ALBERT GIDARI, JR., *pro hac vice*
(AGidari@perkinscoie.com)
PERKINS COIE LLP
1201 Third Avenue, Suite 4800
Seattle, WA 98101
Telephone: (206) 359-8000
Facsimile: (206) 359-9000

LISA A. DELEHUNT, Bar No. 228551
(LDelehunt@perkinscoie.com)
PERKINS COIE LLP
180 Townsend Street, 3rd Floor
San Francisco, California 94107-1909
Telephone: (415) 344-7000
Facsimile: (415) 344-7050
Attorneys for Respondent
GOOGLE INC.

UNITED STATES DISTRICT COURT

NORTHERN DISTRICT OF CALIFORNIA

SAN JOSE DIVISION

ALBERTO R. GONZALES, in his official capacity as ATTORNEY GENERAL OF THE UNITED STATES, Movant, v. GOOGLE INC., Respondent.	CASE NO. 5:06-mc-80006-JW GOOGLE'S OPPOSITION TO THE GOVERNMENT'S MOTION TO COMPEL Hearing: March 13, 2006 Time: 9:00 a.m.

CONTENTS

TABLE OF AUTHORITIES

Cases

Statutes

Regulations and Rules

Other Authorities

I. INTRODUCTION

Google users trust that when they enter a search query into a Google search box, not only will they receive back the most relevant results, but that Google will keep private whatever information users communicate absent a compelling reason. The Government's demand for disclosure of untold millions of search queries submitted by Google users and for production of a million Web page addresses or "URLs" randomly selected from Google's proprietary index would undermine that trust, unnecessarily burden Google, and do nothing to further the Government's case in the underlying action.

Fortunately, the Court has multiple, independent bases to reject the Government's Motion. *First*, the Government's presentation falls woefully short of demonstrating that the requested information will lead to admissible evidence. This burden is unquestionably the Government's. Rather than meet it, the Government concedes that Google's search queries and URLs are not evidence to be used at trial at all. Instead, the Government says, the data will be "useful" to its purported expert in developing some theory to support the Government's notion that a law banning materials that are harmful to minors on the Internet will be more effective than a technology filter in eliminating it.

Google is, of course, concerned about the availability of materials harmful to minors on the Internet, but that shared concern does not render the Government's request acceptable or relevant. In truth, the data demanded tells the Government absolutely nothing about either filters or the effectiveness of laws. Nor will the data tell the Government whether a given search would return any particular URL. Nor will the URL returned, by its name alone, tell the Government whether that URL was a site that contained material harmful to minors.

But, the Government's request would tell the world much about Google's trade secrets and proprietary systems. This is the *second* independent ground upon which the Court should reject the subpoena. Google avidly protects every aspect of its search technology from disclosure, even including the total number of searches conducted on any given day. Moreover, to know whether a

given search would return any given URL in Google's database, a complete knowledge of how Google's search engine operates is required, inevitably further entangling Google in the underlying litigation. No assurances, no promises, and no confidentiality order, can protect Google's trade secrets from scrutiny and disclosure during the course of discovery and trial.

Finally, the Government's subpoena imposes an undue burden on Google without a sufficiently countervailing justification. Perhaps the Government can be forgiven its glib rejection of this point because it is unfamiliar with Google's system architecture. If the Government had that familiarity, it would know that its request will take over a week of engineer time to complete. But the burden is not mechanical alone; it includes legal risks as well. A real question exists as to whether the Government must follow the mandatory procedures of the Electronic Communications Privacy Act in seeking Google users' search queries. The privacy of Google users matters, and Google has promised to disclose information to the Government only as required by law. Google should not bear the burden of guessing what the law requires in regard to disclosure of search queries to the Government, or the risk of guessing wrong.

For all of these reasons, the Court must reject the Government's Motion.

II. BACKGROUND

Google was served with the Subpoena on or about August 25, 2005, demanding disclosure of two full months' worth of search queries entered into Google's search engine by Google's users and production of *all* URLs in Google's index that could possibly be found by a search query using Google's search engine at www.google.com. To put this request in context, Google provides the world's most-used search engine at www.google.com. A search engine provides the capability for users to submit text queries to find materials that may be available from an index of World Wide Web addresses or URLs collected by the search engine provider. Declaration of Matt Cutts, ¶ 6 ("Cutts Decl."). Google treats the total number of, and other information about, the search queries it receives as highly confidential. Google treats its methods of searching its index and returning URLs similarly. Cutts Decl., ¶¶ 26-27.

By way of background, the Web is composed of billions of publicly accessible Web sites from around the world and other information sources that Web browsers can access. Each of these sites or other online documents, pages or resources, has an address known as a "URL," short for Uniform Resource Locator. URLs, technically speaking, are comprised of a protocol (e.g., http://) and an Internet Protocol address or domain name. Cutts Decl., ¶ 3. A URL name does not tell the viewer what content may be available on the Web page itself. Cutts Decl., ¶¶ 20-21.

Google timely objected, both in writing and in telephone discussions with the Government's counsel, as the Government acknowledges, to the over-breadth of the request, the seeming irrelevance of the data sought to the claims of the Government, the potential for compromise of Google's trade secrets, and the impact on the privacy of Google search users. Motion, *passim*. While negotiations with the Government narrowed the scope of the Subpoena, the Government apparently could not, or would not, answer Google's concerns. Declaration of Ashok Ramani, ¶¶ 3-6 ("Ramani Decl."). The Government then moved to compel Google's compliance.

III. ARGUMENT

A. The Standard for Compelling a Third Party to Respond to a Subpoena

A court must quash or modify a subpoena issued to a non-party if it subjects that person to undue burden. *Mattel Inc. v. Walking Mountain Prods.*, 353 F.3d 792, 813 (9th Cir. 2003). In analyzing burden, a court must balance the needs of the requesting party with the burden imposed on the non-party. *See Premium Serv. Corp. v. Sperry & Hutchinson Co.*, 511 F.2d 225 (9th Cir. 1975). In performing this balancing test, a court must consider, *inter alia,* the relevance of the requested information to the claim or defenses in the underlying action. *Moon v. SCP Pool Corp.*, 2005 WL 3526513, at *4 (C.D. Cal. Dec. 7, 2005); *see DIRECTV Inc. v. Trone*, 209 F.R.D. 455, 460 (C.D. Cal. 2002).

Unequivocally, the Government bears the burden of establishing relevance. *Echostar Commc'n Corp. v. News Corp. Ltd.*, 180 F.R.D. 391, 394 (D. Colo. 1998). The burden of

establishing relevance is heavier when the disclosure would reveal the protected trade secrets of a non-party. *Id.* In such cases, the Government must show that it has a "substantial need" for the discovery which "cannot be otherwise met without undue hardship." *Id.* (internal quotations and citations omitted).

Even then, the status of a person as a non-party weighs heavily against disclosure. Rule 45 "was intended to provide 'appropriate protection for the intellectual property of [a] non-party witness.'" *Mattel*, 353 F.3d at 814 (citations omitted). "The word 'non-party' serves as a constant reminder of the reasons for the limitations that characterize 'third-party' discovery." *Dart Indus. Co. v. Westwood Chem. Co.*, 649 F.2d 646, 649 (9th Cir. 1980) (internal quotations omitted).

The Government, of course, has told the Court none of this. Instead, it relies on a talismanic incantation that the standard of relevance is met "so long as [the request] is reasonably calculated to lead to the discovery of admissible evidence." Motion, at 5 (citing *U. S. ex rel. Schwartz v. TRW, Inc.*, 211 F.R.D. 388, 392 (C.D. Cal. 2002)). Remarkably, the case the Government cites stands for exactly the opposite proposition. Rather than holding that the relevance standard is met in third party discovery cases with the mere assertion of need, *Schwartz* actually granted the motion of the third party – there, the Government itself – to withhold information on the ground of privilege if the court found the underlying information to be privileged. 211 F.R.D. at 393. The case imposed a heavy burden on the party seeking discovery from the third party to make "a strong showing of necessity" for the information. *Id.* (internal quotations and citations omitted).

Here the Government fails to show that the data it seeks actually will lead to anything admissible in the underlying case because the data simply is not relevant to any claim or defense. Having not crossed that initial threshold, it almost goes unsaid that they have not met the heavier burden they bear to force the disclosure of Google's confidential information. In the end, their purported need pales in comparison to the burden imposed on Google in meeting the request.

B. The Government Fails to Establish That the Data Will Lead to Admissible Evidence

The Government's rationale for its Motion is two-fold:

> (1) the production of [a week of search queries] would permit the Government to evaluate whether COPA or filtering software is more effective in restricting access to harmful-to-minor materials in response to searches as they are actually performed by present day users of the Internet (Motion, at 6); and
>
> (2) the production of [randomly selected URLs] will permit the Government to review a sample set of Internet addresses available to be retrieved from the search engines operated by Google and by other entities. From that set, the Government will be able to review the sample to draw conclusions as to the prevalence of harmful-to-minor material on the portion of the Internet that is retrievable through search engines. Motion, at 8.

This, the Government asserts, is enough to pass Rule 45 muster and to overcome all of Google's objections to production. It is not.

The Government's showing is mere argument, not the required proof of the demanded data's relevance to their claim. What the Government has failed to understand or admit is that the data it seeks from Google has no conceivable use in furthering either of the Government's points. In the absence of a coherent theory of relevance, the Government's demand must be denied. *See Robin Singh Educ. Servs., Inc. v. Excel Test Prep*, 2004 WL 2554454, at *2 (N.D. Cal. Nov. 9, 2004) (denying motion to compel since, *inter alia*, "[o]ther than the naked statement of this argument, Plaintiff does not even attempt to show how either of these two bases for relevance in fact obtain"); *Nicholas J. Murlas Living Trust v. Mobil Oil Corp.*, 1995 WL 124186, at *5 (N.D. Ill. March 20, 1995) (denying motion to compel in part since, *inter alia,* plaintiff failed to articulate a coherent theory which would explain how the requested information was relevant).

We explain why the Government's theory is wrong below, supported throughout with reference to the Declaration of Matt Cutts, a Senior Engineer with Google who has direct knowledge of Google's search engine operation.[1]

1. Search Queries

The Government apparently wants to evaluate the effectiveness of filters by evaluating "searches as they are actually performed by present day users of the Internet" against a database of available URLs. Motion, at 6, 8; Declaration of Phillip Stark, ¶¶ 3-4 ("Stark Decl."). Set aside that such a theoretical comparison could be done without regard to any of Google's data, the query data requested by the Government has no easily computed correlation to how Google would generate a search result based on that same data. This is because when a user enters a search, Google runs a system of proprietary and confidential methodologies and algorithms that allow Google to crawl and index a portion of the Web, and return the most relevant search results to users. Cutts Decl., ¶ 9. These crawling, collecting, and sorting techniques are Google trade secrets. It is therefore impossible for Professor Stark to develop a test or study from the requested Google queries that would reflect realistic search results, without knowing how Google itself would produce a search result based on that query.

Similarly, Google users can and do modify their environment to generate certain types of search results. For example, users may employ Google's SafeSearch pornography filter to limit results. They may use Google's advanced search programs to deliver personalized or customized search results as well. Cutts Decl., ¶ 10. Therefore, the same query will generate different search

---

[1] The court should view the Cutts Declaration as standing in strong contrast to the Government's declarant, Professor Phillip Stark, a statistician who apparently has been hired to produce a study to support the Government's contentions. The Stark Declaration is vague, cursory, and uninformed about the operation of Google's search engine. In any event, Professor Stark's opinion ought to be viewed with some scrutiny. Although positioned as the Government's expert, he has not yet been qualified as a reliable expert by the Pennsylvania court trying the underlying case pursuant to Federal Rule of Evidence 702 or *Daubert v. Merrell Dow Pharms., Inc.*, 509 U.S. 579 (1993). The Pennsylvania court has thus not yet determined whether Professor Stark's testimony is reliable and of any assistance to the trier of fact. *Id.*

results at different computers, depending on the user's preferences, again making it impossible for the Government to develop a test or study that reflects the results from a given search query.

Furthermore, the Government says that its requested search queries will assist it to understand "the search behavior of current web users." Stark Decl., ¶ 4. This statement is so uninformed as to be nonsensical. Search queries run on Google's databases come from such a wide variety of sources that Google's query data, stripped of personally identifying information, will not reveal whether the search query was run by a minor or adult, human or non-human, or on behalf of an individual or business. No conclusion can accurately be drawn from this data about individual behavior. Cutts Decl., ¶¶ 11 – 15.

Indeed, the search query data demanded by the Government would include all "real" queries entered by individual users and automatic queries generated by computer programs called "bots." Identifying and removing all bot and other non-human generated queries will be difficult – if not impossible – for most researchers. Retaining the bot inquiries, which can generate many times the number of searches as an individual, will skew any data set beyond usability and will generate search results that are meaningless if not misleading. Cutts Decl., ¶¶ 12-13.

In addition to bot queries, an individual may run hundreds of queries on Google, not for routine search purposes, but to check the ranking of a website or to deliberately skew Google's query log. Some Google users have actually deliberately sent pornography queries to Google in reaction to the Government's Subpoena. One striking example is that of an individual who wrote a feature for the Firefox (Mozilla) web browser that will send a random pornography query to Google whenever a user enters a query, as if the pornography query had also been entered by the user. Cutts Decl., ¶¶ 14-15, Ex. A.

Finally, Google's system is not static. Algorithms regularly change. The identical search query submitted today may yield a different result than the identical search conducted yesterday. In no meaningful way can it be said that a past week's worth of search queries will yield URL responses as performed by "present day users of the Internet." Motion, at 6. Past searches tell the

Government nothing about URLs available from those searches, now or in the future. Cutts Decl., ¶ 16.

2. URLs

As Mr. Cutts' Declaration elaborates, there is no superficial correlation between (a) the presence of a URL in Google's index and (b) the likelihood of that URL being returned as part of a search result, or being accurately indicative of the Web page to which it links. Google only attempts to crawl the "best of the Web" to create a useful repository of Web pages. Google then implements a structure of complex systems and policies that build on each other for scoring, ranking, returning, or blocking URLs in response to queries. Cutts Decl., ¶¶ 17-19, 22. In short, unless you know *how* Google works, you cannot possibly know *what* Google will return in response to any query. Any assumption to the contrary would be inadmissible speculation.

In addition, the Government will not be able to ascertain the content of a Web page from its descriptive URL name. A Web site's name that suggests potential harmful material may be benign. Conversely, a URL that seems innocent may actually return pornographic material. The classic example is www.whitehouse.com, which was a pornography site. Here, the adage "you can't judge a book by its cover" applies. A URL such as http://www.pbs.org/wgbh/pages/prontline/shows/porn/etc/links.html contains the word "porn" but actually provides links to anti-pornography organizations. Cutts Decl., ¶ 20.

Web page content also changes, or can be changed, every day or more frequently. For example, unscrupulous Web site owners will program their Web page to show innocent content to Google, in order to improve the ranking of their pages, only to swap out that content later to display pornographic material. To ensure the relevancy of its search results, Google puts significant effort into finding and removing those documents, but the process demonstrates the point that the URL itself is not indicative of what content will be displayed in response to a search. Cutts Decl., ¶ 21.

3. The Data Is Not Useful for Any Study

As Google has explained to the Government for months, search queries and randomly-selected URLs from Google's index not only have no correlative value in a vacuum, but any study conducted without understanding Google's system will produce misleading and grossly inaccurate test results, and as a consequence, will not lead to admissible evidence. The Government has made no showing in its Motion of relevance and instead has posited only the *goals* of its proposed study. But if used as the Government suggests, the requested data would so distort Professor Stark's analysis as to render it useless and misleading. This does not merely go to the weight to be accorded the opinion by the trier of fact; it goes to the very relevance of the data requested to support the opinion in the first instance. Accordingly, the Motion should be denied. *See Robin Singh*, 2004 WL 2554454 at *2; *Heidelberg Americas, Inc. v. Tokyo Kikai Seisakusho, Ltd.*, 333 F.3d 38, 41 (1st Cir. 2003) (affirming that subpoena cast too wide a net when "some considerable question exist[ed] as to how discovery of the materials would lead to . . . other discoverable materials"); *Palumbo v. Schulman*, 1998 WL 436367, at *5 (S.D.N.Y. July 27, 1998) (acknowledging unlikelihood that non-parties' documents alone would be relevant to the underlying issues); *In re Surety Ass'n of America*, 388 F.2d 412, 414 (2d Cir. 1967) (affirming trial court's denial of motion to compel where party sought documents regarding types of bonds outside of the one at issue).

C. The Subpoena Demands Google's Valuable Trade Secrets

The Government seeks compelled disclosure of Google's valuable trade secrets and highly confidential commercial information that, if even inadvertently revealed, would seriously damage Google. The burden rests on the Government to show that it has a "substantial need" for the information and that the information cannot be otherwise obtained without "undue hardship." *Compaq Computer Corp. v. Packard Bell Elecs, Inc.*, 163 F.R.D. 329, 338 (N.D. Cal. 1995); *Echostar,* 180 F.R.D. at 394; *Katz v. Batavia Marine & Sporting Supplies, Inc.*, 984 F.2d 422, 425 (Fed. Cir. 1993) (affirming district court's quashing of subpoena that sought confidential

information and did not show a need for the broad range of information requested). If the Government shows such "substantial need" and absence of alternatives – and it cannot – the Court must balance the Government's need with the injury that would result to Google. *DIRECTV*, 209 F.R.D. at 459; *Echostar,* 180 F.R.D. at 394. The balance here plainly favors Google.

1. The Demanded Data Contains Valuable Trade Secrets and Confidential Commercial Information

The Government has not and cannot dispute that Google has devoted enormous amounts of time and expense to protect its valuable trade secrets and confidential commercial information. Google's query and URL data is as secret as any data in the company and must be protected. *DIRECTV*, 209 F.R.D. at 460; s*ee also Katz*, 984 F.2d at 424 (citing *Am. Standard Inc. v. Pfizer, Inc.*, 828 F.2d 734, 740 (Fed. Cir. 1987) (product formulas, product fabrication and marketing plans are trade secrets and should not be subject to discovery)).

The Government acknowledges that Google asserts information about search queries is a trade secret, but says Google identified no reason why it would suffer harm from compelled disclosure. Motion, at 7. But that harm is plain, because a week's worth of query data reflects a sizable number of queries. Taken together (or even in significant groupings), those queries reflect a wealth of information about aspects of Google's business that, if revealed, would injure Google's competitive position. An analysis of Google's query data would reveal proprietary information such as the number of queries that Google can or does process, its capabilities of processing certain lengths and types of queries, its market share in the United States and other countries, and even the demographics of its users. Cutts Decl., ¶ 26. Competition with Google is fierce. Google's competitors could use Google's confidential query data to manipulate their search engines to accommodate Web users and run queries similar to Google's.

Like queries, from even a sample of URLs that Google has indexed, one could estimate, among other things, the size of Google's index; how deeply Google crawls in different countries or languages; and the ability of Google's crawl metrics to measure the reputation of pages or

domains. Information about how Google crawls, or visits the different sub-pages on a website to collect the best URLs, is essential to Google's success. Google has developed its methods and technologies over many years and at considerable expense. Cutts Decl., ¶ 27. If Google's competitors were to access this information, they could conform their size and crawling metrics to Google's, thereby generating search results that mimic Google's and competing more effectively with Google.

Google takes extraordinary measures to protect its trade secrets and confidential commercial information.[2] Both Mr. Cutts and Marty Lev, Google's Director of Safety and Security, offer numerous examples in their Declarations that illustrate the measures of protection ranging from Google's facilities and computer systems to its employees. Cutts Decl., ¶¶ 29-35; Declaration of Marty Lev, *passim*. For example, Google protects its valuable trade secrets at the most basic level by not disclosing the number of computers it maintains to run its search engine, the nature of the search strings typed by users, the type of browsers its users rely upon, the mix of languages that its search engine handles, or how many queries it processes in any given day. Cutts Decl., ¶¶ 24, 26. Access to Google's internal systems, and, in particular, Google's query log and index are each restricted to a small group of trusted employees with special clearance based, in part, on the length of their employment and demonstrated need for access. Cutts Decl., ¶ 32.

The very fact that the Government is so uninformed about the value of search and URL information and so dismissive of Google's interest in protecting it speaks volumes about why the Court should protect Google from this compelled disclosure. The Government's cavalier attitude undermines any credibility in the assertion it later makes that it can or will protect Google against

---

[2] Google routinely receives and refuses requests from researchers and analysts for search query and URL data. For example, Google has denied researchers query logs to protect both its trade secrets and confidential commercial information, and to protect the privacy of its users. Cutts Decl., ¶ 35. Ironically, almost six years ago, Professor Stark obtained a small sample of URLs and queries from a particular Google engineer for what he described as a research project. Ramani Decl., ¶ 7. Unequivocally, it is and has been Google's policy for years not to share any such information with third parties. Ramani Decl., ¶ 7.

loss or further disclosure of the information – a promise that is hollow in the context of litigation in any event.

2. Disclosure of Google's Trade Secrets Is A Significant Possibility

Wide dissemination or outright disclosure of Google trade secrets is inevitable if the Court grants the Government's Motion, because Google necessarily will become entangled in the underlying litigation.

Disclosure to the Government, and Professor Stark, is only half the story. Once the Government tries to support its proposed theory with Google's query and URL data, the ACLU will question the theory's validity and supporting data. The ACLU has issued a subpoena to AOL apparently doing just that in regard to AOL's production in response to a similar Government subpoena. Ramani Decl., ¶ 8. This places Google – an unwilling non-party – in the witness chair, and exposes Google's intellectual property to cross-examination *in open court* by the ACLU, its counsel, experts, and consultants. No protective order can safeguard Google from the eventual eroding of the secrecy of Google's operations and of its competitive advantages as a result of the sharing of its information through or in litigation. At the very least, Google should not have to rely on a Protective Order that was signed by the parties before Google was ever issued the Subpoena.[3] Google neither agreed to nor negotiated the Protective Order and therefore had no control over its terms.

Moreover, Google has no control over whom the parties may identify as expert witnesses or consultants that, according to the Protective Order, will have access to confidential information.

---

[3] Unlike the Protective Order in *Compaq*, which prohibited access by the parties' employees, agents, and even in-house counsel, the Protective Order signed by the Government and the ACLU gives the parties' employees, witnesses, consultants and counsel access to the information and in no way protects Google's proprietary trade secrets and confidential commercial information. *See Compaq*, 163 F.R.D. at 339. Further, the Government's reliance on *Truswal Sys. Corp. v. Hydro-Air Engineering, Inc.*, 813 F.2d 1207, 1211(Fed. Cir. 1987) for the assumption that counsel will not violate the terms of a protective order is misplaced, as Google's concern reaches farther than mere exposure to parties' counsel.

For example, a party may hire a consultant with expertise on search engines who likely was or will be employed by a Google competitor. Even the Government acknowledges in the little it has disclosed about Professor Stark's study that Google's data will be viewed in "accounting for the potential of any variations in the types of queries that are entered into different search engines."[4] Motion, at 6. The Protective Order offers no shield at all against the array of consultants the parties can hire who will rely on and testify as to Google's trade secrets and confidential commercial information. The parties have not even yet designated their experts and consultants; therefore it is impossible for Google to know who will have access to its information.

Disclosure to Professor Stark is a perfect example of the significant threat of harm to Google. Professor Stark's pursuits include consulting in the private sector. Ramani Decl., Ex. A. One example that deeply concerns Google is his involvement with Cogit.com regarding targeted Web advertising. Cogit describes itself as a Web analyst that "provides insight about your customers" and "reveals facts about how they find your site, how they interact with it, and how they leave," available at http://www.cogit.com/. Professor Stark's involvement with Cogit and similarly situated companies may pose a serious threat to the protection of Google's trade secrets and confidential commercial information.

3. The Government Has Not Shown a Substantial Need

The Government asserts that Google's query and URL data "would be of value to the Government in its development of its overall sample of queries" because Google has the largest market share of the Web search Market. Motion, at 6. But this is no showing of "substantial need" at all. Professor Stark does not say the data is "essential," that there is no alternative to it, or that the study he proposes will not stand without it; nor does he explain why a study based on

---

[4] Of course, variations in search queries entered are meaningless. What matters are the URLs returned in response to a given search. As noted above, Professor Stark cannot surmise what URL will be returned without knowing how the particular search engine works. Certainly, any comparison of queries run, to be relevant, would have to use identical terms, would depend on the entity running it (e.g., bot or human), the origin of the query, and other variables that statistically or otherwise render the exercise one of futility. Cutts Decl., ¶¶ 8-16.

samples of randomly selected data must include Google's data to be valid. This is because none of these points is true.

Google's data cannot be essential, because the Government did not demand data from Ask Jeeves, one of the four major U.S. search engines. Professor Stark does not explain the lack of need for that data because he has not in fact disclosed how *any* data will be used in his putative study. Conversely, Google's data cannot be essential because the Government itself has narrowed its request to Google to a smaller sample. If a smaller sample is adequate from Google, and the Government hasn't asked for data from Ask Jeeves, the Government should explain why it doesn't have enough data from the search engines that already have provided millions of search queries and URL data in response to this very Government subpoena. Stark Decl., ¶ 8. There is no showing of necessity because there is no explanation of the study itself or how a sampling of data proves any fact reliably. In the absence of such a showing of necessity, the Motion should be denied. *See DIRECTV*, 209 F.R.D. at 460 (arguments not supported by specific facts or sufficient explanation of why defendants need plaintiff's proprietary information are insufficient to show need).

Nor can there be "substantial need" where, as Google has told the Government, and it acknowledges as much, there are better alternative sources of information. Motion, at 9. The Government concludes without explanation that those sources are "incomplete" and the "most readily available source for those materials are [sic] the operators of search engines themselves." *Id.* On the first point, the Government is simply wrong; on the second point, it chooses its own convenience over the burden it imposes on Google.

Mr. Cutts presents numerous examples of alternative sources for queries and URLs in his Declaration. Cutts Decl., ¶¶ 36 – 38. Google describes a few of these alternatives here. Metasearch engines Dogpile and MetaCrawler each provide services that allow anyone to view lists of queries through their search services, "SearchSpy" and "Metaspy" respectively. SearchSpy, available at: http://www.dogpile.com/info.dogpl/searchspy/, allows one to choose to

view either filtered or unfiltered search queries. With a simple click of a mouse, a user may view search queries as they are run in real time. Metaspy, available at http://www.metacrawler.com/info.metac/searchspy, offers the same feature. Ask Jeeves' "Ask Jeeves Take A Peek" service, available at http://www.ask.com/docs/peek/, lists recently run queries and refreshes automatically twice a minute.

Regarding URLs, the data the Government claims to seek is readily and abundantly available from Alexa.com ("Alexa"). Alexa is specifically intended for "[r]esearchers who wish to tackle problems related to Web content," allowing users to process over four billion URLs and over 1.7 billion full-text documents. A researcher or developer could use this system to test software code, including pornography filters, across much more than the one million URLs sought here, and to test code over the full text of documents. Cutts Decl., ¶ 37, Ex. B.

The query and URL data sought by the Government is available from other sources in form and content more suited to its proposed study, for whatever value there may be in the enterprise. Whatever else that can be said, there is no necessity when similar data is available from other sources. *Compare Compaq*, 163 F.R.D. at 338 (substantial need for a portion of requested information concerning industry practice was established when information could be obtained by no source other than third party industry member) *with Am. Standard,* 828 F.2d at 743 (need not established when information was publicly available); *Travelers Indem. Co. v. Metro Life Ins. Co.*, 228 F.R.D. 111, 114 (D. Conn. 2005) (granting motion to quash subpoena served on non-party since, *inter alia*, requested information was otherwise available).

D. The Subpoena Imposes an Undue Burden on Google

A court must quash or modify a subpoena issued to a non-party if it subjects a person to an undue burden. *Mattel*, 353 F.3d at 813; *see also Echostar*, 180 F.R.D. at 394; *Travelers Indem. Co. v. Metro. Life Ins. Co.*, 228 F.R.D. 111 (D. Conn. 2005). Google's non-party status weighs heavily against the Government in a burden analysis. *See Dart*, 649 F.2d at 649; *Travelers*, 228 F.R.D. at 113 ("courts also give special weight to the burden on non-parties of producing

documents to parties involved in litigation") (citing *Cusumano v. Microsoft Corp.*, 162 F.3d 708, 717 (1st Cir. 1998)).

1. The Time and Resources Required to Pull the Requested Information Would Be Significant

The Government asserts that Google's burden of complying with the Subpoena is "minimal," "not complicated," and "straight-forward." Motion, at 7, 9. That other search engines have complied with the Subpoena and not reported difficulties has nothing to do with Google or its burden. Contrary to the Government's dismissive statements, the Subpoena would require significant time and resources, may hinder Google's basic operations, and may affect Google's performance.

First and most basically: Google does not maintain query or URL data in the ordinary course of business in the format requested by the Government. For this reason alone, the Government's Motion should be denied. *Instituform Technologies, Inc. v. CAT Contracting, Inc.*, 168 F.R.D. 630, 633 (N.D. Ill. 1996) ("Rule 45 does not contemplate that a nonparty will be forced to create documents that do not exist"). Nor is there a program that could simply gather the requested data. Therefore, Google would have to create new code to format and extract the query and URL data from its many computer banks. Pulling each type of data would require multiple teams of Google engineers, diverted from their normal job responsibilities, to research, develop, write, implement, test, fix and execute new computer code. Finally, the selected data must be extracted and copied into a format that can be provided to the Government. In total, the process of gathering the queries and URLs would likely consume up to eight full-time days of engineering time. This time, of course, would have to be covered by other engineers within Google. Cutts Decl., ¶¶ 39-42.

Even if the Government were to pay Google for its engineers' time, executing the searches required by the Government's requests would command extended hours of processing time on Google's computers. Running these programs above and beyond the normal demand on Google's

computers is likely to cause slowdowns, interference and even interruption of Google's normally efficient flow of operations and service, resulting in lower quality of service to users of Google's search engine and to Google's advertisers. Cutts Decl., ¶ 42.

While again we don't know because Professor Stark hasn't said, if it is the Government's intention to use the Google data to then run the same search queries on Google.com, this would put a further enormous and undue burden upon Google. To run the search queries would essentially add a week's worth of searches on Google.com. If the Government were to do this within a short period of time, it would put an enormous strain on Google's computer systems. Cutts Decl., ¶ 23. *See, e.g.*, *Palumbo*, 1998 WL 436367, at *5 (finding that, since non-party data could not be viewed in isolation, it would be unduly burdensome to ask defendants to examine additional information).

2. The Government's Offer to Collaborate Is Inadequate and Unrealistic

Suggesting that creating some multi-stage sample makes it easy on Google, the Government offers that it and Professor Stark are "willing to work with Google to specify a multi-stage sample of the queries" and URLs to reduce "any burden" on Google. Motion, at 7; Stark Decl., ¶ 3. True to form, neither the Government nor Stark proposes a method by which to specify such samples or how to determine randomness. Defining "random" could involve days or months of negotiations on how to determine selection, which could involve months of research and weeks of negotiations on a matter that is currently debated in journals and among authorities. The Government's conclusory statements of the willingness to collaborate with Google are unrealistic and in no way reduce the potential burdens on Google; it further entangles Google in the litigation, further exposes it to pre-trial deposition and cross-examination at trial, and makes it an unwilling witness and participant in the development of an expert's theory. This the law does not require nor should the Court set in motion by passing on the Government's Motion. *Premium Serv.*, 511 F.2d at 229 (plaintiff's offer to reduce burden by sifting through non-parties' documents was unrealistic).

3. The Production of the Requested Data Will Result in a Chilling Effect on Google's Business and User Trust

If Google is forced to compromise its privacy principles and produce to the Government on such a flimsy request, its search query and URL data, Google will, without a doubt, suffer a loss of trust among users. Google's success can be attributed in large part to the high volume of Web users attracted to Google.com every day. The privacy and anonymity of the service are major factors in the attraction of users – that is, users trust Google to do right by their personal information and to provide them with the best search results. If users believe that the text of their search queries into Google's search engine may become public knowledge, it only logically follows that they will be less likely to use the service.

This is no minor fear because search query content can disclose identities and personally identifiable information such as user-initiated searches for their own social security or credit card numbers, or their mistakenly pasted but revealing text. Cutts Decl., ¶¶ 24-25. What will the Government do with this information? While the Protective Order says it should do nothing, at least one Department of Justice spokesperson has said: "I'm assuming that if something raised alarms, we would hand it over to the proper [authorities]." *Technology: Searching for Searches*, Newsweek, Jan. 30, 2006; Ramani Decl., Ex. B.

Because this chilling effect on Google's business is potentially severe, the Motion should be denied.

4. Google Should Not Bear the Burden of Responding to Potentially Inadequate Process Based on ECPA

In addition to the compelling arguments already presented, there remains a substantial question as to whether the Government's request for search queries invokes the mandatory procedures of the Electronic Communications Privacy Act ("ECPA"), 18 U.S.C. §§ 2701-2712. ECPA "creates statutory privacy rights for customers and subscribers of computer network service providers." *Searching and Seizing Computers and Obtaining Electronic Evidence in Criminal Investigations*, United States Dep't of Justice, Computer Crime and Intellectual Property Section

Criminal Division (July 2002), *available at* www.Cybercrime.gov/s&smanual2002.htm. The procedures defined in ECPA for governmental access to stored electronic communications and associated transactional data are not discretionary. If search queries are covered by the statute – and there is good reason to believe they are – the Government must follow the mandatory procedures of either obtaining a court Order or giving notice to every Google user and issuing a subpoena.

The privacy of Google users matters and Google has promised to disclose information to the Government only as required by law or where some imminent harm is threatened. Ramani Decl., Ex. C. Google should not bear the burden and the risk of having to decide whether ECPA applies to this request.

Google provides a service to the public that gives users the ability to send electronic communications in the form of search queries and to receive electronic communications in the form of search results. Google users may initiate recurring searches with results sent to their Google GMail or other email accounts at user – defined intervals. Under ECPA, any service that "provides to users thereof the ability to send or receive wire or electronic communications" is an electronic communication service ("ECS").[5] 18 U.S.C. § 2510(15). An ECS cannot disclose the content of such communications absent strict government compliance with the procedures outlined in Section 2703. Under those procedures, a mere subpoena for this information is not enough.

Further, ECPA places similar restrictions on the disclosure of stored communications to the government by providers of remote computing services. A "'remote computing service' means the provision to the public of computer storage or processing services by means of an electronic

---

[5] Some courts have held that a mere "user" of an ECS provided by another is not itself an ECS. *See Crowley v. Cyberspace Corp.,* 166 F. Supp. 2d 1263, 1270 (N.D. Cal. 2001) (holding that Amazon.com is not an ECS because it did not provide users the ability to communicate); *see also In re Jetblue Airways Corp. Privacy Litigation,* 379 F. Supp. 2d 299, 310 (E.D.N.Y. 2005) (airline reservation website is not an ECS). Google is no mere user of another's ECS of course. It provides the very communications capability at issue here – search – and the capability to receive to receive or direct receipt to oneself or others of search results.

communications system." 18 U.S.C. § 2711(2). An "electronic communications system" is any facility used "for the transmission of wire or electronic communications, and any computer facilities or related electronic equipment for the electronic storage of such communications." 18 U.S.C. § 2510(14).

ECPA applies to a remote computing service if the communication that is held or maintained on that service is –

> (A) on behalf of, and received by means of electronic transmission from (or created by means of computer processing of communications received by means of electronic transmission from), a subscriber or customer of such remote computing service; and
> (B) solely for the purpose of providing storage or computer processing services to such subscriber or customer, if the provider is not authorized to access the contents of any such communications for purposes of providing any services other than storage or computer processing.

18 U.S.C. § 2703(b)(2).

Google users routinely store or establish repeat search queries. Google processes search requests as directed by, and for, its users who in turn retrieve the search results of their choosing from Google's index, or Google sends the results by email or text messages to individuals, to wireless phones or other designated mobile devices. Cutts Decl., ¶ 6. Said in plain language, users rely on the remote computer facilities of Google to process and store their search requests and to retrieve by electronic transmission their search results.

That the Government has asked Google to remove any personally identifiable information from the content of the search queries is of no moment. ECPA makes no exception for anonymous or anonymized content. Surely, the Government does not mean to suggest that it could obtain millions of emails stored in Google's servers simply by asking Google to remove the "To" and "From" lines. It matters not that it might even be helpful or relevant to the Government's case to show that email is used to send content harmful to minors. Content is off limits under ECPA except in rare cases and when procedural safeguards are followed. Google should not bear

the burden of interpreting whether the Government is entitled to the search query results under the form of process it has issued.

IV. CONCLUSION

The Government seeks trade secrets from Google without coming close to proving that these secrets would be relevant in the underlying litigation, that the Government faces a "substantial need" that would not impose an "undue burden" on Google, and that federal law does not blunt the disclosure. The Government's Motion must fail.

DATED: February 17, 2006.

Respectfully submitted,

PERKINS COIE LLP

By: /s/

ALBERT GIDARI, JR.
LISA A. DELEHUNT
180 Townsend Street, 3rd Floor
San Francisco, California 94107-1909
Telephone: (415) 344-7000
Facsimile: (415) 344-7050
Email: AGidari@perkinscoie.com
Email: LDelehunt@perkinscoie.com

Attorneys for Google