

Can RECAP Turn PACER Around?

by Blackwell N. Shelley Jr.

PACER is the acronym for Public Access to Court Electronic Records, a Web-based service run by the Administrative Office of the United States Courts (AO). PACER began in 1989 as a pilot program serving a few U.S. district and bankruptcy courts. Beginning in 1990, the federal Judicial Conference, under the direction of Congress, prescribed fees for the use of PACER. In those dark days before the Internet, PACER was a bulletin board service with dial-up access and it cost a dollar a minute to use. The per minute fee decreased during the 1990s, until 1998, when the federal judiciary implemented the new Case Management/Electronic Case Files (CM/ECF) system. CM/ECF, which was Web-based, dropped the per minute charge entirely and substituted a seven cents per page user fee to download or view documents in case files. Currently, the fee is eight cents per page, with some exceptions.

The PACER fees trace their origin to 1988, when the judiciary sought congressional funding to establish electronic public access services. Rather than appropriating funds for this purpose, Congress directed the judiciary to fund that initiative through the collection of user fees. (See 28 U.S.C. § 612.) As a result, PACER has always relied on fee revenue. These revenues, however, have far outstripped expectations. According to the 2006 annual report of the Judicial Information Technology Fund, the federal judiciary collected \$62.3 million in electronic public access fees in 2006, resulting in a budget surplus (for PACER alone) of \$32.2 million.¹ By 2008, the director of the Administrative Office of the U.S. Courts reported that revenues from the PACER user fees would be “used to finance other expenses related to electronic public access to the courts in areas such as courtroom technology and the Bankruptcy Noticing Center.”² According to the *New Jersey Law Journal*, PACER’s unspent revenues were \$76.8 million for the 2008 fiscal year.³ Clearly,

PACER’s user fees are a significant source of revenue for the third branch, but do PACER’s users get what they pay for?

Law librarians have criticized PACER, saying that documents downloaded from PACER cannot be authenticated. They have been circulating a petition asking the AO to digitally sign each document filed on the system using readily available technology. And law librarians have criticized the cost and poor design of PACER and have requested that depository libraries get free access.⁴

Librarians are not alone in criticizing PACER.

Enter Carl Malamud. Malamud, depending on your perspective, is either a hero or villain of the Internet. He is partially responsible for creating the first Internet radio station, for putting the U.S. Securities and Exchange Commission’s EDGAR database online and, recently, for persuading the Government Printing Office (GPO) to create a standard for publishing the *Federal Register* online, for free, in XML format. Malamud operates the nonprofit Public.Resource.Org, which, among other things, advocates that public records should be freely available on the Web.

Not surprisingly, PACER’s per page charge irks Malamud, who believes that public access to court records should be open and free. (Malamud has also complained publicly about PACER’s hit-or-miss record related to the publication of personal identifying information, such as addresses and Social Security numbers.) Malamud’s organization has been amassing case law, codes, and treatises from public domain sources and, in some instances, by purchasing the rights and making the collection available online for free in the Internet Archive.⁵ Malamud’s efforts have been controversial.

In the fall of 2008, the GPO experimented with giving PACER away for free at seventeen select libraries around the country. Twenty-two year old pro-

grammer Aaron Swartz seized the opportunity to make a contribution to Public.Resource.Org. On one of the computers at the Seventh Circuit U.S. Court of Appeals library, Swartz installed a small Perl (dynamic programming language) script that, every three seconds, downloaded a new PACER document. Over the course of several weeks, Swartz moved 780 gigabytes of data — 19,856,160 pages of text — from PACER to an Amazon cloud server. Swartz then donated the documents to Public.Resource.Org.

The GPO and the AO were not pleased. The free access experiment was abruptly discontinued. Amazon identified Swartz to the FBI. According to a Freedom of Information Act request made by Swartz, the FBI checked Swartz’s Facebook page, his work history with the U.S. Department of Labor, any outstanding warrants and prior convictions, and his mobile phone number against its federal wiretap or pen register records. They checked him against the records in a private data broker’s database and considered a stakeout of his house. On the advice of his counsel, Swartz declined invitations to discuss his exploit and, ultimately, the FBI dropped the investigation.⁶

Swartz’s Perl script, referenced above, originated with Stephen Schultze, a fellow at the Berkman Center for Internet and Society at Harvard University. Schultze has also been a critic of PACER’s user fees, and suggests that they violate Section 205 (e) of the E-Government Act of 2002, which amended then-existing law to state that “the Judicial Conference may, *only to the extent necessary*, prescribe reasonable fees [for PACER].” In August 2009, Schultze became the associate director of Princeton University’s Center for Information Technology Policy (CITP).

RECAP continued on page 58

Also in August, the CITP started the RECAP project.⁷ RECAP (motto: “Turning PACER around”) is a free plug-in for the Firefox Web browser. RECAP works like this: If you run a PACER search, RECAP checks your query against the free database at Internet Archive (www.archive.org). If the document is already there, RECAP will show an on-screen icon, you can get the document from the public source, and you can skip the user fee. If the document is not in the public database and you choose to download it, RECAP automatically posts the new document to the free database.

The AO’s response to RECAP was terse, but the office did not summon the FBI.⁸ The position of the U.S. Courts is that if a PACER user is exempt from the user fees, then the user may not redistribute a downloaded document; otherwise, any PACER user who pays the user fee may save a copy of the document on the public database. So, officially, RECAP is legitimate and legal.

The documents in the RECAP database at Internet Archive are, however, heirs to the flaws of the documents in the PACER database. There is no way to know whether documents in the RECAP database are genuine copies of the documents in the PACER system. Likewise, if a document in the PACER system contains unredacted personal identifying information, then the RECAP document will also contain unredacted information. (For now, the RECAP creators have requested that

the RECAP database not be indexed by search engines, in order to keep the information relatively unknowable.) Finally, while documents created by a court are not subject to copyright protection, there is no clear answer to the question of whether copyright protection can be extended to pleadings drafted by counsel or pro se parties.

Although there is no apparent connection to RECAP, the federal judiciary’s Electronic Public Access Program is conducting a self-assessment of PACER to end in 2010.⁹ The survey asks interested PACER users the participant to rate his or her satisfaction, explain the rating, and pick one thing to change about the system. According to the AO, the survey results will help define the next generation of PACER. Meanwhile, PACER will cost eight cents per page for the foreseeable future.

Endnotes:

- 1 See Judiciary Information Technology Fund Annual Report for Fiscal Year 2006, at <http://www.scribd.com/doc/2436289/>, last visited November 12, 2009.
- 2 See Annual Report of the Director, Activities of the Administrative Office of the U.S. Courts, James C. Duff, Director, at <http://www.uscourts.gov/library/annualreports/2008/index.cfm>, last visited November 12, 2009.
- 3 See *New Jersey Law Journal*, September 2, 2009, “Free Web Access to Judicial Records Gladdens Public but Worries Some Courts,” available at http://www.law.com/jsp/law/sfb/lawArticleSFB.jsp?id=1202433517232&Free_Web_Access_to_Judicial_Records_Gladdens_Public_but_

- 4 The petition is available at <http://www.thepetitionsite.com/1/improve-PACER>, last visited November 12, 2009.
- 5 The current collection is at: <http://bulk.resource.org/courts.gov/>, last visited November 12, 2009.
- 6 See “FBI Investigated Coder for Liberating Paywalled Court Records,” available at <http://www.wired.com/threatlevel/tag/aaron-swartz/>, last visited November 12, 2009. See, The New York Times News Blog, “The Lede: Steal These Federal Records — Okay, Not Literally,” Schwartz, J. and Mackey, R., available at <http://thelede.blogs.nytimes.com/2009/02/13/steal-these-federal-records-okay-not-literally/?ref=us>, last visited November 12, 2009. See New York Times, “An Effort to Upgrade a Court Archive System to Free and Easy,” Schwartz, J., at http://www.nytimes.com/2009/02/13/us/13records.html?_r=1, last visited November 12, 2009.
- 7 See <https://www.recapthelaw.org/>, last visited November 12, 2009.
- 8 See <http://pacer.psc.uscourts.gov/announcements/general/exemptnotice.html>, last visited November 12, 2009.
- 9 Electronic Public Access Program/PACER Assessment Begun, The Third Branch, available at http://www.uscourts.gov/ttb/2009-09/article05.cfm?WT.cg_n=TTB&WT.cg_s=Sep09_article05_tableOfContents, last visited November 12, 2009.