

Disclaimer: This is a position statement for a class debate. It does not necessarily reflect the actual views of any of the authors.

October 28, 2007

Power Corrupts: Don't Tempt Google with Evil
Debate Stance Document

Chris Sosa Duane Merrill
Adrienne Felt Kevin Binswanger

Google is poised to assert itself as the world's dominant information aggregator. Google not only serves as indexer of and liaison to world's published information, it also actively and passively collects information regarding all those who use its services (e.g., web search, email, desktop search, calendaring, shopping, games, etc.).

It is certainly alarming that, as a gateway to the Internet, Google has the ability to influence public perception with search ranking results. Yet in this regard Google can be thought of like any other media outlet: it is well within its right to "spin" information to its customers (as long as it does not exhibit a monopoly over the genre).

Instead, our group focuses on the personal privacy aspect of Google's position as the dominant aggregator of electronic information. Although they are not yet *clearly* running afoul of U.S. privacy laws, the U.S. has a responsibility to have stricter, less-ambiguous privacy laws codifying how both personal and personally-identifying information can be used. Canada and the E.U. have already begun working in this direction and we want to follow suit.

Specifically, we address three issues:

First, Google's privacy policy is based on the assumption that personally-identifiable information is information that, at time of delivery to Google, personally identifies you, (e.g., your name, email address, billing information, etc.). However, Google fails to consider that when aggregating user input from its various sources and published information that it indexes, it has the ability to infer personally-identifiable data. For most of its services, any input by the user should be considered personal information that can potentially be used to construct personally-identifiable information.

Second, children under 13 are very comfortable using web search engines such as Google for both recreational and educational purposes. Most sixth grade teachers would assert that children are more inclined to use Google than their local library. Google *is* used by children, and Google *knows* this. Google is therefore subject to The Children's Online Privacy Protection Act of 1998 (COPPA). COPPA is designed to prevent website operators from aggregating personal information of children without parental consent, and Google fails to meet these provisions with its policies.

Third, users don't understand the ramifications of Google's technology well enough for the market to perfectly self-correct Google's privacy problems. This puts the technology in a niche similar to food/health products and services that require government regulation.

To solve these issues Google must comply with the following:

- 1) Google be required to post a conspicuous "warning label" about the dangers of aggregate search data. This is similar in notion to warning labels on drugs and food.

- 2) Google (and other online organizations) must meet the requirements of a new US regulation body similar to the Department of Health and Human Services that creates and enforces policies regarding the information that can be both recorded as well as disseminated to third-parties. This includes both non-aggregate and aggregate data.

In addition, in order for Google to obey COPPA they must further comply with the following:

- 3) Google needs to provide an option to "opt-out" of collection of personal information for minors. In addition to actively-collected personal information (names, addresses, email, etc.), COPPA indicates that "personal information" includes passively-collected data such as IP addresses, cookies, search terms, etc.
- 4) Google needs to provide parents with access to their child's personal information and the opportunity to delete the child's personal information. This requirement connotes *absolute* deletion, not "unless it comes at undue expense, like removing from a backup tape" as is stated in Google's current policy's "best effort" access to personal information. Personal information is any information that can be used to distinguish children (in addition to personally-identifying information).
- 5) If Google does not collect personal information, but rather "anonymized aggregate information" (as they call it in their privacy policy) regarding children, they **MUST** publicly publish their anonymization methodologies.

As a reminder, COPPA violators could be liable for up to \$11,000 *per violation*. We argue that a violation consists of an unconsented recording of a child's personal information. For example, each recorded web query by a minor can be considered a unique violation.

Witnesses:

- Annie I. Antón: Associate Professor of Software Engineering, NC State. Director of ThePrivacyPlace.org.
- Marc Klaas: Father of kidnap/murder victim Polly Klaas and founder of Klaaskids.org
- Thelma Arnold: AOL Search data lady
- Brad Templeton: Chairman of the Electronic Frontier Foundation (EFF)

Documents:

COPPA

<http://www.ftc.gov/ogc/coppa1.htm>

<http://www.ftc.gov/bcp/conline/pubs/online/kidsprivacy.shtm>

HIPAA

<http://www.hhs.gov/ocr/hipaa>

Personal Information

<http://www.nytimes.com/2006/08/09/technology/09aol.html?ex=1312776000&en=f6f61949c6da4d38&ei=5090&partner=rssuserland&>

<http://googleblog.blogspot.com/2006/03/judge-tells-doj-no-on-search-queries.html>