Judge rejects government demand for Google search terms

Mike Ingram
20 March 2006

In a ruling issued Friday, March 17, Judge James Ware denied a demand from the Department of Justice that search giant Google turn over samples of search terms entered into its web site.

In his ruling Ware said there were three "vital interests" raised by the case: "(1) the national interest in a judicial system to reach informed decisions through the power of a subpoena to compel a third party to produce relevant information; (2) the third-party's interest in not being compelled by a subpoena to reveal confidential business information and devote resources to a distant legislation; and (3) the interest of individuals in freedom from general surveillance by the Government of their use of the Internet or other communications media."

Though not central to Google's legal arguments, the last point here was always the most crucial in relation to this case, as the ruling itself recognizes. Judge Ware states, "On March 14, 2006, this Court held a hearing on the Government's Motion. At that hearing, the Government made a significantly scaled-down request from the information it originally sought. For the reasons explained in this Order, the motion to compel, as modified, is GRANTED as to the sample URLs from Google search index and DENIED as to the sample of users' search queries from Google's query log."

The US Department of Justice had asked the District Court of San Jose, California, to compel Google to comply with a subpoena issued last year to turn over records that detail millions of Internet searches.

Following negotiations with Google, the initial request for all URLs (web addresses) available through Google's index and all search queries between June 1, and July 31, 2005 was eventually scaled down to 50,000 URLs and searches made over a one-week period. While seeking to limit the impact of the subpoena, Google continued to oppose the government request on the grounds that it was unduly burdensome and "not reasonably calculated to lead to evidence admissible in the underlying litigation."

The underlying litigation refers to statements by the US government that the data was needed to bolster its claims that the Child Online Protection Act (COPA) does not violate the Constitution. That act was passed during the Clinton administration in 1998 under the auspices of protecting children from online pornography. It established criminal penalties for any commercial distribution of material harmful to minors. The legislation was suspended a year later after a successful suit by the American Civil Liberties Union and others claiming the act violated the constitutional right to free speech. Like all such legislation, its scope was far broader than its supposed target, making it an offense for web sites to post material deemed "harmful to minors," which, as civil rights campaigners said at the time, could criminalize sites of some art galleries and book stores.

Google also argued that handing over the data to the government could result in a loss of trust on the part of its users. In comments posted on the Internet March 17, Google Associate General Counsel, Nicole Wong said of the ruling, "This is a clear victory for our users and for our company, and Judge Ware's decision regarding search queries is especially important. While *privacy was not the most significant legal issue in this case* [emphasis mine] (because the government wasn't asking for personally identifiable information), privacy was perhaps the most significant to our users. As we noted in our briefing to the court, we believe that if the government was permitted to require Google to hand over search queries, that could have undermined confidence that our users have in our ability to keep their information private. Because we resisted the subpoena, the Department of Justice will not receive any search queries and only a small fraction of the URLs it originally requested."

Though Google's stance in refusing to hand over the data—in contrast to that of rivals AOL, Microsoft and Yahoo—should be applauded, Wong's comments reveal a dangerous level of complacency as to the seriousness of the case. From the text of the final ruling one must conclude that the federal judge was far more conscious of the implications of this case for privacy than Google's legal counsel.

In his ruling Ware states, "Google primarily argues that the information sought by the subpoena is not reasonably calculated to lead to evidence admissible in the underlying litigation, and that the production of information is unduly burdensome. The Court discusses each of these objections in turn, as well as *the Court's own concerns* about the potential interests of Google's users. [emphasis mine]"

In fact the denial of the search terms data was based upon these concerns rather than the central legal argument presented by Google. In relation to the turning over of 50,000 URLs, the judge states, "The Court finds that 50,000 URLs randomly selected from Google's database for use in a scientific study of the effectiveness of filters is relevant to the issues." On the question of undue burden, the judge states, "The Court is particularly concerned any time enforcement of a subpoena imposes an economic burden on a

non-party," but because the government had agreed to compensate Google for engineering time required to extract the data, "the Court does not find that the technical burden of production excuses Google from complying with the subpoena. Later in this Order, the Court addresses other concerns with respect to this information, however."

Those "other concerns" all focused on the issue of user privacy. After analyzing Google's privacy policy and pointing out that it guaranteed the privacy only of "personal information" and was in no way a commitment on Google's part to "guard the query log," the judge found nevertheless that "even if an expectation by Google users that Google would prevent disclosure to the Government of its users' search queries is not entirely reasonable, the statistic cited by Dr. Stark [a statistician hired by the government to analyze the Google data] that over a quarter of all Internet searches are for pornography . . . indicates that at least some of Google's users expect some sort of privacy in their searches."

He went on to state that "the Government has not demonstrated . . a substantial need for *both* the information contained in the sample URLs and sample of search query text [emphasis in original]," adding that "both the sample of URLs and the set of search queries are aimed at providing a list of URLs which will be categorized and run through the filtering software in an effort to determine the effectiveness of filtering software as to certain categories."

Judge Ware therefore finds, "this Court exercises its discretion . . and determines that the marginal burden of loss of trust by Google's users based on Google's disclosure of its users' search queries to the Government outweighs the duplicative disclosure's likely benefit to the Government's study. Accordingly, the Court grants the Government's motion to compel only as to the sample 50,000 URLs from Google's search index."

In a section subtitled "Privacy," the judge states, "Although the Government has only requested the text strings entered (Subpoena at 4), basic identifiable information may be found in the text strings when users search for personal information such as their social security numbers or credit card numbers through Google in order to determine whether such information is available on the Internet . . The Court is also aware of the so-called 'vanity searches,' where a user queries his or her own name perhaps with other information. Google's capacity to handle long complex search strings may prompt users to engage in such searches on Google . . Thus, while a user's search query reading '[user name] stanford glee club' may not raise serious privacy concerns, a user's search for '[user name] third trimester abortion san jose,' may raise certain privacy issues as of yet unaddressed by the parties' papers. This concern, combined with the prevalence of Internet searches for sexually explicit material . . generally not information that anyone wishes to reveal publicly—gives this Court pause as to whether the search queries themselves may constitute potentially sensitive information."

In conclusion, the judge turned to the issue of the possible uses of the information gathered by subpoena. "Even though counsel for the Government assured the Court that the information received will only be used for the present litigation," the judge states, "it is conceivable that the Government may have an obligation to pursue information received for unrelated litigation purposes under certain circumstances regardless of the restrictiveness of a protective order. The Court expressed this concern at oral argument as to queries such as 'bomb placement white house,' but queries such as 'communist berkeley parade route protest war,' may also raise similar concerns. In the end, the Court need not express an opinion on this issue because the Government's motion is granted only as to the sample of URLs and not as to the log of search queries."

While taking a clear stance on the issue of user privacy, Ware was careful not to set any precedent regarding the right of government to subpoena search data, stating in the conclusion to his ruling, "In particular, this Order does not address the Plaintiffs' concern articulated at the hearing about the appropriateness of the Government's use of the Court's subpoena power to gather and collect information about what individuals search for over the Internet."

When the government subpoena became public in January of this year, the *World Socialist Web* Site stated that the request for search data was essentially a fishing operation among random Internet users that served to highlight the extent of the Bush administration's attacks upon privacy and democratic rights.

In demanding such data the US government was essentially testing the water. While the court's refusal to allow access to search terms is to be welcomed, there are many troubling aspects to this case. Firstly, the court did not deny the government's demand out right, nor offer any opinion as to the legitimacy of the request being made. In agreeing to the handing over of 50,000 URLs, the court essentially upheld the right of government to forcefully solicit the assistance of Internet companies in pursuit of legal cases in which they have no interests.

But by far the most troubling aspect of the case occurred last August when the original request for data was made and three out of four of those who received the request silently complied. Even after the hearing which resulted from Google's refusal to comply, it is not known what data was handed to the US government by three of the four largest search engines.