Forgot What You Searched For? Google Didn't
By Leslie Walker

Saturday, January 21, 2006; Page D01

The Justice Department may have done us all a big favor by issuing subpoenas to Internet search engines to find out what people are researching online.

Not because that data could help shield children from online porn, which was the government's stated goal in demanding data from Google and three other search firms.

Rather, the request -- and Google's refusal to fork over its search data -- is putting a helpful public spotlight on the vast amount of personal information being stored, parsed and who knows what else by the Web services we increasingly rely on to manage our lives.

Even though the government has demanded no personal information -- only a list of Web queries divorced from the names of those submitting them -- Google is resisting partly on grounds that turning over the data might create a public perception that it would readily cough up personal factoids, if asked.

So that raises the question: What, exactly, does Google know about us?

In my case, a lot.

I've done a great deal of beta testing of Google services, including Gmail, Orkut social networking, Froogle shopping lists, personal search and a custom home page. Most are linked by my Gmail address and account name.

Google has a wealth of data about me, especially through its personal search service, a tool that only collects data on you if you elect to turn it on, as I have.
That service gives me -- along with Google, and maybe the government should it ever suspect me of a crime -- access to every query I've typed while signed into Google, organized by a clickable calendar.

Clicking on "Nov. 3" produces a page listing all 27 queries I submitted while signed into Google that day. I'm not sure I'd want the government to see the ones on "panties" and "underpants." (Sorry, but I'm not going to tell you why I entered those words, except to say it was unrelated to porn.) And it's no one's business why I looked up "Herman Miller chair," "redhead" or "Ocean City" either.

My stored history is so detailed it shows I clicked on none of the results from those queries, but I did click on results from four searches that day. The five sites I visited are even listed.

Google doesn't keep such detailed data on anonymous users who don't sign in. Unless users tweak their Web browser settings, Google stores a "tracking cookie" or small file on each user's computer to store items such as the address of their computer, type of Web browser used, and date and time of each query submitted.

A Google spokesman said that data are not currently correlated with each user's search query, but Google's technology and privacy policies would allow the company to do so if it chose.

Search histories already are creeping into criminal trials. A North Carolina man, Robert Petrick, who was convicted in November of murdering his wife, ran suspicious Internet searches immediately before and after she was dumped in a lake. His queries? "Body decomposition," "rigor mortis," "neck," "snap" and "break," along with topics relating to the depth of the lake where her body turned up.

Those searches were stored on the hard drives of the computers Petrick used, but they could just as easily have been stored by Google had Petrick turned on the archiving feature that I use.

Our personal search histories are highly sensitive information -- and obviously open to misinterpretation -- because they offer such a unique view into what we are thinking. Most of us routinely ask Google questions about religion, social behavior, sex, work -- whatever pops into our heads.

And those queries are mere rocks in a growing mountain of profiling data about us being compiled by many other Web services, not just Google. Over at Amazon, hackers or government investigators might have a field day if they gained access to the 171 items on my supposedly private "wish list." (I'm too lazy to ever delete anything, and I use Amazon's wish list as a bookmarking tool.)

It's one thing for our personal data to be stored on our own computers, which theoretically we could erase (a harder task than it seems, actually) whenever we choose. It's quite another to have so much personal activity logged and analyzed by distant, impersonal Web sites. There is simply no telling how much long-term control we are giving up over our digital reputations in these still-early days of the Web.
So if the government scares people into thinking more about their own Internet histories by slapping subpoenas on the search engines, maybe that's not a bad thing.