PETER D. KEISLER
Assistant Attorney General
THEODORE HIRT
Assistant Branch Director
JOEL McELVAIN, D.C. Bar No. 448431
Trial Attorney
U.S. Department of Justice
Civil Division, Federal Programs Branch
20 Massachusetts Ave., NW
Washington, DC 20001
Telephone: (202) 514-2988
Fax: (202) 616-8202
Email: Joel.L.McElvain@usdoj.gov

Attorneys for Alberto R. Gonzales

IN THE UNITED STATES DISTRICT COURT
FOR THE NORTHERN DISTRICT OF CALIFORNIA
(SAN JOSE DIVISION)

ALBERTO R. GONZALES, in his official capacity as ATTORNEY GENERAL OF THE UNITED STATES,	Case No. 5:06-mc-80006-JW
Movant,	**Reply Memorandum in Support of the Motion to Compel Compliance with Subpoena Duces Tecum**
v.	
GOOGLE INC.,	Hearing: March 13, 2006 Time: 9:00 a.m.
Respondent.	

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page

Cases

Page

Statutes

Miscellaneous

INTRODUCTION

The Child Online Protection Act (COPA), 47 U.S.C. § 231, serves to protect minors from the harmful effects of their exposure to sexually explicit materials on the World Wide Web. A substantial question has arisen, however, as to whether COPA satisfies the requirements of the First Amendment. The United States Supreme Court accordingly has directed the federal district court in the Eastern District of Pennsylvania to make certain factual determinations with respect to that issue, such as the relative efficacy of that statute and of filtering software. To assist that court in its determination, and to defend the constitutionality of a statute that would serve a compelling public purpose, the government is preparing a study that will address the prevalence of harmful sexually explicit material on the Web and the effectiveness of filtering software in screening that material. To develop a data set for that study, the government has sought materials from various sources, including Google Inc. Despite the manifest importance of COPA to families throughout the United States, Google now balks at the government's requests for production, raising three arguments, each of which is meritless.

Before turning to those arguments, it should first be noted what is *not* at issue here. The government has not asked Google to produce any information that would personally identify its users. Instead, it asks only for a sample of URL's available from Google's database, and for the text – and the text only, without any additional identifying information – of a sample of the queries, or search strings, entered onto the Google search engine. The government seeks this information only to perform a study, in the aggregate, of trends in the Internet. No individual user of Google, or of any other search engine, need fear that his or her personal identifying information will be disclosed in response to the subpoena.

Turning to the points that are at issue, none of the three arguments that Google raises suffices to defeat its obligation to provide relevant evidence in response to the subpoena. It first contends that the government's requests are irrelevant to the underlying litigation; as explained below, the materials that the government seeks plainly will have a direct bearing on the Pennsylvania district court's evaluation of the factual questions with which it has been charged.

Google next contends that the subpoena seeks the disclosure of its trade secrets. However, it cannot link the information that the government has requested to any supposed trade secrets, and it is in any event fully protected by the protective order entered by the Pennsylvania district court. Finally, it claims that it is subject to an undue burden in complying with the request, but in fact its own pleadings show that it can comply with the subpoena with relative ease.

ARGUMENT

I. The Requested Samples of URL's and of Search Strings Are Relevant to the Underlying Litigation

Google initially claims that the materials the government has requested are irrelevant to the underlying litigation concerning the constitutionality of COPA. This is a remarkable assertion, given that there is a broad presumption in favor of disclosure, for parties and non-parties alike, in the federal judicial system. "This broad right of discovery is based on the general principle that litigants have a right to every man's evidence, and that wide access to relevant facts serves the integrity and fairness of the judicial process by promoting the search for truth." *Shoen v. Shoen*, 5 F.3d 1289, 1292 (9th Cir. 1993) (internal quotations omitted). A court should be reluctant to accept a non-party's views as to what evidence is relevant to an action, particularly where that action is pending before a different judge in a separate district. "A district court whose only connection with a case is supervision of discovery ancillary to an action in another district should be especially hesitant to pass judgment on what constitutes relevant evidence thereunder. Where relevance is in doubt the court should be permissive." *Compaq Computer Corp. v. Packard Bell Electronics, Inc.*, 163 F.R.D. 329, 335 (N.D. Cal. 1995) (internal quotation and ellipses omitted).

Google's opposition brief amply demonstrates this need for caution. Although it is not a party to the underlying action, it claims to know that its production of URL's or of queries would have "no conceivable use" there. (Opp. at 5.) Google, however, shows nothing more than its own misunderstanding of the underlying action, and of the purposes of the study that the government will undertake to prepare for trial there.

A. The Requested Sample of URL's Is Relevant

As part of its defense of the constitutionality of COPA in *ACLU v. Gonzales*, Civ. Action No. 98-5591 (E.D. Pa.), the government has commissioned a study that will evaluate the effectiveness of content filtering software in protecting minors from sexually explicit material on the Web. (Supplemental Declaration of Philip B. Stark, Ph.D., ¶ 2.) Part of this study will involve the collection of a representative sample of websites, the categorization of those websites that contain such explicit material, and the evaluation of the effectiveness of filtering software in preventing access to that sample of websites. (*Id.*) For this purpose, the government seeks to obtain a random sample of URL's available from Google's index of URL's, for the purpose of evaluating the websites associated with those URL's. (*Id.*) Similarly, the government seeks to obtain a random sample of the text of queries, or search strings, entered on to Google's search engine over a one-week period. (*Id.*) A sample of those queries will be run through the Google search engine for the purposes of drawing the URL's returned by such a search and of evaluating the websites associated with those URL's as well. (Supp'l Stark Decl., ¶ 3.)

This study will permit the government to assist the federal district court in answering a central question that the Supreme Court has asked that court to consider – the relative efficacy of COPA and filtering software in protecting minors from harmful sexually explicit material. For the purpose of this study, the production of a random sample of URL's from Google's database will assist the government in developing a relevant sample of the websites that will serve as the test set for the study. (Supp'l Stark Decl., ¶¶ 5, 18.) The government's request for the production of a sample of URL's from Google is thus plainly reasonably calculated to lead to the discovery of admissible evidence. *See Compaq*, 163 F.R.D. at 339 (discovery from non-party to develop one source of data regarding standard prevailing in non-party's industry, is reasonably calculated to lead to admissible evidence).

Google denies the relevance of its URL's to the government's planned study, asserting that the content of a website cannot necessarily be determined from its descriptive URL name

alone. (Opp. at 8.)[1] This assertion is true, but it is utterly beside the point. As stated above, the government plans to review a random sample of the websites that are associated with the URL's that Google will provide to it, for the purpose of determining, in the aggregate, the prevalence of sexually explicit material on those websites. (Supp'l Stark Decl., ¶¶ 3, 9.) Google further asserts that a sample of URL's from its database would not reveal the likelihood that any particular URL would be returned in response to a search on its search engine. (Opp. at 8.) This, again, misses the point of the government's planned study. The government intends to estimate both the aggregate properties of the websites that search engines have indexed, and the aggregate properties of websites that are returned by search engine queries. (Supp'l Stark Decl., ¶ 8.) The request for URL's is relevant for the former estimate, of course, while the request for queries – discussed in greater detail below – is relevant for the latter estimate. (*Id.*)

Finally with respect to the request for a sample of URL's, Google asserts that the content of websites can be changed, citing to the example of website owners who substitute pornographic materials for innocent material. (Opp. at 8.) This objection also misses the mark. The Internet may be a fluid entity, and the content of the websites associated with particular URL's may be subject to change, but the drawing of a random sample of URL's at a particular moment in time would still produce a relevant sample for the study. (Supp'l Stark Decl., ¶ 18.)

B. The Requested Sample of Queries Is Relevant

In addition to a review of the websites associated with a random sample of URL's from Google's database, the government's study will also use a sample of queries that are entered into the Google search engine, for the purpose of evaluating the URL's that are returned from those

---

[1] None of the arguments that Google now raises regarding the relevance of the discovery requests was presented in its letter stating its objections. (See McElvain Decl., Ex. B.) Rule 45 "'require[s] the recipient of a subpoena to raise all objections at once, rather than in staggered batches, so that discovery does not become a 'game.'" *McCoy v. Southwest Airlines Co.*, 211 F.R.D. 381, 385 (C.D. Cal. 2002) (quoting *In re DG Acquisition Corp.*, 151 F.3d 75, 81 (2d Cir. 1998)).

queries. (Supp'l Stark Decl., ¶ 3.) The production of a sample of queries from Google would assist the government in this study.

Google again claims to know that a sample of the queries from its search engine could not be relevant to the action pending before the Pennsylvania district court. Its arguments with respect to the relevance of the queries are as mistaken as are its arguments with respect to the URL's. First, it claims that the text of any particular search will have no necessary correlation to the results of that search, and that the government could not know what the results of that search would be without also knowing the details of Google's proprietary algorithms. (Opp. at 6.) This is simply untrue. Millions of people use Google's search engine each day, and they learn the results of their searches simply by viewing Google's response; they need have no particular knowledge of Google's methodology in order to make sense of the results that are returned to them. (Supp'l Stark Decl., ¶ 7.) Similarly, for the purpose of the study that the government has commissioned, all that is needed is the drawing of a random sample of URL's from Google's index, a random sample of queries from its search engine, and the results from running those queries through that search engine. (*Id.*)

Google next asserts that some of its users customize the parameters of their searches. (Opp. at 6-7.) This claim is irrelevant for the purposes of the government's study; the running of a random sample of queries through Google's search engine, set to its default parameter settings, will allow for an estimate of the amount of sexually explicit materials that are available for the user of a search engine to encounter, and will provide a sample of a relevant population of websites that can be categorized and used to test filtering software. (Supp'l Stark Decl., ¶ 18.)

Third, Google asserts that the text of queries on its search engine will not reveal information regarding the source of those queries, such as whether the query was entered by a minor or an adult, or whether it was entered directly by an individual or by a computer program on the individual's behalf. (Opp. at 7.) (In the same pleading, however, Google also asserts that a sample of its search terms would reveal proprietary demographic information about its users. Opp. at 10. As a matter of simple logic, these claims cannot both be true. Supp'l Stark Decl.,

¶ 6.) This, again, misses the point of the government's planned study. The government seeks to study the indexed Web and the results of searches at an aggregate level, and to measure how well filtering software works to block sexually explicit materials on a sample drawn from a relevant population; for that purpose, it is not essential to determine the source of any particular query that was run on the Google search engine. (Supp'l Stark Decl., ¶ 18.)

Google further claims that its algorithms for its search engine will change over time. (Opp. at 7-8.) This, too, is simply beside the point. The running of a random sample of Google queries through the Google search engine will still allow for an estimate of the fraction of queries that can return results with sexually explicit materials, and also will still allow for the testing of filtering software against that relevant sample. (Supp'l Stark Decl., ¶ 18.) In other words, the mere fact that the Internet changes over time, or that a search engine's algorithms change over time, does not mean that it is impossible to draw relevant conclusions as to the nature of the internet from data gathered at any one particular time.[2]

The government thus has plainly shown that its requests for a sample of URL's and for a sample of queries are reasonably calculated to lead to the discovery of admissible evidence.

---

[2] The plaintiffs in the underlying litigation have filed a brief stating their views of the merits of that case, and also questioning the relevance of the subpoena to Google. They, however, have no standing to raise such arguments here; the subpoena does not implicate any particular interests of theirs, such as a claim of privilege. *See United States v. Tomison*, 969 F. Supp. 587, 596 (E.D. Cal. 1997). If the plaintiffs believed that the subpoena harmed them in some way relating to the underlying litigation, their remedy would not be to appear here, but instead to object before the Pennsylvania district court. *See Thomas v. Marina Assocs.*, 202 F.R.D. 433, 434 (E.D. Pa. 2001). They did no such thing; instead, at a recent case management conference, that court directly asked the plaintiffs to state whether they objected to the Google subpoena, and their counsel responded that he had no objection. (2d McElvain Decl., Ex. A, at 11-12.) On the basis of that representation, and on the basis of the government's explanation of the purpose of the subpoena, the court stated its understanding that the Google subpoena was relevant for the purposes of discovery in the action before it (while, of course, reserving judgment as to later questions of admissibility at trial). (*Id.* at 15.) The plaintiffs should not be allowed to contradict their representation to the Pennsylvania district court here.

II. Google's Trade Secrets Would Not Be Revealed by the Production of a Sample of URL's or of a Sample of Search Strings

Google fares no better with its next argument, as its compliance with the subpoena would not subject it to any meaningful risk of harm from the disclosure of its trade secrets. Google bears the burden to demonstrate, first, that the subpoena seeks the production of its trade secrets, and that the disclosure of its trade secrets, under the terms requested in the subpoena, would be harmful to it. *See, e.g., Trevino v. ACB American, Inc.*, 232 F.R.D. 612, 617 (N.D. Cal. 2006). If it fulfills that burden, the government would then be required to show that the information sought is "relevant and necessary." *See, e.g., Coca-Cola Bottling Co. v. Coca-Cola Co.*, 107 F.R.D. 288, 293 (D. Del. 1985). Contrary to Google's claims (Opp. at 13), the government need not demonstrate that the evidence sought would be outcome-determinative, or even that it would be admissible at trial. Instead, the government need only show that it seeks materials that are "reasonably necessary for a fair opportunity to develop and prepare the case for trial." *American Standard, Inc. v. Pfizer, Inc.*, 828 F.2d 734, 743 (Fed. Cir. 1987). For example, this court has held that a party may seek discovery – even of trade secrets, and even from a non-party – to support an expert report analyzing the prevailing standard in the non-party's industry, even though the material sought would only "be illustrative, albeit not dispositive" of the ultimate issues at trial. *Compaq Computer Corp.*, 163 F.R.D. at 339 n.25.

If this showing is made, the court then weighs the party's need for the requested materials against the possible harm from the disclosure. *See, e.g., Trevino*, 232 F.R.D. at 617. However, "*[b]alancing* is perhaps the wrong word to describe [this] task." *Mycogen Plant Sci. v. Monsanto Co.*, 164 F.R.D. 623, 626 (E.D. Pa. 1996) (emphasis in original). This is because "[o]rders forbidding any disclosure of trade secrets or confidential commercial information are rare. More commonly, the trial court will enter a protective order restricting disclosure to counsel or the parties." *Federal Open Market Comm. v. Merrill*, 443 U.S. 340, 362 n. 24 (1979) (internal citations omitted). Thus, "a survey of the relevant case law reveals that discovery is

virtually always ordered once the movant has established that the secret information is both relevant and necessary." *Compaq*, 163 F.R.D. at 338 (internal quotation omitted).

Google cannot demonstrate that the subpoena implicates any of its trade secrets or that it faces any appreciable risk of harm from their disclosure. In any event, the material that the government seeks is reasonably necessary to its defense of the constitutionality of COPA. The balance thus weighs heavily in favor of disclosure, subject to the appropriate protections that Google will enjoy under the protective order entered by the district court.

A. Google Does Not Face a Risk of Disclosure of Its Trade Secrets

The subpoena does not seek the production of any of Google's confidential business information, but instead only two forms of data, URL's from its database, and the text of queries from the query logs for its search engine. Google asserts that its production in response to the subpoena would reveal its trade secrets, but fails to explain convincingly why this would occur. (The purported trade secrets that Google now claims are implicated by the subpoena, it should be noted, are entirely different from the alleged secrets that it had claimed were at stake in its objections letter; it consequently has waived these new claims. *See McCoy*, 211 F.R.D. at 385.)

Google asserts that a sample of URL's from its index would reveal the size of its index, the ability of its crawl metrics to measure the reputation of webpages, or the depth of its crawling capabilities. (Opp. at 10-11.) None of these claims is plausible. As a matter of basic statistical principles, the drawing of a random sample of a particular size from a population reveals nothing about the size of that population other than that it is at least as large as the sample. (Supp'l Stark Decl., ¶ 11.) It is, of course, common knowledge that Google's index is much larger than the sample that the government has asked to be drawn. (*Id.*) Nor could any conclusions be drawn regarding Google's methodology for measuring the reputation of webpages merely from the drawing of a random sample of URL's. (*Id.*)

With respect to the depth of Google's crawling, it is unclear whether Google means to refer to the use of "click depth" or "folder depth"; with respect to the former method, no conclusions could be drawn from a random sample as to the click depth of Google's crawling

without also knowing what other pages in Google's index link to the URL's in that sample. (*Id.*) With respect to the latter method, while it is conceivable that some conclusions as to the folder depth of Google's crawling could be drawn from the sample by reviewing the URL titles in the sample, the same conclusions could be drawn simply by restricting a Google search to a specific domain and reviewing the results of that search. (*Id.*) Despite the fact that Google did not see fit to raise this objection previously, the government is willing to accommodate it by accepting a smaller random sample of Google's URL's. The uncertainty in any estimates of folder depth from that smaller sample would be large enough to avoid any concerns that Google might have as to this allegedly proprietary information. (*Id.*)

Google fares no better in its claim that a sample of queries would somehow reveal its trade secrets. As an initial matter – as will be explained in greater detail below – the text of queries on Google's search engine are routinely revealed to third parties every day. Such public disclosure, of course, defeats any claim that these materials are trade secrets. *See Religious Tech. Ctr. v. Netcom On-Line Commc'n Serv.*, 923 F. Supp. 1231, 1254 (N.D. Cal. 1995). In any event, it is plain that the request for production of queries does not implicate the alleged secrets that Google claims are at stake here. It asserts that the production of its queries would reveal its capabilities of processing those queries, such as its ability to process certain lengths or types of queries. (Opp. at 10.) This is a non sequitur. The mere fact that a query was entered in Google's search engine would not reveal that Google processed the query in any particular way. (Supp'l Stark Decl., ¶ 10.) Also, Google's claim that a sample of its queries would reveal the demographics of its users is inconsistent with other claims that Google has made here, and is in any event implausible; no reliable demographic data would be revealed from such a sample. (*Id.*)[3]

[3] Nonetheless, the government is willing to accommodate Google's concerns by specifying a smaller random sample to be drawn from a population consisting of one week's worth of queries entered onto its search engine. (Supp'l Stark Decl., ¶ 3.)

Google also argues that it production of queries would allow for an estimate of its market share in the United States or other countries. (Opp. at 10.) Google apparently bases this claim on a belief that an analysis of searches run in particular languages would reveal this information. But it would be tenuous to infer, from the language of a query, the country from which a search has been run. (Supp'l Stark Decl., ¶ 10.) In any event, even if this inference were appropriate, any estimate of Google's market share would require analysis of the proportion of searches in particular languages in samples not only from Google but from multiple search engines. The government, however, will not review the queries (apart from quality control checks) for the language of those queries or for any other aspect of their content. (Supp'l Stark Decl., ¶ 12.) Thus, there is no prospect that such an estimate could be drawn from the query samples. In any event, detailed information regarding Google's market share in particular languages or in particular countries is already publicly available. (Supp'l Stark Decl., ¶ 10.)

B. Google Is Fully Protected by the Protective Order Already in Place

It is thus apparent that Google faces no risk of the disclosure of any trade secrets from its compliance with the subpoena. Moreover, even if Google could demonstrate that any of its trade secrets would be implicated by the subpoena, it still faces no appreciable risk of the disclosure of those secrets. A comprehensive protective order has already been entered by the district court in Pennsylvania, and that order fully protects Google's interests. (McElvain Decl., Ex. D ("Protective Order").)[4] Google asserts that the protective order is insufficient, because nothing would prevent the disclosure of its alleged trade secrets at trial "*in open court.*" (Opp. at 12; emphasis in original.) Nothing prevents such a disclosure, that is, except a closer reading of that order; Google has overlooked the fact that the order also affords it protections from the disclosure of confidential materials at trial. (Protective Order, ¶ 10.)

[4] Google need only designate its production as "confidential" in order to benefit from the terms of the protective order; thus, despite the real doubt that any legitimate trade secrets are implicated here, Google will nonetheless be protected from disclosure. (Protective Order, ¶ 3.)

Google also contends that the parties' witnesses or consultants might also be employed by one or more of its competitors, and that those persons could violate the terms of the protective order by sharing its confidential information with those competitors. (Opp. at 12-13.) In support of this claim, Google impugns the integrity of the government's declarant, Dr. Philip Stark, asserting that it is "deeply concern[ed]" by his work for what it purports to be its competitor, Cogit.com. (Opp. at 13.) Contrary to the aspersions that Google has cast, Dr. Stark has signed a declaration attesting that he will obey the protective order entered by the Pennsylvania district court. (Supp'l Stark Decl., Ex. A.) He has frequently signed similar confidentiality agreements, and he has abided by each of those agreements. (Supp'l Stark Decl., ¶ 14.) In any event, Cogit.com has been out of business for several years, and the web site that Google quotes in its brief is run not by Cogit, but by a different entity. (*Id.*)

C. The Materials Sought Are Reasonably Necessary For the Preparation of the Government's Defense

Google cannot demonstrate that any of its trade secrets would be implicated by its production in response to the subpoena, or that any such secrets – given the protective order already in place – would be at risk of disclosure. Consequently, its effort to avoid compliance with the subpoena on this ground must fail. But even if Google had made such a showing, the government would still be entitled to disclosure, as it has shown that the materials it seeks are reasonably necessary for the preparation of its defense of the constitutionality of COPA.

As the government has explained, the production of these materials from Google would be of significant assistance to it for the purposes of the study that it has commissioned. Google is the largest search engine, and some parties estimate that it is one of the largest gateways to access pornography on the internet. (Supp'l Stark Decl., ¶ 4.) The data from the Google index, and from its search terms, thus plainly provides a relevant population that may serve as a basis for the government's planned study. (Supp'l Stark Decl., ¶¶ 4, 5.) *See Compaq*, 163 F.R.D. at 339 n. 25 (movant showed need for production of trade secrets from non-party by demonstrating that information would be relevant as part of larger study).

Google asserts that the government has already received "millions" of queries and URL data from other search engines, and thus could not possibly need more. (Opp. at 14.) But, in the field of statistics, the volume of data is not itself meaningful; instead, data must be drawn from a relevant population, and Google is of course a relevant population for the purpose of evaluating the character of the internet. Google also asserts that there must be no need for its data, as the government instead could have subpoenaed the search engine Ask Jeeves. (Opp. at 14.) But Google has vastly more search traffic than does Ask Jeeves, and so it plainly is an appropriate source of relevant data for the purpose of the statistical study. (Supp'l Stark Decl., ¶ 5.)[5]

D. The Balance Weighs in Favor of Disclosure

In light of the fact that Google cannot demonstrate that it suffers any real danger of the disclosure of its trade secrets, and the fact that the government has a legitimate need for the disclosure of data that is uniquely in Google's possession, the balance certainly weighs in favor of disclosure of any alleged trade secrets. Given the strong public interest in allowing the Pennsylvania district court to have a full and fair opportunity to assess the factual questions that the Supreme Court has charged to it on remand, it is particularly appropriate here to weight the balance in favor of disclosure of Google's alleged trade secrets. As discussed above, the balance almost always weighs in favor of such disclosure, *see Merrill*, 443 U.S. at 362 n. 24, and Google can demonstrate no reason to depart from that general principle here.

In any event, even if Google had truly made a showing that its trade secrets were threatened, it would be required further to show that its interests could not be protected by modifying the subpoena, as opposed to quashing the subpoena in its entirety. *See Northrop Corp. v. McDonnell Douglas Corp.*, 751 F.2d 395, 404 (D.C. Cir. 1984) (court must first consider modification of subpoena before quashing). "The quashing of a subpoena is an

[5] Google's additional claim that relevant data could be obtained from other sources such as Alexa.com fails for a similar reason. (Opp. at 14-15.) Whatever the relevance that those other sources might have, they do not provide information regarding Google's index, or regarding Google's queries. Google's data is a relevant population for the purposes of the government's study. (Supp'l Stark Decl., ¶ 5.) Of course, Google data can only be obtained from Google.

extraordinary measure, and is usually inappropriate absent extraordinary circumstances. A court should be loathe to quash a subpoena if other protection of less absolute character is possible." *Flanagan v. Wyndham Int'l, Inc.*, 231 F.R.D. 98, 102 (D.D.C. 2005). Google has made no such showing, and its effort to avoid compliance with the subpoena in any form should be rejected.

III. Google Would Not Face an Undue Burden in Complying with the Subpoena

Finally, Google asserts that it should not be subjected to what it claims to be the "significant" burden of complying with the subpoena, for three reasons. (Opp. at 16.) It claims that it should not be required to devote its time or its computing resources to complying with the subpoena. It next claims that it would face a loss of trust among the users of its search engine if it were to comply with the subpoena. It also asserts that its compliance with the subpoena might violate the Electronic Communications Privacy Act (ECPA), 18 U.S.C. §§ 2701-2712. None of these claims suffices to carry Google's burden to prove that it cannot be reasonably expected to comply with the subpoena. *See Green v. Baca*, 226 F.R.D. 624, 653 (C.D. Cal. 2005).

A. Google Would Not Be Required to Expend Significant Resources to Comply with the Subpoena

Google asserts that it would be required to devote between a minimum of three and a maximum of eight days of the time of its engineers to comply with the subpoena. (Opp. at 16.) It is likely that it would take Google substantially less time than it estimates to comply with the subpoena, for several reasons. First, there is no dispute that Google already maintains query logs and an index of its URL's. (*See* Cutts Decl., ¶¶ 29, 32-33.)[6] Second, Google has been able to produce samples of its index in the past. (Supp'l Stark Decl., ¶ 15.) Further, Google regularly generates reports from its query logs and publishes those reports, for example, on its Google Zeitgeist page; the generation of a random sample from the query logs would almost certainly

[6] Google claims that its data is not maintained in the ASCII file format for which the government has requested production, and thus that it cannot be compelled to produce electronic data in that format. (Opp. at 16.) This argument is mistaken. Federal Rule of Civil Procedure 34(a) expressly provides for the discovery of "data compilations from which information can be obtained, translated, if necessary, by the respondent through detection devices into reasonably usable form." *See also PHE, Inc. v. Dep't of Justice*, 139 F.R.D. 249, 257 (D.D.C. 1991).

require only minor modifications from the tools that Google uses to produce those reports. (*Id.*) Finally, other search engine providers have been able to produce URL's and queries to the government in this litigation without complaining of undue burden. (*Id.*) It is likely that Google has technical capabilities that are at least comparable to those of its competitors. In any event, as the government has previously explained, any variations between the structure of Google's databases and those of its competitors could be accounted for by the specification of a multi-stage sample, which would greatly diminish any potential burden that Google would face.

Google asserts that it cannot be expected to engage in the "months of research" that would be required before it would negotiate with the government as to the definition of a random sample. (Opp. at 17.) This contention is preposterous. The principles governing the definition of a random sample, for the purposes of defining the samples to be drawn from Google, are well established in the scientific field of statistics. (Supp'l Stark Decl., ¶ 16.) Moreover, there is no need for the government and Google to negotiate the point; the government can simply specify the methodology that is needed and permit Google to draw the sample consistent with that specification. (*Id.*) It would then fall to the government, not to Google, to defend its methodology should any challenge arise later in the Pennsylvania district court.

In any event, even if Google is accurate in its estimate that it would need to devote as many as 64 of its employees' hours to comply with the subpoena, and even if that estimate could not be limited by the use of a multi-stage sample, this plainly does not approach the showing that would be required to justify quashing the subpoena. The government is willing to compensate Google for its reasonable expenses in complying with the subpoena. Given this fact, Google, whose most recent report to the Securities and Exchange Commission reflects that its assets are worth over $9 billion (2d McElvain Decl., Ex. B), can hardly claim that the cost of the subpoena would be unduly burdensome. *See Compaq*, 163 F.R.D. at 339 (subpoena requiring non-party to devote 1,000 of its employees' hours to compliance was not unduly burdensome, although compensation would be ordered); *see also McCoy v. Whirlpool Corp.*, 214 F.R.D. 642, 645 (D.

Kan. 2003) (discovery requiring 160 hours and $10,400 in cost was not unduly burdensome to corporation with net income comparable to Google).

Google further argues that its execution of a program to draw samples could cause interference with the operations of its search engine. (Opp. at 17.) It fails to specify, however, the extent to which it believes those operations would suffer. *See, e.g., Diamond State Ins. Co. v. Rebel Oil Co.*, 157 F.R.D. 691, 696 (D. Nev. 1994) (alleged burden must be identified with specificity to justify quashing subpoena). This oversight is telling. Google's database processes hundreds of millions of search terms every day. (Supp'l Stark Decl., ¶ 17.) Given the size of Google's computing capabilities, any impact of the program needed to draw the samples would almost certainly be vanishingly small.[7] Thus, because Google cannot show that it would face any uncompensated burden, let alone a burden that would be "undue," in complying with the subpoena, its objections should be rejected.

B. Google Would Not Risk the Loss of Its Users' Confidence if It Were to Comply with the Subpoena

Google argues that it should not be "forced to compromise its privacy principles" by complying with the subpoena. (Opp. at 18.) This statement is highly misleading, as Google has asserted no "privacy principles" that would prevent it from disclosing search terms to the government, or to any other party. (Google, to its credit, does not raise the even more tenuous claim that its production of URL's would raise privacy concerns.)

It bears repeating here that the government has not asked Google to produce *any* information that could identify the users of its search engines, or the computers from which any search terms have been entered. Instead, the government has asked for the production *only* of the actual text of a sample of queries entered on to the Google search engine, without any additional information identifying the source of that text. Of course, without this additional

---

[7] Google also asserts that its computing capabilities would suffer if the government were to run its search queries back through its search engine. This misapprehends the government's planned study. A smaller sample of approximately 1,000 queries, and not the entire set, will be run through the Google search engine. (Supp'l Stark Decl., ¶ 3.)

information, the text of a query entered on a search engine reveals nothing about the author of that text. (Supp'l Stark Decl., ¶ 12.) There is therefore, simply, no basis for any fear that the subpoena seeks the disclosure of the identity of any particular user of Google's search engine.

Given Google's prominent declarations in its brief as to its purported commitment to the confidentiality of the queries on its search engine, one might have expected it to refer to the policy statement that it has published describing its treatment of its users' personal information. There is no such reference, however, and the omission is telling. While Google does make certain representations in its privacy policy as to the circumstances in which it will disclose "personal information," it makes no such representations as to any other kind of data, including "aggregate non-personal information." (2d McElvain Decl., Ex. C.) For this purpose, it defines "personal information" to mean "information that you provide to us which personally identifies you, such as your name, email address or billing information, or other data which can be reasonably linked to such information by Google." (*Id.*)

Google plainly does not consider the content of search terms to be "personal information" for the purpose of this privacy policy. To the contrary, queries that are entered into Google's search engine are *routinely* revealed to other websites, and Google makes no efforts to prevent this. This disclosure occurs as follows. First, when a user runs a Google search, Google returns a page of results with an address that includes the entered search terms (*e.g.*, http://www.google.com/search?q=my+search+terms). The user may then click on a link as displayed on this results page. When the user does so, under the specification for the Hypertext Transfer Protocol (RFC 2616), his or her web browser will pass several pieces of information to the new website that he or she is visiting; one of these fields, known as the "referer," or HTTP_REFERER, specifies the address of the previous web page that directed the user to the current website.[8] As a result, when a user clicks on a link in a Google search results page, the

[8] *See* http://www.w3.org/Protocols/rfc2616/rfc2616-sec14.html#sec14.36 *See also* http://www.alistapart.com/articles/searchhighlight

address of that page – including the search terms embedded in the address (*e.g.*, my+search+terms) – is disclosed to the operator of the linked-to website. And this occurs despite the fact that the operator of that website will also receive information regarding the user's IP address, which may be associated with the search terms. The government's request for production here, in contrast, seeks no such identifying information.

Google itself does not transmit this search information to other websites; instead, the individual user's browser does so, in accordance with the official HTTP specification. However, Google could easily prevent users' search terms from leaking out in this fashion, but it chooses not to do so, and thus tacitly allow user search queries to be disclosed to websites visited by Google search users.[9] Moreover, Google affirmatively encourages its advertisers to use referrer logging to track the traffic on their websites. (Supp'l Stark Decl., ¶ 13.) This is, of course, inconsistent with Google's present assertion as to the value it places on the confidentiality of the text of queries on its search engine.

C. The Subpoena Does Not Violate the Electronic Communications Privacy Act

Google finally and half-heartedly suggests that there is a "substantial question" as to whether the government's request for a sample of queries complies with the Electronic Communications Privacy Act ("ECPA"), 18 U.S.C. §§ 2701-2712. (Opp. at 18.)[10] It cites no case law or legislative history whatsoever in support of this theory. The reason for Google's lack of legal support is clear; there is none. Section 2703 of ECPA regulates government access to electronic communications stored by two defined types of network service providers: "electronic communication services," *see* 18 U.S.C. § 2510(15), and remote computing services,

---

[9] Google could construct its search input form to use the HTTP POST method instead of the GET method. *See generally* Shishir Gundavaram, *CGI Programming on the World Wide Web*, ¶ 4.2 (O'Reilly 1996) (available at http://www.oreilly.com/openbook/cgi/ch04_02.html).

[10] This argument, as well, is entirely new, and hence Google has waived it.

see 18 U.S.C. § 2711(2). Google's search engine does not fall within either of these categories, and therefore it is not subject to the statute.

ECPA defines "electronic communication service" as a service that "provides to users thereof the ability to send or receive wire or electronic communications." 18 U.S.C. §§ 2510(15), 2711(1). When Congress enacted ECPA, it identified telephone companies and email providers as providers of electronic communication service. *See* S. Rep. No. 99-541 (1986) at 14, *reprinted in* 1986 U.S.C.C.A.N. 3555, 3568. With the growth of the Internet, the definition of electronic communication service has come to include services offered by Internet service providers (ISP's). *See, e.g., Freedman v. America Online, Inc.*, 325 F. Supp. 2d 638, 643 & n.4 (E.D. Va. 2004). Like telephone companies and email providers, ISP's enable users to communicate with others.

A party that merely maintains a website or utilizes Internet access does not provide an electronic communication service under ECPA. Websites are users of communication services, rather than providers. For example, in *Crowley v. Cybersource Corp.*, 166 F. Supp. 2d 1263, 1270 (N.D. Cal. 2001), this court held that Amazon.com was not a provider of electronic communication service, despite the fact that Amazon received e-mails from its customers; the mere fact that Amazon's website allowed for communication over the Internet did not transform Amazon into a provider. *See id.*; *see also State Wide Photocopy v. Tokai Fin. Servs.*, 909 F. Supp. 137, 145 (S.D.N.Y. 1995) (company's use of a computer or fax machine did not make that business an electronic communication service provider under ECPA). Similarly, a party – such as an airline – that merely provides products or services over the internet, without providing access to the Internet itself, is not a provider of an electronic communications service. *See, e.g., In re JetBlue Airways Corp. Privacy Litig.*, 379 F. Supp. 2d 299, 307 (E.D.N.Y. 2005) (airline "does not become an 'electronic communication service' provider simply because it maintains a website that allows for the transmission of electronic communications between itself and its customers"); *Dyer v. Northwest Airlines Corp.*, 334 F. Supp. 2d 1196, 1999 (D.N.D. 2004).

The reasoning of *Crowley* and the airline cases applies with equal force to Google's search engine. Google operates its search engine through a website; it does not provide Internet access or otherwise enable direct communications between Google users. Google, like Amazon.com and the airlines, is a user rather than a provider of electronic communication service. Google, through its search engine, process a query and then sends the results back to the user over the Internet; it thus acts as a party to communications with the user, rather than providing users with a channel of communication.

Google seeks to distinguish the solid line of case law by characterizing its search engine as a "communications capability." (Opp. at 19 n. 5.) This is simply wrong; the definition of an electronic communication service – a service that provides users "the ability to send or receive wire or electronic communications" – does not encompass conducting searches. 18 U.S.C. § 2510(15). Indeed, Amazon.com's website allows users to search its products, and airline website users can query the availability and cost of flights, but such search capabilities do not convert Amazon.com or the airlines into providers of electronic communication service.

Google also notes that its users may initiate recurring searches and have the results sent to specified email accounts. (Opp. at 19.) This capability does not transform Google into a provider of electronic communication service. In emailing users the results of periodic searches, Google merely becomes a periodic user of electronic communication service. Similarly, even if Google sends query results to a party other than the user initially making the query, it still is acting as a user rather than a provider of electronic communication service. In such cases, Google is not acting as a conduit to transmit a message from one user to another. Instead, the results of the query are created by Google, and Google then uses the Internet to transmit the results to the specified party.

Nor is Google is a provider of "remote computing service" under ECPA. By definition, a remote computing service provides "to the public . . . computer storage or processing services by means of an electronic communications system." 18 U.S.C. § 2711(2). Essentially, a remote computing service is a service which handles outsourced computer storage or processing. *See* S.

Rep. No. 99-541 at 10-11 (1986), *reprinted in* 1986 U.S.C.C.A.N. at 3564-65 (noting that firms face a choice over "whether to process data in house on the user's own computer or on someone else's equipment," and that "businesses of all sizes . . . use remote computer services for computer processing"). For example, a service provider that processes data in a time-sharing arrangement provides a remote computing service. *See id.* The mere operation of a website, however, is not a "remote computing service" for this purpose. *JetBlue*, 379 F. Supp. 2d at 310.

Similarly, Google's website search engine is not a remote computing service under ECPA. In its basic form, Google's search engine does not provide basic computer storage, and it does not handle outsourced computer processing. Although Google uses its computer resources to respond to search queries, the same can be said for Amazon.com's website or an airline web site. In fact, *every* website uses computers in response to communications from users. At a minimum, a website must receive data from users, analyze the data and formulate the appropriate response (for example, retrieving a stored web page requested by the user), and transmit the response back to the user. Websites thus are not remote computing services under ECPA, as their function is not to perform outsourced computer processing.

This result – that Google is not a remote computing service – is not changed by the fact that Google may "store or establish repeat search queries" on behalf of some users. (Opp. at 20.) "Storage" within the meaning of the definition of "remote computing service" must be storage for general archival purposes, not merely storing some information to fulfill some customer request. Otherwise, nearly every business on the Internet would become a remote computing service, as they store basic information and preferences regarding their customers. This result was rejected in *Crowley*, *JetBlue*, and the other airline cases, and it should be rejected here.

Finally, even if Google were a remote computing service under ECPA, the government's subpoena would not violate ECPA. Section 2703(b) of ECPA governs compelled disclosure of the contents of communications held by a provider of remote computing service. However, the disclosure restrictions of § 2703(b) only apply to the contents of certain communications. In particular, the restrictions of § 2703(b) do not apply unless the communications are stored

"solely for the purpose of providing storage or computer processing services to such subscriber or customer." 18 U.S.C. § 2703(b)(2)(B). Here, the Government seeks all queries entered into Google's search engine during a specified time frame and a sample of URLs selected from Google's index. Assuming arguendo that this information is the contents of communications, this information is not stored on behalf on any particular subscriber or customer. Instead, Google stores such information for its own purposes. Thus, regardless of whether Google is a remote computing service, the Government's subpoena here does not violate ECPA.

CONCLUSION

For the foregoing reasons, the Movant, Alberto R. Gonzales, in his official capacity as Attorney General of the United States, respectfully requests that this motion be granted and that the Respondent, Google Inc., be compelled to comply with the subpoena issued to it. In order to accommodate the case management schedule in the pending case of *ACLU v. Gonzales*, Civ. Action No. 98-5591 (E.D. Pa.), the Movant further respectfully requests that the Court expedite its consideration of this motion, and that it order Google Inc. to comply with the subpoena within 21 days of this Court's order.

Dated: February 24, 2006

Respectfully submitted,

PETER D. KEISLER
Assistant Attorney General
THEODORE HIRT
Assistant Branch Director

/s/
JOEL McELVAIN
Trial Attorney
United States Department of Justice
Civil Division, Federal Programs Branch
20 Massachusetts Ave., NW, Room 7130
Washington, D.C. 20001
Telephone: (202) 514-2988
Fax: (202) 616-8202
Email: Joel.L.McElvain@usdoj.gov

Attorneys for the Movant, Alberto R. Gonzales

CERTIFICATE OF SERVICE

I hereby certify that I have made service of the foregoing Reply Memorandum in Support of the Motion to Compel Compliance with Subpoena, and of a Proposed Order, by depositing in Federal Express at Washington, D.C., on February 24, 2006, true, exact copies thereof, enclosed in an envelope with postage thereon prepaid, addressed to:

Albert Gidari, Jr., Esquire
Perkins Coie, LLP
1201 Third Avenue
Seattle, Washington 98101-3099
(Counsel for Respondent Google Inc.)

Lisa Delehunt, Esquire
Perkins Coie, LLP
180 Townsend Street, Third Floor
San Francisco, CA 94107
(Counsel for Respondent Google Inc.)

Aden J. Fine, Esquire
American Civil Liberties Union Foundation
125 Broad Street
New York, New York 10004
(Counsel for Plaintiffs, *ACLU v. Gonzalez*, E.D. Pa. No. 98-cv-5591)

/s/
JOEL McELVAIN
Attorney