about it when it occurs. In my experience, the notice always comes by email and can be at varying levels of formality, ranging from a stern "form email" to a casual message asking something along the lines of "Hey, Andrew. What's up with this?" Also varying is the threat level on the part of the licensor; some say that access to the licensed content will end off if no satisfactory resolution to the problem emerges while others cut off access immediately, at varying IP levels (denying access to the entire subscribing institution is not uncommon), and will turn access back on only when the issue has been dealt with. Accordingly, libraries respond to systematic downloading incidents with some urgency.

At the University of Calgary, we've found that systematic downloading occurs a few to several times a year (such as the same day I finished writing this article!) and often involves an open proxy server being left on the system; sometimes this is deliberate, sometimes it is accidental. There is also evidence in many incidents that some sort of software is employed to aid in the rapid downloading of content; this can be seen in the faster-than-humans-can-type downloading statistics. In addition, we've recently experienced systematic downloading that has come about as a result of login information being "loaned" by authorised users to non-authorised friends or, in fact, being outright stolen.

While many systematic downloading incidents are clear, overt attempts to grab as much content as possible in as little time as possible, there are situations where this may not be the case.   I mention three here:

Threshold Problems
Some licensors set the thresholds that will indicate systematic downloading too low, both in terms of the amount of content and the amount of time in which the downloading occurs.

Looking at the former, at the University of Calgary, our download profile is often more like a larger institution (e.g. Harvard) than a typical library of our size (just under 30,000 students); like users elsewhere, ours very much like electronic resources. On at least one occasion, this has caused a publisher to deem downloading activity by our users to be systematic even though it was well within academic user norms and practices (i.e. many users were involved, the content downloaded was from a large number of products, etc.). The problem was rectified when the publisher raised their download threshold.

Downloading Entire Issues
The downloading of an entire issue of a journal, or even a majority of a journal issue, is sometimes prohibited in licenses; it is viewed as a form of systematic downloading (see the MIT license above). But should this be the case? There are many valid reasons why someone would want to download a complete issue; for example, a journal could produce a theme issue, where all of the articles are are on the same topic, or a journal could be key to a user's area of interest, to the point where all of the content in an issue is of value to that user.

Future Uses
The prohibitions in licenses on systematic downloading may conflict with future demands from users. For instance, I suspect that there will be greater call in the years to come for data mining and deep textual analysis of licensed content, activities that could easily be perceived as systematic downloading. There may well be important scholarly reasons for these activities but they would still be forbidden by license and, accordingly, would still set off alarms.

To conclude, I think we are still in the "early days" of the systematic downloading and there is much that both licensees and licensors can do to better deal with the problem. I think that there are also many questions that merit further examination. These include:

- *Are the instances of systematic downloading increasing, decreasing, or staying about the same?* I'm not sure. My impression is that, at my institution, there is roughly the same amount every year but others may have different experiences.

- *Why does systematic downloading occur?* I think many assumptions are made about the reasons why this activity goes on, in part because we usually don't catch the perpetrators. Yes, the student just about to go back to his or her information-poor home country exists as does the recent graduate newly-arrived in a workplace without the resources he or she was accustomed to at school; both sometimes engage in systematic downloading but I suspect that there are other reasons why it happens.
- *Are all incidents of systematic downloading noticed?* Most publishers and vendors seem to have a good handle on how their content is being used but are some better at others in doing this? Are some users far ahead of licensors and libraries in terms of hiding excessive downloading activities?
- *What do we know about the software used for systematic downloading?* From what I have heard, this software can be fairly easily accessed on the Web.   What do we know about this software?
- *What is happening in other types of libraries, outside of academia?* My experience and knowledge about systematic downloading comes almost solely from the university environment. But other libraries license electronic products as well and they can be subject to the same sort of attacks. It would be interesting to hear what has happened to them, if anything, and how they dealt with the experience.
- *How do we prevent systematic downloading?* User education and technical solutions are always mentioned but are there other options? How can we employ user education and technical solutions better?

I welcome thoughts in response to these questions and any other aspects of systematic downloading. I can be reached at waller@ucalgary.ca or at (403) 220-8133.

Further reading:

Baker, Gayle, and Carol Tenopir. "Managing the unmanageable: systematic downloading of electronic resources by library users". *Journal of Library Administration*, 44, no. 3/4 (2006): 11-24.

Duranceau, Ellen Finnie, Deanna Graham, and Stephen Martin. "After the license is signed: collaboration to resolve license breaches". *The Serials Librarian*, 28, no. 3/4 (2005):

Copyright 2007 by Andrew Waller

About the author:
Andrew Waller is Serials Librarian at the University of Calgary Library, where he oversees the serials collection. He has written and presented on topics such as e-journals, electronic product licensing, and Open Access, among others. Later

in 2007, his first licensing and access column will appear *The Serials Librarian*. In his spare time, Andrew hosts a radio show, "Breaking the Tethers", on CJSW FM (www.cjsw.com), playing mostly freaky jazz.