

This is PDF of story that may still be available online at:  
<http://radar.oreilly.com/2006/06/nsa-datamining-social-networki.html>

[Home](#) | [Shop](#) | [Radar: News & Commentary](#) | [Answers](#) | [Safari Books Online](#) | [Conferences](#) | [School of Technology](#) | [Community](#)



## NSA Datamining Social Networking Sites

by [Tim O'Reilly](#) | [@timoreilly](#) | [+Tim O'Reilly](#) | [Comments: 7](#) | June 9, 2006

[Richard Forno](#) wrote on Dave Farber's [IP List](#): "[New Scientist Magazine has discovered](#) that Pentagon's National Security Agency, which specialises in eavesdropping and code-breaking, is funding research into the mass harvesting of the information that people post about themselves on social networks. And it could harness advances in internet technology – specifically the forthcoming "semantic web" championed by the web standards organisation W3C – to combine data from social networking websites with details such as banking, retail and property records, allowing the NSA to build extensive, all-embracing personal profiles of individuals."

[Slashdot picked up the story](#), so many of you have probably already seen it. However, both the New Scientist story and much of the followup laid a misleading trail in the focus on social networking sites. The point is that one of the side effects of Web 2.0, the internet as platform, is that we all leave tracks in cyberspace, whatever sites we use. The kind of work described in this story is going on not just at the NSA, but in commercial operations doing data mining and profiling for everything from direct sales to political fundraising. And while a lot of this work is going on behind closed doors, some of it is open to the public. Look at [fundrace](#) and [zillow](#) (and trace an individual from one to the other), and you'll see how the tools of surveillance and privacy invasion are becoming "democratized" (so to speak.)

And this isn't all bad — [the paper that triggered the story](#) was not about datamining MySpace and similar networks at all. It was about looking for conflicts of interest among reviewers of scientific papers, which would be a useful tool indeed, and one very relevant to the story I posted the other day about [Nature's experiments with open peer review](#). (What led New Scientist to the NSA connection was a footnote that identified an NSA group called ARDA as one of the funders of the research.)

Another interesting tidbit to me was the fact that the DOD has renamed that NSA group "The Disruptive Technology Office." I'm sure that they are looking at a lot more than data mining ... as should all of you.

P.S. Slashdot clearly got the story from IP, since it paraphrases Richard rather than the New Scientist story itself, but doesn't credit him as the source. I'd love to see more bloggers using (and acknowledging) mailing lists as their source, because there are a lot of very interesting conversations that happen in other places than blogs...

Semantic Analytics on Social Networks:  
<http://www2006.org/programme/files/pdf/4068.pdf>