



You Are What You Search

AOL's data leak reveals the seven ways people search the Web.



By Paul Boutin
Friday, Aug. 11, 2006, at 5:30 PM ET

AOL researchers recently published the search logs of about 650,000 members—a total of 36,389,629 individual searches. AOL's search nerds intended the files to be an academic resource but didn't consider that users might be peeved to see their private queries become a research tool. Last weekend, the Internet service provider tried to [pull back the data](#), but by that point it had leaked all over the Web. If you've ever wanted to see what other people type into search boxes, now's your chance.

The search records don't include users' names, but each search is tagged with a number that's tied to a specific AOL account. The *New York Times* quickly [sussed out](#) that AOL Searcher No. 4417749 was 62-year-old Thelma Arnold. Indeed, Arnold has a "dog who urinate on everything," just as she'd typed into the search box. [Valleywag](#) has become one of many clearinghouses for funny, bizarre, and painful user profiles. The searches of AOL user No. 672368, for example, morphed over several weeks from "you're pregnant he doesn't want the baby" to "foods to eat when pregnant" to "abortion clinics charlotte nc" to "can christians be forgiven for abortion."

While these case studies are good voyeuristic fodder, snooping through one user's life barely scratches the surface of this data trove. The startup company I work for, [Splunk](#), makes software to search computer-generated log files. AOL's 36 million log entries might look like an Orwellian nightmare to you, but for us it's a user transaction case study to die for. Using the third-party site [splunkd.com](#), I've parsed the AOL data to create a typology of AOL Search users. Which of the seven types of searcher are you? (Click [here](#) for tips on how to do this yourself.)

The Pornhound. Big surprise, there are millions of searches for mind-bendingly kinky stuff. User No. 927 is already an Internet legend—[click here](#) if you're not faint of heart (and not at the office). When I clicked Splunk's "Show Events by Time" button, though, I found that porn searchers vary not only by *what* they search for, but *when* they search for it. Some users are on a quest for pornography at all hours, seeking little else from AOL. Another subgroup, including No. 927, search only within reliable time slots. The data doesn't list each user's time zone, but 11 p.m. Eastern and 11 p.m. Pacific appear to be prime time for porn on AOL's servers. My favorite plots show hours of G-rated searches before the user switches gears—what I call the [Avenue Q Theory](#) of Internet usage. User No. [190827](#) goes from "talking parrots jokes" and "poems about a red rose" before midnight to multiple clicks for "sexy dogs and hot girls" a half hour later. An important related discovery: Nobody knows how to spell "bestiality."

The Manhunter. The person who searches for other people. Again, I used Splunk's "Show Events by Time" function to plot name searches by date and time. Surprisingly, I didn't uncover many long-term stalkers. Most of the data showed bursts of searches for a specific name only once, all within an hour or a day, and then never again. Maybe these folks are background-checking job candidates, maybe they're looking up the new cutie at the office, or maybe they just miss old friends. Most of the names in AOL's logs are too ambiguous to pinpoint to a single person in the real world, so don't get too tweaked if you find your own name and hometown in there. I've got it much worse. There are 36 million searches here, but none of them are for me.

The Shopper. The user who hits "treo 700" [37 times](#) in three days. Here, the data didn't confirm my biases. I'd expected to find window shoppers who searched for Porsche Cayman pages every weekend. But AOL's logs reveal that searches for "[coupons](#)" are a lot more common. My favorite [specimen](#) is the guy who mostly looked up food brands like Dole, Wendy's, Red Lobster, and Turkey Hill, with an occasional break for "asian movie stars." How much more American could America Online get?

The Obsessive. The guy who searches for the same thing over and over and over. Looking at the search words themselves can obfuscate a more general long-term pattern—A, A, A, A, B, A, A, C, A, D, A—that suggests a user who can't let go of one topic, whether it's Judaism, real estate, or Macs. Obsessives are most likely to craft [advanced search terms](#) like "craven randy fanfic -wes" and "pfeffern**sse."

The Omnivore. Many users aren't obsessive—they're just online a lot. My taxonomy fails them, because their search terms, while frequent, show little repetition or regularity. Still, I can spot a few subcategories. There are the trivia buffs who searched "imdb" hundreds of times in three months and the nostalgia surfers on the hunt for "pat benatar helter skelter lyrics."

The Newbie. They just figured out how to turn on the computer. User No. 12792510 is one of many who confuses AOL's search box with its browser address window—he

keeps seaching for "www.google." Other AOLers type their searches without spaces between the words ("newcaddillacdeville") as if they were 1990s-era [AOL keywords](#).

The Basket Case. In college I had to write a version of the classic [ELIZA](#) program, a pretend therapist who only responds to your problems ("I am sad") with more questions ("Why do you say you are sad?"). AOL Search, it seems, serves the same purpose for a lot of users. I stumbled across queries like "i hate my job" and "why am i so ugly." For me, one log entry stands above the rest: "i hurt when i think too much i love roadtrips i hate my weight i fear being alone for the rest of my life." Me too, 3696023. Me too.

[Paul Boutin](#) is a Silicon Valley-based writer who also contributes to Business Week, Wired, and Engadget. Illustration by Robert Neubecker.