

1 KASSRA P. NASSIRI (215405)  
 (knassiri@nassiri-jung.com)  
 2 CHARLES H. JUNG (217909)  
 (cjung@nassiri-jung.com)  
 NASSIRI & JUNG LLP  
 3 47 Kearny Street, Suite 700  
 San Francisco, California 94108  
 4 Telephone: (415) 762-3100  
 Facsimile: (415) 534-3200  
 5

6 MICHAEL J. ASCHENBRENER  
 (mja@aschenbrennerlaw.com) (277114)  
 ASCHENBRENER LAW, P.C.  
 7 795 Folsom Street, First Floor  
 San Francisco, CA 94107  
 8 Telephone: (415) 813-6245  
 Facsimile: (415) 813-6246  
 9

10 Attorneys for Plaintiff and the Putative Class

11 **UNITED STATES DISTRICT COURT**  
 12 **NORTHERN DISTRICT OF CALIFORNIA**  
 13 **SAN JOSE DIVISION**  
 14

15 PALOMA GAOS, an individual, on behalf of  
 16 herself and all others similarly situated,  
 17 **Plaintiff,**  
 18 v.  
 19 GOOGLE INC., a Delaware Corporation,  
 20 **Defendant.**  
 21

Case No. 5:10-cv-04809-EJD

CLASS ACTION

**DECLARATION OF KASSRA P.  
 NASSIRI IN SUPPORT OF  
 PLAINTIFF'S OPPOSITION TO  
 GOOGLE'S MOTION TO DISMISS  
 FIRST AMENDED COMPLAINT  
 PURSUANT TO RULES 12(b)(1)  
 AND 12(b)(6)**

Hearing Date: October 28, 2011  
 Time: 9:00 a.m.  
 Place: Courtroom 1, 5th Floor  
 Judge: Hon. Edward J. Davila

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

I, Kassra P. Nassiri, declare as follows:

1. I am a member in good standing of the Bar of the State of California and a partner with the firm of Nassiri & Jung, LLP, counsel for plaintiff in this matter. I submit this Declaration in support of Plaintiff's Opposition to Google's Motion to Dismiss the First Amended Complaint Pursuant to Rules 12(b)(1) and 12(b)(6). I make this declaration on my own personal knowledge and, if called as a witness, would testify to the matters set forth herein.
2. Attached hereto as Exhibit 1 is a true and correct copy of Google's Opposition to the Government's Motion to Compel in *Gonzales v. Google, Inc.*, Case No. 5:06-mc-80006-JW, Dkt. No. 12 (N.D. Cal. February 17, 2006), as retrieved from the Public Access to Court Electronic Records website on October 7, 2011.
3. Attached hereto as Exhibit 2 is a true and correct copy of the Declaration of Matt Cutts in *Gonzales v. Google, Inc.*, Case No. 5:06-mc-80006-JW, Dkt. No. 14 (N.D. Cal. February 17, 2006), as retrieved from the Public Access to Court Electronic Records website on October 7, 2011.

I declare under penalty of perjury under the laws of the United States and the State of California that the foregoing is true and correct. Executed on October 7, 2011 in San Francisco, California.

/s/ Kassra P. Nassiri  
Kassra P. Nassiri

# EXHIBIT 1

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

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

10 **UNITED STATES DISTRICT COURT**  
11 **NORTHERN DISTRICT OF CALIFORNIA**  
12 **SAN JOSE DIVISION**

14 ALBERTO R. GONZALES, in his official  
capacity as ATTORNEY GENERAL OF THE  
15 UNITED STATES,

16 Movant,

17 v.

18 GOOGLE INC.,

19 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.

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28

**CONTENTS**

I. INTRODUCTION ..... 1

II. BACKGROUND ..... 2

III. ARGUMENT ..... 3

    A. The Standard for Compelling a Third Party to Respond to a  
        Subpoena ..... 3

    B. The Government Fails to Establish That the Data Will Lead  
        to Admissible Evidence ..... 5

        1. Search Queries ..... 6

        2. URLs ..... 8

        3. The Data Is Not Useful for Any Study ..... 9

    C. The Subpoena Demands Google's Valuable Trade Secrets ..... 9

        1. The Demanded Data Contains Valuable Trade  
            Secrets and Confidential Commercial Information ..... 10

        2. Disclosure of Google's Trade Secrets Is A  
            Significant Possibility ..... 12

        3. The Government Has Not Shown a Substantial  
            Need ..... 13

    D. The Subpoena Imposes an Undue Burden on Google ..... 15

        1. The Time and Resources Required to Pull the  
            Requested Information Would Be Significant ..... 16

        2. The Government's Offer to Collaborate Is  
            Inadequate and Unrealistic ..... 17

        3. The Production of the Requested Data Will Result  
            in a Chilling Effect on Google's Business and User  
            Trust ..... 18

        4. Google Should Not Bear the Burden of Responding  
            to Potentially Inadequate Process Based on ECPA ..... 18

IV. CONCLUSION ..... 21

**TABLE OF AUTHORITIES**

**Cases**

1		
2		
3		
4	<i>Am. Standard Inc. v. Pfizer, Inc.</i> , 828 F.2d 734 (Fed. Cir. 1987).....	10, 15
5	<i>Compaq Computer Corp. v. Packard Bell Elecs, Inc.</i> , 163 F.R.D. 329 (N.D. Cal. 1995) .....	9, 12, 15
6	<i>Crowley v. Cyberspace Corp.</i> , 166 F. Supp. 2d 1263 (N.D. Cal. 2001) .....	19
7	<i>Cusumano v. Microsoft Corp.</i> , 162 F.3d 708 (1 <sup>st</sup> Cir. 1998).....	16
8	<i>Dart Indus. Co. v. Westwood Chem. Co.</i> , 649 F.2d 646 (9 <sup>th</sup> Cir. 1980).....	4, 15
9	<i>Daubert v. Merrell Dow Pharms., Inc.</i> , 509 U.S. 579 (1993).....	6
10	<i>DIRECTV Inc. v. Trone</i> , 209 F.R.D. 455 (C.D. Cal. 2002).....	3, 10, 14
11	<i>Echostar Commc'n Corp. v. News Corp. Ltd.</i> , 180 F.R.D. 391 (D. Colo. 1998).....	3, 9, 10, 15
12	<i>Heidelberg Americas, Inc. v. Tokyo Kikai Seisakusho, Ltd.</i> , 333 F.3d 38 (1 <sup>st</sup> Cir. 2003).....	9
13	<i>In re Jetblue Airways Corp. Privacy Litigation</i> , 379 F. Supp. 2d 299 (E.D.N.Y. 2005) .....	19
14	<i>In re Surety Ass'n of America</i> , 388 F.2d 412 (2d Cir. 1967) .....	9
15	<i>Instituform Technologies, Inc. v. CAT Contracting, Inc.</i> , 168 F.R.D. 630 (N.D. Ill. 1996) .....	16
16	<i>Katz v. Batavia Marine &amp; Sporting Supplies, Inc.</i> , 984 F.2d 422 (Fed. Cir. 1993).....	9, 10
17	<i>Mattel Inc. v. Walking Mountain Prods.</i> , 353 F.3d 792 (9th Cir. 2003) .....	3, 4, 15
18	<i>Moon v. SCP Pool Corp.</i> , 2005 WL 3526513 (C.D. Cal. Dec. 7, 2005).....	3
19	<i>Nicholas J. Murlas Living Trust v. Mobil Oil Corp.</i> , 1995 WL 124186 (N.D. Ill. March 20, 1995).....	5
20	<i>Palumbo v. Schulman</i> , 1998 WL 436367 (S.D.N.Y. July 27, 1998).....	9, 17
21	<i>Premium Serv. Corp. v. Sperry &amp; Hutchinson Co.</i> , 511 F.2d 225 (9 <sup>th</sup> Cir. 1975).....	3, 17
22		
23		
24		
25		
26		
27		
28		

1 *Robin Singh Educ. Servs., Inc. v. Excel Test Prep*, 2004 WL 2554454  
 2 (N.D. Cal. Nov. 9, 2004) .....5, 9

3 *Travelers Indem. Co. v. Metro Life Ins. Co.*, 228 F.R.D. 111 (D. Conn.  
 4 2005)..... 15

5 *Truswal Sys. Corp. v. Hydro-Air Engineering, Inc.*, 813 F.2d 1207 (Fed.  
 6 Cir. 1987)..... 12

7 *U. S. ex rel. Schwartz v. TRW, Inc.*, 211 F.R.D. 388 (C.D. Cal. 2002) .....4

8 **Statutes**

9 18 U.S.C. § 2510(14)..... 20

10 18 U.S.C. § 2510(15)..... 19

11 18 U.S.C. § 2703 ..... 19

12 18 U.S.C. § 2703(b)(2) ..... 20

13 18 U.S.C. § 2711(2)..... 20

14 18 U.S.C. §§ 2701-2712 ..... 18

15 **Regulations and Rules**

16 Fed. R. Civ. P. 45.....4, 5

17 Fed. R. Evid. 702 .....6

18 **Other Authorities**

19 *Searching and Seizing Computers and Obtaining Electronic Evidence in*  
 20 *Criminal Investigations*, United States Dep't of Justice, Computer  
 21 Crime and Intellectual Property Section Criminal Division (July 2002) ..... 18

22 *Technology: Searching for Searches*, Newsweek, Jan. 30, 2006 ..... 18

## I. INTRODUCTION

1  
2 Google users trust that when they enter a search query into a Google search box, not only  
3 will they receive back the most relevant results, but that Google will keep private whatever  
4 information users communicate absent a compelling reason. The Government's demand for  
5 disclosure of untold millions of search queries submitted by Google users and for production of a  
6 million Web page addresses or "URLs" randomly selected from Google's proprietary index would  
7 undermine that trust, unnecessarily burden Google, and do nothing to further the Government's  
8 case in the underlying action.

9  
10 Fortunately, the Court has multiple, independent bases to reject the Government's Motion.  
11 *First*, the Government's presentation falls woefully short of demonstrating that the requested  
12 information will lead to admissible evidence. This burden is unquestionably the Government's.  
13 Rather than meet it, the Government concedes that Google's search queries and URLs are not  
14 evidence to be used at trial at all. Instead, the Government says, the data will be "useful" to its  
15 purported expert in developing some theory to support the Government's notion that a law banning  
16 materials that are harmful to minors on the Internet will be more effective than a technology filter  
17 in eliminating it.

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

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



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

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

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

## 16 **II. BACKGROUND**

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

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

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

### 15 III. ARGUMENT

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

17 A court must quash or modify a subpoena issued to a non-party if it subjects that person to  
18 undue burden. *Mattel Inc. v. Walking Mountain Prods.*, 353 F.3d 792, 813 (9th Cir. 2003). In  
19 analyzing burden, a court must balance the needs of the requesting party with the burden imposed  
20 on the non-party. *See Premium Serv. Corp. v. Sperry & Hutchinson Co.*, 511 F.2d 225 (9<sup>th</sup> Cir.  
21 1975). In performing this balancing test, a court must consider, *inter alia*, the relevance of the  
22 requested information to the claim or defenses in the underlying action. *Moon v. SCP Pool Corp.*,  
23 2005 WL 3526513, at \*4 (C.D. Cal. Dec. 7, 2005); *see DIRECTV Inc. v. Trone*, 209 F.R.D. 455,  
24 460 (C.D. Cal. 2002).

25 Unequivocally, the Government bears the burden of establishing relevance. *Echostar*  
26 *Comm'n Corp. v. News Corp. Ltd.*, 180 F.R.D. 391, 394 (D. Colo. 1998). The burden of  
27

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

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

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

21 Here the Government fails to show that the data it seeks actually will lead to anything  
22 admissible in the underlying case because the data simply is not relevant to any claim or defense.  
23 Having not crossed that initial threshold, it almost goes unsaid that they have not met the heavier  
24 burden they bear to force the disclosure of Google's confidential information. In the end, their  
25 purported need pales in comparison to the burden imposed on Google in meeting the request.  
26

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

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

4 (1) the production of [a week of search queries] would permit the  
5 Government to evaluate whether COPA or filtering software is more  
6 effective in restricting access to harmful-to-minor materials in response to  
7 searches as they are actually performed by present day users of the Internet  
(Motion, at 6); and

8 (2) the production of [randomly selected URLs] will permit the  
9 Government to review a sample set of Internet addresses available to be  
10 retrieved from the search engines operated by Google and by other entities.  
11 From that set, the Government will be able to review the sample to draw  
12 conclusions as to the prevalence of harmful-to-minor material on the  
13 portion of the Internet that is retrievable through search engines. Motion,  
14 at 8.

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

17 The Government's showing is mere argument, not the required proof of the demanded  
18 data's relevance to their claim. What the Government has failed to understand or admit is that the  
19 data it seeks from Google has no conceivable use in furthering either of the Government's points.  
20 In the absence of a coherent theory of relevance, the Government's demand must be denied. *See*  
21 *Robin Singh Educ. Servs., Inc. v. Excel Test Prep*, 2004 WL 2554454, at \*2 (N.D. Cal. Nov. 9,  
22 2004) (denying motion to compel since, *inter alia*, "[o]ther than the naked statement of this  
23 argument, Plaintiff does not even attempt to show how either of these two bases for relevance in  
24 fact obtain"); *Nicholas J. Murlas Living Trust v. Mobil Oil Corp.*, 1995 WL 124186, at \*5 (N.D.  
25 Ill. March 20, 1995) (denying motion to compel in part since, *inter alia*, plaintiff failed to  
26 articulate a coherent theory which would explain how the requested information was relevant).  
27

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

#### 4 **1. Search Queries**

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

17 Similarly, Google users can and do modify their environment to generate certain types of  
18 search results. For example, users may employ Google's SafeSearch pornography filter to limit  
19 results. They may use Google's advanced search programs to deliver personalized or customized  
20 search results as well. Cutts Decl., ¶ 10. Therefore, the same query will generate different search  
21

---

22 <sup>1</sup> The court should view the Cutts Declaration as standing in strong contrast to the  
23 Government's declarant, Professor Phillip Stark, a statistician who apparently has been hired to  
24 produce a study to support the Government's contentions. The Stark Declaration is vague,  
25 cursory, and uninformed about the operation of Google's search engine. In any event, Professor  
26 Stark's opinion ought to be viewed with some scrutiny. Although positioned as the Government's  
27 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.*

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

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

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

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

23 Finally, Google's system is not static. Algorithms regularly change. The identical search  
24 query submitted today may yield a different result than the identical search conducted yesterday.  
25 In no meaningful way can it be said that a past week's worth of search queries will yield URL  
26 responses as performed by "present day users of the Internet." Motion, at 6. Past searches tell the  
27

1 Government nothing about URLs available from those searches, now or in the future. Cutts Decl.,  
2 ¶ 16.

## 3 2. URLs

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

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

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







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

5  
6 **1. The Demanded Data Contains Valuable Trade Secrets and  
Confidential Commercial Information**

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

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

24 Like queries, from even a sample of URLs that Google has indexed, one could estimate,  
25 among other things, the size of Google's index; how deeply Google crawls in different countries or  
26 languages; and the ability of Google's crawl metrics to measure the reputation of pages or  
27

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

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

18 The very fact that the Government is so uninformed about the value of search and URL  
19 information and so dismissive of Google's interest in protecting it speaks volumes about why the  
20 Court should protect Google from this compelled disclosure. The Government's cavalier attitude  
21 undermines any credibility in the assertion it later makes that it can or will protect Google against  
22

---

23  
24 <sup>2</sup> Google routinely receives and refuses requests from researchers and analysts for search  
25 query and URL data. For example, Google has denied researchers query logs to protect both its  
26 trade secrets and confidential commercial information, and to protect the privacy of its users.  
27 Cutts Decl., ¶ 35. Ironically, almost six years ago, Professor Stark obtained a small sample of  
28 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.

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

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

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

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

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

21  
22  
23 <sup>3</sup> Unlike the Protective Order in *Compaq*, which prohibited access by the parties'  
24 employees, agents, and even in-house counsel, the Protective Order signed by the Government and  
25 the ACLU gives the parties' employees, witnesses, consultants and counsel access to the  
26 information and in no way protects Google's proprietary trade secrets and confidential commercial  
27 information. See *Compaq*, 163 F.R.D. at 339. Further, the Government's reliance on *Truswal Sys.*  
28 *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.

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

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

### 17 **3. The Government Has Not Shown a Substantial Need**

18 The Government asserts that Google's query and URL data "would be of value to the  
19 Government in its development of its overall sample of queries" because Google has the largest  
20 market share of the Web search Market. Motion, at 6. But this is no showing of "substantial  
21 need" at all. Professor Stark does not say the data is "essential," that there is no alternative to it, or  
22 that the study he proposes will not stand without it; nor does he explain why a study based on  
23

---

24 <sup>4</sup> Of course, variations in search queries entered are meaningless. What matters are the  
25 URLs returned in response to a given search. As noted above, Professor Stark cannot surmise  
26 what URL will be returned without knowing how the particular search engine works. Certainly,  
27 any comparison of queries run, to be relevant, would have to use identical terms, would depend on  
28 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.

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

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

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

22 Mr. Cutts presents numerous examples of alternative sources for queries and URLs in his  
23 Declaration. Cutts Decl., ¶¶ 36 – 38. Google describes a few of these alternatives here.  
24 Metasearch engines Dogpile and MetaCrawler each provide services that allow anyone to view  
25 lists of queries through their search services, "SearchSpy" and "Metaspy" respectively.  
26 SearchSpy, available at: <http://www.dogpile.com/info.dogpl/searchspy/>, allows one to choose to  
27

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

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

12       The query and URL data sought by the Government is available from other sources in form  
13 and content more suited to its proposed study, for whatever value there may be in the enterprise.  
14 Whatever else that can be said, there is no necessity when similar data is available from other  
15 sources. *Compare Compaq*, 163 F.R.D. at 338 (substantial need for a portion of requested  
16 information concerning industry practice was established when information could be obtained by  
17 no source other than third party industry member) *with Am. Standard*, 828 F.2d at 743 (need not  
18 established when information was publicly available); *Travelers Indem. Co. v. Metro Life Ins. Co.*,  
19 228 F.R.D. 111, 114 (D. Conn. 2005) (granting motion to quash subpoena served on non-party  
20 since, *inter alia*, requested information was otherwise available).

21 **D. The Subpoena Imposes an Undue Burden on Google**

22       A court must quash or modify a subpoena issued to a non-party if it subjects a person to an  
23 undue burden. *Mattel*, 353 F.3d at 813; *see also Echostar*, 180 F.R.D. at 394; *Travelers Indem.*  
24 *Co. v. Metro. Life Ins. Co.*, 228 F.R.D. 111 (D. Conn. 2005). Google's non-party status weighs  
25 heavily against the Government in a burden analysis. *See Dart*, 649 F.2d at 649; *Travelers*, 228  
26 F.R.D. at 113 ("courts also give special weight to the burden on non-parties of producing  
27

1 documents to parties involved in litigation") (citing *Cusumano v. Microsoft Corp.*, 162 F.3d 708,  
2 717 (1<sup>st</sup> Cir. 1998)).

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

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

11 First and most basically: Google does not maintain query or URL data in the ordinary  
12 course of business in the format requested by the Government. For this reason alone, the  
13 Government's Motion should be denied. *Instituform Technologies, Inc. v. CAT Contracting, Inc.*,  
14 168 F.R.D. 630, 633 (N.D. Ill. 1996) ("Rule 45 does not contemplate that a nonparty will be  
15 forced to create documents that do not exist"). Nor is there a program that could simply gather the  
16 requested data. Therefore, Google would have to create new code to format and extract the query  
17 and URL data from its many computer banks. Pulling each type of data would require multiple  
18 teams of Google engineers, diverted from their normal job responsibilities, to research, develop,  
19 write, implement, test, fix and execute new computer code. Finally, the selected data must be  
20 extracted and copied into a format that can be provided to the Government. In total, the process of  
21 gathering the queries and URLs would likely consume up to eight full-time days of engineering  
22 time. This time, of course, would have to be covered by other engineers within Google. Cutts  
23 Decl., ¶¶ 39-42.

24 Even if the Government were to pay Google for its engineers' time, executing the searches  
25 required by the Government's requests would command extended hours of processing time on  
26 Google's computers. Running these programs above and beyond the normal demand on Google's  
27



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

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

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

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



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

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

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

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

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

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

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

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

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

19 Further, ECPA places similar restrictions on the disclosure of stored communications to  
20 the government by providers of remote computing services. A "'remote computing service' means  
21 the provision to the public of computer storage or processing services by means of an electronic  
22

---

23  
24 <sup>5</sup> Some courts have held that a mere "user" of an ECS provided by another is not itself an  
25 ECS. *See Crowley v. Cyberspace Corp.*, 166 F. Supp. 2d 1263, 1270 (N.D. Cal. 2001) (holding  
26 that Amazon.com is not an ECS because it did not provide users the ability to communicate); *see*  
27 *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.

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

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

7 (A) on behalf of, and received by means of electronic transmission from (or  
8 created by means of computer processing of communications received by  
9 means of electronic transmission from), a subscriber or customer of such  
remote computing service; and

10 (B) solely for the purpose of providing storage or computer processing  
11 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.

12 18 U.S.C. § 2703(b)(2).

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

19 That the Government has asked Google to remove any personally identifiable information  
20 from the content of the search queries is of no moment. ECPA makes no exception for  
21 anonymous or anonymized content. Surely, the Government does not mean to suggest that it  
22 could obtain millions of emails stored in Google's servers simply by asking Google to remove the  
23 "To" and "From" lines. It matters not that it might even be helpful or relevant to the Government's  
24 case to show that email is used to send content harmful to minors. Content is off limits under  
25 ECPA except in rare cases and when procedural safeguards are followed. Google should not bear  
26

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

3  
4 **IV. CONCLUSION**

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

9 DATED: February 17, 2006.

10 Respectfully submitted,

11 **PERKINS COIE LLP**

12 By: /s/\_\_\_\_\_

13 ALBERT GIDARI, JR.

14 LISA A. DELEHUNT

15 180 Townsend Street, 3<sup>rd</sup> Floor

16 San Francisco, California 94107-1909

17 Telephone: (415) 344-7000

18 Facsimile: (415) 344-7050

19 Email: [AGidari@perkinscoie.com](mailto:AGidari@perkinscoie.com)

20 Email: [LDelehunt@perkinscoie.com](mailto:LDelehunt@perkinscoie.com)

21 Attorneys for Google

# EXHIBIT 2

1 ALBERT GIDARI, JR., *pro hac vice*  
([AGidari@perkinscoie.com](mailto:AGidari@perkinscoie.com))

2 PERKINS COIE LLP  
1201 Third Avenue, Suite 4800  
3 Seattle, WA 98101  
Telephone: (206) 359-8000  
4 Facsimile: (206) 359-9000

5 LISA A. DELEHUNT, Bar No. 228551  
([LDelehunt@perkinscoie.com](mailto:LDelehunt@perkinscoie.com))

6 PERKINS COIE LLP  
180 Townsend Street, 3rd Floor  
7 San Francisco, California 94107-1909  
Telephone: (415) 344-7000  
8 Facsimile: (415) 344-7050

9 Attorneys for Respondent  
GOOGLE INC.

10

11

**UNITED STATES DISTRICT COURT  
NORTHERN DISTRICT OF CALIFORNIA  
SAN JOSE DIVISION**

12

13

14

15 ALBERTO R. GONZALEZ, in his official  
16 capacity as ATTORNEY GENERAL OF THE  
UNITED STATES,

CASE NO. 5:06-mc-80006-JW

17

Movant,

DECLARATION OF MATT CUTTS

18

v.

19

GOOGLE INC.,

20

Respondent.

21

22

MATT CUTTS states as follows:

23

24

25

26

27

28

**I. INTRODUCTION**

1  
2 1. I am a Senior Staff Software Engineer and head of the Webspam group at Google.  
3 I have held this position since November 2004. I have worked with Google since February 2000  
4 as a technical engineer in one capacity or another. I have personal knowledge of the facts set forth  
5 below and, if called as a witness about those facts, could testify competently thereto.

6 2. Currently, I manage the Webspam group. "Webspam" does not refer to spam  
7 emails, but rather to pages from the World Wide Web ("Web") that violate Google's quality  
8 guidelines in an attempt to rank higher on Google. For example, a page containing pornographic  
9 material that attempted to show up as a result to a search for "Disney cartoons" would be  
10 considered webspam. Accordingly, I am knowledgeable about how Google crawls, indexes, and  
11 ranks web pages. I have written a variety of programs for Google, including programs having to  
12 do with filtering pornography and detecting web pages that attempt to bypass Google's quality  
13 guidelines. I have a Master of Science in Computer Science, *summa cum laude*, from University  
14 of North Carolina at Chapel Hill. I am recipient of a National Science Foundation Graduate  
15 Research Fellowship, a Link Foundation Fellowship, and a Gaines Fellowship in the Humanities.  
16 I also have a Bachelor of Science in Computer Science and Mathematics, *summa cum laude*, from  
17 the University of Kentucky. I have co-authored and published a variety of articles relating to  
18 computer graphics and computer vision.

19 3. I have reviewed the Motion to Compel Compliance with Subpoena Duces Tecum  
20 brought by Alberto R. Gonzales (the "Government") against Google, Inc. and the supporting  
21 declarations and exhibits, including the Declaration of Professor Philip B. Stark, Ph.D. The  
22 Government has requested that Google produce a multistage random sample of one million URLs  
23 from Google's databases. URL stands for "Uniform Resource Locator" and acts as the global  
24 address of documents and other resources on the World Wide Web ("Web"). URLs consist of the  
25 protocol (e.g. HTTP) and the Internet Protocol ("IP") address or domain name. The Government  
26 has suggested that Google select at random 100 of its data centers containing URLs, and then  
27

1 select at random 10,000 URLs from each of those data centers. In his declaration, Professor Stark  
2 has stated that he could be involved in the random selection process.

3 4. Additionally, the Government has requested that Google produce every single  
4 search query entered at www.google.com over a given week.

5 **II. SUMMARY OF DECLARATION**

6 5. In my declaration, I testify as to the following:

7 A. In Section III, I summarize how Google's search engine works.

8 B. In Section IV, I explain how examining the search queries Google's users  
9 have typed into Google's search engine at www.google.com does not reveal accurate  
10 information about the likelihood of minors' exposure to material the Government deems  
11 Harmful to Minors ("HTM") or about the search patterns of users.

12 C. In Section V, I explain that there is little to no correlation between a URL's  
13 presence in Google's index and if or when that URL would be returned as a result  
14 responsive to a specific set of query words.

15 D. In Section VI, I discuss the fact that, in order to protect the privacy of its  
16 users, Google does not disclose to third parties the raw records of searches entered into its  
17 search engine.

18 E. In Section VII, I explain that Google does not disclose the queries it  
19 receives from its users because it offers information about the nature of Google's users and  
20 of Google's market share in both the United States and other countries and languages,  
21 which could allow Google's competitors to compete more effectively with Google.  
22 Additionally, Google does not disclose its index of URLs or any such large samples from  
23 its index, because one could glean from a sample of the URLs from Google's index  
24 confidential and proprietary information about how Google's search engine works.

25 F. In Section VIII, I explain those measures of which I am personally aware  
26 that Google undertakes to protect the confidentiality of its proprietary information.



1 G. In Section IX, I discuss reasonable, viable alternative public sources of the  
2 information the Government is seeking from Google.

3 H. In Section X, I explain the burden that would be imposed on Google if it  
4 were required to respond to the Government's subpoena.

5 **III. GOOGLE'S SEARCH ENGINE**

6 6. Google provides the world's most-used search engine at www.google.com. In  
7 general terms, search engines allow users of the Web to search a considerable portion of the vast  
8 amount of material located on the Web by entering text queries into the engine. The Web is  
9 composed of billions of publicly accessible web sites from around the world and other information  
10 sources that web browsers can access. Search engines offer this search capability by creating an  
11 index of certain Web content through the use of programs that "crawl" the Web and automatically  
12 fetch Web pages. The search engine then allows users to search that index of certain Web content  
13 through an interface such as that found at www.google.com. When a user enters a query, the  
14 engine searches its Web index and retrieves results relevant to the query. Google's search engine  
15 is open to the public. Google is the leader among the major search engines (including Yahoo!,  
16 Ask Jeeves, and MSN Search) to a large degree because of its sophisticated and proprietary  
17 technology that returns the most relevant and useful results in response to user queries.

18 7. Some of Google's services, like Google News, allow users to store or establish  
19 repeat search queries. Google users may set up recurring searches with results sent to email  
20 accounts at defined intervals. Google then processes these requests for its users and sends the  
21 results to the email accounts, wireless phones, or other designated mobile devices.

22 **IV. SEARCH QUERIES**

23 8. Examining the search queries Google's users have typed into Google's search  
24 engine at www.google.com does not reveal accurate information about how likely it is that minors  
25 might be exposed to material the Government deems Harmful to Minors ("HTM") or about the  
26 search patterns of users.

1           9.       First, queries entered into Google's search engine have no easily computed  
2 correlation to how likely it is that any user might be exposed to what the Government might  
3 designate as HTM. This is because Google uses a number of proprietary and confidential methods  
4 to return search results to users. Over a period of years, and at enormous expense, Google has  
5 developed and implemented multiple unparalleled technologies to crawl and index a considerable  
6 portion of the Web and rank the results so as to ensure that Google's users receive the most  
7 relevant search results possible. Google employs proprietary methods to determine which  
8 documents to crawl from the web and sophisticated techniques to decide which documents to  
9 retrieve for a query. Google uses a combination of algorithms and a unique network architecture  
10 (among other functionalities) to accomplish this. Most of these crawling, collecting, and sorting  
11 techniques are not known outside Google, and they are particularly kept secret from competitors.

12           10.       Additionally, users entering the same query at different computers may receive  
13 different search results. The search results presented to the user may vary based on whether  
14 SafeSearch (Google's pornography filter) is enabled for that user, the query language requested by  
15 the user, or whether the user belongs to Google's program to deliver personalized search results.  
16 Google employs additional proprietary techniques that can return different results to different  
17 users for the same query. In short, a query alone is not sufficient to determine the results that  
18 Google will present to a user.

19           11.       There is no way to tell what queries have been entered by minors and what queries  
20 have been entered by adults.

21           12.       Moreover, the Government's request for search queries in an attempt to understand  
22 how individuals use search engines is flawed because it does not distinguish between "real"  
23 queries entered by individual users, automatic queries generated by a program, known as "bot"  
24 queries, and artificial queries generated by real users.

25           13.       For example, a single person may write a short program to send the query  
26 "underage sex resources" hundreds, thousands, or even tens of thousands of times a day to search  
27 engines using a program, or "bot query." I am also aware of individuals who have written

1 programs that detect whether a computer is connected to the Web by sending periodic queries to  
2 Google; such programs can send thousands of queries to Google each day. Google also receives a  
3 fraction of its queries from malicious programs, such as the Santy worm.

4 14. In addition to queries sent by "bots," test programs, and worms, an individual web  
5 page owner (also known as a "Webmaster") may send dozens of queries by hand to check on how  
6 his or her websites rank in Google. Many website owners check their rankings every day, which  
7 can cause further skew in the query log. I am also aware of efforts by some Google users to  
8 deliberately send pornography queries to Google in reaction to the Government's subpoena. One  
9 individual wrote a feature for the popular Firefox web browser that will send a random  
10 pornography query to Google whenever a user performs a normal query, as if the pornography  
11 query had also been entered by the user. Attached as Exhibit A is a true and correct copy of this  
12 blog entry which can be found at <http://www.hughes-family.org/wordpress/2006/01/23/help-the-justice-department-out-via-greasemonkey/>, and which was printed on February 16, 2006.

14 15. Without removing these artificial and automatic queries, one week of raw query  
15 logs will be skewed beyond usability for many purposes. It would be difficult—perhaps  
16 impossible—for most researchers to distinguish between such artificial queries and real queries  
17 without considerable effort and the use of proprietary techniques.

18 16. Google's proprietary techniques for returning search results are not static, and  
19 Google's algorithms change regularly. Thus, the identical search query run in Google's search  
20 engine today is likely to yield different search results than an identical search conducted yesterday,  
21 last week, or last month.

## 22 V. GOOGLE'S INDEX OF URLS

23 17. The presence of a URL in Google's index is not in any way representative of the  
24 frequency that the URL will be shown to a user. There is little to no correlation between a URL's  
25 presence in the index Google maintains of copies of documents it has collected from crawling the  
26 web and if or when that URL would be returned as a result responsive to a specific set of query  
27 words. Using its proprietary and confidential technology, Google employs more than 100 factors

1 in scoring documents for relevancy besides just the document's URL. Some of the factors are  
2 straightforward. For example, Google considers whether query words are present in the title of the  
3 document, how many times the query words appear in the document, the proximity of the query  
4 words to each other, as well as a document's PageRank, which is Google's patented method of  
5 measuring the reputation of a page. Most of Google's scoring factors and how they are combined  
6 is confidential.

7 18. Beyond scoring factors, some documents in Google's index will have additional  
8 demotions or will be blocked for spam or legal reasons. For example, Google's index contains  
9 documents for which we have received valid complaints under the Digital Millennium Copyright  
10 Act, and Google does not return those documents in our search results.

11 19. For all these reasons, it does not follow that a sample of URLs from Google's index  
12 will indicate how often searches will return HTML material. Simply put, the fact that a URL in  
13 Google's index may appear to contain content the Government considers HTML is not  
14 representative of whether and how often HTML content appears as a result of a query. For  
15 example, if 5 percent of the URLs listed in Google's index contain HTML content, that fact alone  
16 does not mean that any given search would yield 5 percent HTML content.

17 20. The adage "you can't judge a book by its cover" applies on the Web: URLs alone  
18 do not indicate whether content at that URL is HTML. A URL may or may not be logically  
19 connected to its content. Sites may have names that suggest explicit sexual material, but actually  
20 do not contain HTML content. A URL such as  
21 <http://www.pbs.org/wgbh/pages/prontline/shows/porn/etc/links.html> contains the word "porn" but  
22 provides links to anti-pornography organizations such as the American Family Association and the  
23 Family Research Council. Likewise, a URL such as <http://www.porn-free.org/> may appear  
24 pornographic from the URL, but is a faith-based site that is against pornography. The reverse is  
25 also true: URLs may have name that seem innocent but actually contain material the Government  
26 might consider HTML. The classic example was that [www.whitehouse.gov](http://www.whitehouse.gov) was a harmless website  
27 about the White House, while, until recently, [www.whitehouse.com](http://www.whitehouse.com) contained pornographic

1 material. Furthermore, the content of web pages at URLs is fluid and dynamic, and the actual  
2 content can be difficult to ascertain. For example, until recently the URL  
3 <http://www.crisiscentersyr.org> was evidently owned by a rape crisis center. It appears that the  
4 domain expired in late 2005 and was instead registered by someone else. The new owner kept the  
5 appearance of the site, making it appear it still belongs to a rape crisis center, but added links to  
6 pornography at the bottom of the document. In September 2005, that URL contained non-  
7 pornographic material, while in January 2006, it contained links to pornographic material. In  
8 short, a URL alone is insufficient to determine whether a document is harmful to minors. This  
9 example also shows the dynamic nature of Web content. The content of a web page can change at  
10 anytime, and some Web pages update their content daily or more frequently.

11 21. Additionally, in an attempt to rank higher in Google's search results, some  
12 Webmasters show different results to Google than they do to individual users. That is, Google  
13 uses a program to crawl the Web to find Web pages for its index. Google's program comes from a  
14 specific set of IP addresses and identifies itself as "Googlebot." Some Webmasters program their  
15 Web pages to show certain content to Google, content that may be innocent and may seem  
16 relevant to a search query. The Webmaster, however, shows different content to individual users,  
17 and the content may be pornographic or material otherwise prohibited by Google's guidelines.  
18 Such "bait and switch" tactics are called "cloaking." Ultimately, through its diligence, Google  
19 finds and removes those documents, but the nature of these Web pages is another example of how  
20 a URL does not indicate what content will be displayed in response to a search query.

21 22. The presence of URLs in Google's index is not reflective of the entire Web.  
22 Google uses its technology to crawl and index only the best documents from the Web. The Web  
23 can be viewed as having an infinite number of pages. For example, a single web server running a  
24 calendar application could generate dynamic web pages with dates going forward forever.  
25 Therefore, Google's mission must be to retrieve and index the most useful pages that it can from  
26 the infinite number of potential pages that can be retrieved.

1           23.     The only way to get a current and accurate snapshot of the search results that would  
2 be returned by a query into Google's search engine would be to run the query on Google's search  
3 engine. Notably, running millions of queries on Google above and beyond the normal use of  
4 Google would burden Google's system and possibly shut it down. Alternatively, one would need  
5 to understand how Google crawls the web, collects URLs, sorts them, indexes them, ranks them,  
6 and returns them as search results—in other words, exactly how Google's crown-jewel trade  
7 secrets function.

#### 8           **VI. SEARCH QUERIES AND PERSONALLY IDENTIFYING INFORMATION**

9           24.     Google does not publicly disclose the searches queries entered into its search  
10 engine. If users believe that the text of their search queries into Google's search engine could  
11 become public knowledge, they may be less likely to use the search engine for fear of disclosure  
12 of their sensitive or private searches for information or websites.

13           25.     There are ways in which a search query alone may reveal personally identifying  
14 information. For example, many internet users have experienced the mistake of trying to copy-  
15 and-paste text into the search query box, only to find that they have pasted something that they did  
16 not intend. Because Google allows very long queries, it is possible that a user may paste a  
17 fragment of an email or a document that would tie the query to a specific person. Users could also  
18 enter information such as a credit card, a social security number, an unlisted phone number or  
19 some other information that can only be tied to one person. Some people search for their credit  
20 card or social security number deliberately in order to check for identity theft or to see if any of  
21 their personal information is findable on the Web.

#### 22           **VII. GOOGLE'S COMPETITIVE AND CONFIDENTIAL INFORMATION**

23           26.     Another reason that Google does not disclose the queries it receives from its users  
24 is because it offers information about the nature and effectiveness of Google's search engine and  
25 Google's market share in the United States and other countries, which could allow Google's  
26 competitors to compete more effectively with Google. Disclosing an entire week's worth of  
27 queries would give an estimate of the number of queries that Google processes, which could be  
28



1 used to deduce market share among search engines. Google protects this information not only by  
2 not disclosing search queries but by not disclosing even the amount of computers Google  
3 maintains to run its search engine. A week's worth of queries would also indicate the percentage  
4 of queries Google tends to receive in each language, which would allow competitors to estimate  
5 Google's relative market share in a given language or country, to learn opportunities for market  
6 growth, and to decide where to allocate resources for each language or country. The data would  
7 also indicate the average length of query that Google's search engine typically receives and how  
8 Google's users search, such as what percentage of queries use special search operators or  
9 punctuation. This information could allow competitors to better understand the type of users who  
10 seek out Google or to allocate resources on specific search operators. Similarly, the queries could  
11 relate information regarding whether users tend to use Google for navigational help, for research,  
12 or for shopping.

13         27. Google also does not disclose its index or such a large sample from its index. One  
14 could estimate from a sample of the URLs from Google's index information such as (i) the size of  
15 Google's index; (ii) how "deeply" Google crawls in different countries or languages (i.e., how  
16 many URLs are crawled from each website on average); and (iii) the ability of Google's crawl  
17 metrics to measure the reputation of pages or domains. For example, the amount of crawling  
18 between different top-level domains such as .com and .uk compared to .pl and .jp would disclose  
19 much about how Google's crawling works. The depth of the crawl, the languages of the URLs  
20 crawled, and the number of distinct sites crawled could all reveal confidential information about  
21 Google's technology. As described above, Google uses multiple methods for crawling the Web,  
22 collecting URLs, indexing them, ranking them, and providing relevant results to Google users.  
23 Google developed these methods over a number of years and at considerable expense. As far as I  
24 know, these methods are not known by Google's competitors. These methods are critical to  
25 Google's success as the world's leading search engine and popularity as the world's most-used  
26 search engine.

1           28.     While Professor Stark has not disclosed what he intends to do with Google's data or  
2 how he intends to do it, it seems logical that Professor Stark will need personal knowledge of how  
3 the URLs are collected and maintained. Given my experience in this area, I believe that to obtain  
4 that level of knowledge, Professor Stark will need to understand how Google crawls the Web and  
5 indexes information. In addition, anyone wishing to attack the randomness of the sample,  
6 Professor Stark's methodology (which he has not disclosed), or his ultimate opinions might want  
7 to probe into everything that goes into maintaining Google's URL database. For example, it might  
8 be important to defend or attack Professor Stark's methodology to know what percentage of  
9 queries are initiated within the United States, what percentage of the Web is found on Google's  
10 URL index, and how a URL is returned as a result to a search query.

11                   **VIII. GOOGLE'S MEASURES TO MAINTAIN THE CONFIDENTIALITY**  
12                   **OF ITS INFORMATION**

13           29.     I am personally aware of some of the measures Google takes to maintain the  
14 confidentiality of its query log, index, and proprietary crawling, indexing, and retrieval  
15 technology. The methods of which I am aware are described below. Additional methods are  
16 described in the Declaration of Marty Lev.

17           30.     The Google buildings that I have visited all have secured doors and entrances; I  
18 have no reason to think that any Google building that houses the information I describe above  
19 would be unsecured. Employees and visitors are required to carry appropriate security badges that  
20 security personnel routinely check. Certain entrances are restricted by access devices. There are  
21 also security guards on duty in the building.

22           31.     Google's information is compartmentalized so that only employees with a need to  
23 know have access to certain information. Thus, the information that a sales employee can access  
24 is different from the information an engineer can access. For example, sales employees do not  
25 have access to Google's source code or to certain parts of Google's technical information. Access  
26 to Google's computers is controlled by employee logins and passwords, and employee logins  
27 allow them to access only the information they need and blocks them from sensitive information



1 they do not need. Interns, independent contractors, and temporary employees require special  
2 clearance before they may access parts of Google's technical information.

3 32. Access to Google's index is restricted to engineers. I believe that most employees  
4 do not even know the specifics of where the index is stored. The index cannot be viewed in  
5 decipherable form unless one uses a source code compiled with specific Google libraries.  
6 Additionally, Google is implementing a system where its most sensitive data—such as  
7 documentation reflecting the methods it uses to measure the reputation of web pages—will be  
8 taken off line and kept in hard copy in a locked location, and an employee wanting access will  
9 have to request such access and sign out the file.

10 33. Likewise, access to Google's query log is particularly restricted to a much smaller  
11 group of employees with special clearance who need access to perform their job duties. These  
12 employees must reaffirm their need for access to the query log periodically, and anyone who does  
13 not reaffirm that need loses their access automatically.

14 34. If a Google employee does not have the correct security clearance for the query log,  
15 he or she must file a request to get it, explaining in writing his or her need to know and promising  
16 to keep the information confidential. I am aware that any abuse of Google's internal security and  
17 privacy policies is grounds for immediate termination.

18 35. I am aware that Google has refused requests for search queries in the past. For  
19 example, a professor by the name of Amanda Spink at the University of Pittsburgh, requested  
20 access to query data from Google and her request was denied. Additionally, Google generally  
21 does not share its index.

## 22 IX. ALTERNATIVE SOURCES

23 36. Both query data and web documents are available from several other sources.  
24 Metasearch engine Dogpile provides a service called "SearchSpy" to see queries done on that  
25 engine at <http://www.dogpile.com/info.dogpl/searchspy/>. Metasearch engine MetaCrawler  
26 provides a service called "Metaspy" which provides a similar list of queries at  
27 <http://www.metacrawler.com/info.metac/searchspy>. The search engines Infospace and

1 WebCrawler also provide similar information. For under \$250 per year, Wordtracker.com sells a  
2 database of over 330 million queries which is updated on a weekly basis. Ask Jeeves recently  
3 offered a service called "Ask Jeeves Take a Peek" at <http://www.ask.com/docs/peek/> which  
4 showed questions being typed at Ask Jeeves, and the page would refresh automatically twice a  
5 minute. According to her website, Prof. Spink successfully obtained query data from Excite,  
6 AltaVista, Ask Jeeves, Fast/AllTheWeb, Vivisimo, Dogpile, Metacrawler, Webcrawler, and  
7 Infospace. Prof. Spink has also written a book discussing trends and characteristics of user  
8 interaction with search engines.

9 37. An index of web documents is also available from other sources. One approach  
10 would be to crawl the public web, and several research groups have done that. In addition, an  
11 Amazon.com subsidiary, Alexa.com ("Alexa"), recently released a service that allows anyone to  
12 access Alexa crawl data. The system allows users to process over 4 billion URLs and over 1.7  
13 billion full-text documents. The system is specifically intended for "[r]esearchers who wish to  
14 tackle problems related to Web content," according to  
15 [http://pages.alexa.com/awsp/docs/WebHelp/Introduction/Who\\_Should\\_Use\\_the\\_Platform.htm](http://pages.alexa.com/awsp/docs/WebHelp/Introduction/Who_Should_Use_the_Platform.htm).  
16 Alexa.com also provides tools to search, process, and publish one's own custom subset of data. A  
17 researcher or developer could use this system to test code, including pornography filters, over the  
18 full text of documents (not just URLs) and to test code with much more than one million URLs.  
19 Attached as Exhibit B is a copy of a web log entry about Alexa which can be found at  
20 <http://battellemedia.com/archives/002116.php>, and which was printed from the internet on  
21 February 16, 2006.

22 38. I believe that Alexa offers Professor Stark a reasonable and viable option to  
23 accomplish his goals as set forth in his declaration. I believe that Alexa data would allow for more  
24 thorough testing than a sampled list of URLs from Google, as, according to Alexa, their system  
25 provides over 300 terabytes of data for researching and testing.

**X. THE DEMAND ON GOOGLE TO COMPILE THE INFORMATION REQUESTED**

1       39. Google does not index URLs in the form the Government has requested. Rather,  
2 Google maintains an index of copies of the billions of web pages that it has crawled. The index  
3 requires vast amounts of computer space and is maintained on multiple computers. There is no  
4 existing method to simply copy and hand over in electronic or paper form a selection of URLs  
5 from Google's index. The type of information sought by the Government is not created or used  
6 internally.  
7

8       40. Accordingly, an engineer would need to write a computer program capable of  
9 gathering the URLs that the Government seeks. The engineer would need to develop the code by  
10 determining a scheme to map the data logs and pull random URLs from it. The code would then  
11 need to be implemented, which involves a potentially lengthy debugging process in which the  
12 engineer tests portions of the code (and eventually the entire program) to identify and fix  
13 problems. It is common for debugging to take longer than the initial coding. Finally, the selected  
14 documents must be extracted somehow and copied in a form that can be provided to the  
15 Government. I estimate that this will take two to five full-time days of an engineer's time. This  
16 procedure could be further complicated if Professor Stark needs to participate to verify the  
17 randomness of the sample. Google does not dedicate any engineers for this type of task, and so an  
18 engineer would have to be diverted from his or her normal job responsibilities.

19       41. Similarly, Google also does not maintain search queries in the form requested by  
20 the Government. The query logs maintained by Google will have to be scrubbed of any personally  
21 identifying information. Thus, as in searching the URLs, responding to this request will require  
22 writing and implementing new code. I estimate that this will take one to three full-time days of an  
23 engineer's time.

24       42. In addition, implementing the code to search and pull random URLs from Google's  
25 index and to search and copy the search queries without identifying information will require  
26 extended hours of processing time on Google's computers. Because these are not routine functions  
27 for Google, the computers will be processing an additional and burdensome program. Running  
28

1 such programs above and beyond the normal demand on Google's computers will interfere to  
2 some unknown degree with the day-to-day operations of Google. The computers that maintain the  
3 index and search queries also process Google's spam filters and process advertising reports.  
4 Accordingly, these functions could be slowed down or completely interrupted by processing the  
5 Government's request, resulting in lower quality of service to users of Google's search engine and  
6 to Google's advertisers.

7  
8 I DECLARE UNDER PENALTY OF PERJURY that the foregoing is true and correct.

9 DATED at Mountain View, California, this 16<sup>th</sup> day of February, 2006.

10  
11  
12   
13 \_\_\_\_\_  
14 MATT CUTTS  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28