

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

UNITED STATES OF AMERICA, *et al.*,

Plaintiffs,

v.

GOOGLE LLC,

Defendant.

Case No. 1:20-cv-03010-APM

HON. AMIT P. MEHTA

**DEFENDANT GOOGLE LLC’S OPPOSITION TO DOJ PLAINTIFFS’ MOTION
TO COMPEL PRODUCTION OF DOCUMENTS AND DATA CITED IN EXPERT
REPORTS OF EDWARD A. FOX AND MARK A. ISRAEL**

DOJ Plaintiffs’ (“Plaintiffs”) August 3, 2022 Motion to Compel Production of Documents and Data is the latest in a series of efforts to reopen the 16-month fact-discovery period that followed a year-plus pre-Complaint investigation. For both issues raised by the motion, Google has exceeded its disclosure obligations under the Case Management Order and the Federal Rules, and Plaintiffs face no obstacle in evaluating and testing the experts’ opinions. Instead of proceeding with expert discovery, Plaintiffs rely on overheated rhetoric and inaccurate facts to seek privileged documents, even more deposition testimony from already-deposed fact witnesses, and data that was not requested during fact discovery and not required to be produced under this Court’s orders.

First, Plaintiffs contend that references in the expert report of Professor Edward Fox to an experiment “conducted at the request of European lawyers” waived the attorney-client privilege with respect to other documents relating to that experiment. ECF No. 376-1 (Pls.’ Mem. ISO Pls.’ Mot. to Compel Prod. of Docs. and Data Cited in Expert Reports of Edward A. Fox and Mark A. Israel (“Pls.’ Mem.”)) at 2. Plaintiffs are mistaken. There can only be a

potential subject-matter waiver if a party uses a privileged document. Plaintiffs' analysis fails at this first step. The documents they cite were never claimed by Google to be privileged.

Moreover, Google's expert did not rely on the European experiment that is the subject of Plaintiffs' motion. Dr. Fox's report explains that he "had no involvement with the experiment" in question, ECF No. 369-1 at 45, and instead conducted his own experiment, which was the "singular focus" of his report, ECF No. 370 at 2.

Plaintiffs' speculation that Dr. Fox reviewed "cherry-picked documents" relating to the European experiment also is false. *See* Pls.' Mem. at 8. Instead, he reviewed two non-privileged documents reflecting the results of that experiment. Neither was withheld as privileged at any point, and both have been produced to Plaintiffs in expert discovery. And contrary to Plaintiffs' overheated rhetoric, there has been no "gamesmanship of discovery and privilege." *See id.* During fact discovery, a handful of privileged documents relating to the European experiment were retrieved by the parties' agreed-upon search parameters, correctly identified as privileged, and included on Google's privilege log. Dr. Fox did not review those documents or any other privileged document. There is nothing unusual, let alone improper, about an expert reviewing non-privileged documents that relate to the same subject as other documents that are properly deemed privileged. Simply put, there was no privilege waiver, and there is no basis for Plaintiffs' request for additional documents and deposition testimony.

Second, Plaintiffs seek to compel production of certain "raw Google Panels data" that Google's expert did not access, let alone rely upon, and therefore is not subject to disclosure under the CMO or any other applicable authority. *See id.* at 9. Plaintiffs' request purportedly arises from a portion of the expert report of Dr. Mark Israel, which analyzes certain aggregated data regarding the web browsing and searching habits of a panel of paid volunteers convened by

Google in the ordinary course of business. Dr. Israel and his support staff only had access to aggregated data, which reflected how the anonymous volunteers as a group generally used Google and other major websites. Dr. Israel did not access private data relating to individual panelists or their search queries because his analysis relied on aggregated data and not the behavior of any individual panelist, and because it raises privacy concerns that are resolved through aggregation.

On June 13, Dr. Israel timely disclosed, in accordance with the CMO, the aggregated panels data that he used in his analysis, along with the “data and programs underlying the calculations” that he performed using the panels data and that appear in his report. ECF No. 108-1 ¶ 24. Accordingly, notwithstanding the framing of their motion, Plaintiffs are not seeking the data “underlying the calculations” in Dr. Israel’s report, as that data has already been produced. They are instead seeking to compel production of all of the more granular data that relates to or underlies the “data and programs underlying the calculations” in the report—data that Dr. Israel has never seen nor used in forming his opinions. Plaintiffs’ request is unsupported by the text of the CMO and inconsistent with the conduct of their own experts, who have relied upon numerous forms of aggregated data produced by Google and third parties, without ever suggesting that the data is somehow incomplete or unreliable unless it is also accompanied by the most granular available form of underlying data. For this reason, Plaintiffs’ motion should be denied as it pertains to this data.

In addition, Google respectfully submits that it will not be necessary to reach this issue, because the motion is now moot in light of subsequent productions made by Google. In order to avoid any possible disruption to the case schedule or further burden the Court, Google has worked on a process for collecting the panel data at issue since Plaintiffs first requested it on July

19. On August 15, Google produced the requested data after implementing privacy measures that are consistent with those that Plaintiffs agreed to last year in the context of a different request for query-related data. Plaintiffs therefore now have all of the data that they reasonably could have received if they had made this request during fact discovery (which they did not do, despite conducting other discovery regarding Google panels).

I. THERE HAS BEEN NO WAIVER WITH RESPECT TO PRIVILEGED DOCUMENTS RELATING TO THE EUROPEAN DATA EXPERIMENT

Plaintiffs seek an order compelling production of what they agree are privileged documents, as well as an order permitting deposition questioning of fact witnesses beyond the limits set at the last status conference. The gravamen of Plaintiffs’ motion—*i.e.*, that Google waived attorney-client privilege when it served an expert report that referred to the output of a European data study—is flawed. As described below, a so-called “selective” waiver requires the disclosure of *privileged* material. No such disclosure has occurred here. Plaintiffs’ motion should therefore be denied.

A. Background

On June 6, Google served the expert report of Professor Edward Fox, who has nearly four decades of experience as a professor of computer science and has received a number of awards relating to his work in information retrieval. As the Court is aware from the most recent status conference, Dr. Fox performed a “data reduction” experiment—that is, an experiment to assess what would happen to Google’s search quality if Google used less data to train its ranking algorithms. Dr. Fox found that even a very large decrease in training data would have only a modest effect on Google’s search quality. As a result, he “conclude[d] that the vast majority of the quality ‘gap’ between Google and Microsoft’s search results must be explained by factors other than the amount of available click-and-query data.” ECF No. 369-1 at 7.

Dr. Fox's report also made passing reference to certain *other* data-reduction experiments that Dr. Fox became aware of while designing and implementing his experiment. *See id.* §§ 5.1.1–5.1.3. One of those other experiments forms the basis of Plaintiffs' present motion. The experiment in question was performed "at the request of European lawyers," and "was completed before [Dr. Fox] started [his report]." *Id.* at 45. As Dr. Fox stated in his report, he "had no involvement with" the European data-reduction experiment that is the subject of Plaintiffs' motion. *Id.* And, as Google has confirmed to Plaintiffs on more than one occasion, "Professor Fox's opinions do not rely upon the European data reduction experiment ... or underlying data associated with that experiment." ECF No. 376-4 (July 25, 2022 Ltr. from F. Rubinstein) at 2.

Nearly a month after receiving Dr. Fox's report, on July 1, Plaintiffs sent Google a letter claiming that Dr. Fox's report had waived privilege over the subject matter of the European data-reduction experiment. Although Google explained why Plaintiffs' assertion was inaccurate in a July 5 letter, it agreed to identify to Plaintiffs documents relating to the European data-reduction experiment that had been withheld from production as privileged. There are four such documents, none of which Dr. Fox has ever seen or is even aware. Google has maintained its privilege claim over those four documents and has identified for Plaintiffs the four pertinent privilege log entries.

At the same time, Plaintiffs asked questions regarding two figures set forth in Dr. Fox's report that summarize the results of the European data-reduction experiment. To address Plaintiffs' questions, Google produced the two documents setting forth the data reproduced in those figures in Dr. Fox's report. *See* ECF No. 376-4 at 2; ECF No. 376-5 (July 29, 2022 Ltr. from F. Rubinstein) at 1–2.

Google has never asserted privilege over the two documents that it produced on July 25 and 29 in response to Plaintiffs' requests, and they have never appeared on any of Google's privilege logs. Those particular documents were not produced during fact discovery because they were not retrieved pursuant to the agreed-upon search parameters. Nor were they produced in the original expert discovery because the CMO provides that in this case an expert need only disclose "documents relied upon by the expert in forming any opinions in his or her report." ECF No. 108-1 ¶ 24(a). As discussed above, Dr. Fox did not rely upon these two documents—or any other documents or information relating to the European data-reduction experiment—in forming any of the opinions set forth in his report. Nevertheless, in the interest of resolving Plaintiffs' complaint without burdening the Court, Google produced the non-privileged documents setting forth the results of the European data-reduction experiment referenced in Dr. Fox's report. Plaintiffs have now received more than they are entitled to under the CMO or any other applicable authority.

B. Legal Argument

The Court should deny Plaintiffs' motion because it rests on the faulty premise that Google "selectively withdr[e]w" a claim of privilege relating to the European data-reduction experiment or "strategically suppl[ied] its expert with information regarding this experiment, while asserting privilege over and denying Plaintiffs access to all other documents and information." Pls.' Mem. at 5. Google did no such thing.

Federal Rule of Evidence 502—which governs the waiver of attorney-client privilege—applies only upon the "disclosure of a communication or information *covered by the attorney-client privilege or work-product protection.*" Fed. R. Evid. 502 (emphasis added). In other words, "[i]n order for disclosure of certain material to effect a waiver of attorney-client privilege, the material disclosed must itself be privileged." *EEOC v. Geo. Wash. Univ.*, 502 F. Supp. 3d

62, 87 (D.D.C. 2020). If it is not, no waiver results. Here, that threshold requirement has gone unmet. Nothing in the text of Dr. Fox's report is privileged, nor are the documents that Dr. Fox reviewed and Google produced in response to Plaintiffs' concerns.

Specifically, the two non-privileged documents that Dr. Fox reviewed contained the final results and summary of the European data reduction experiment referenced in his report. Unlike the privileged documents that were properly withheld, the materials that were reviewed by Dr. Fox (and have been produced to Plaintiffs) were created with the understanding that they may be disclosed to the European Commission or certain other authorities. There is nothing anomalous about the fact that the attorney-client privilege attaches to some, but not all, communications regarding a given subject matter.¹

The cases Plaintiffs cite do not support the notion that production of non-privileged documents can cause a waiver, as none appears to concern the disclosure of such documents. *See In re Sealed Case*, 676 F.2d 793, 824 (D.C. Cir. 1982) (disclosed materials included privileged report prepared by law firm); *US Airline Pilots Ass'n v. Pension Ben. Guar. Corp.*, 274 F.R.D. 28, 30 (D.D.C. 2011) (disclosed material summarizing "legal conclusions"); *In re Papst Licensing GMBH & Co., KG Litig.*, 250 F.R.D. 55, 58 (D.D.C. 2008) (non-movant

¹ For example, inventors frequently engage in privileged exchanges with an attorney about which technical data to include in a patent application, even though the final version of the application is not privileged. *See, e.g., In re Spalding Sports Worldwide, Inc.*, 203 F.3d 800, 806 (Fed. Cir. 2000). Documents created in the course of preparing securities filings or other periodic reports likewise may be privileged, even though the final report intended for submission to a regulator typically is not. *See, e.g., Roth v. Aon Corp.*, 254 F.R.D. 538, 541 (N.D. Ill. 2009); *Carey-Canada, Inc. v. Cal. Union Ins.*, 118 F.R.D. 242, 247–48 (D.D.C. 1986). For much the same reason, materials prepared at the request of attorneys conducting an internal investigation or drafting a white paper may be privileged, even if some of those materials are later included in a final, non-privileged summary intended for disclosure to a third party. *See, e.g., In re Linerboard Antitrust Litig.*, 237 F.R.D. 373, 388–89 (E.D. Pa. 2006); *In re Air Crash Disaster at Sioux City, Iowa on July 19, 1989*, 133 F.R.D. 515, 518 (N.D. Ill. 1990).

acknowledged that numerous privileged documents had been and would be produced, leading to a consideration of whether additional privileged documents also would need to be produced); *Minebea Co. v. Papst*, 2005 WL 8179732, at *2 (D.D.C. June 23, 2005) (expert witness had been “provid[ed] . . . with privileged documents”). By contrast, in this case Plaintiffs seek an order compelling the production of privileged materials based on the disclosure of materials that Google has never withheld as privileged. That is not how Rule 502 works, and it is “flat out wrong” to suggest that “the disclosure of non-privileged information” acts “as a waiver . . . of privileged information,” even if it “deals with the same subject matter.” *Trs. of Elec. Workers Loc. No. 26 Pension Tr. Fund v. Tr. Fund Advisors, Inc.*, 266 F.R.D. 1, 11 (D.D.C. 2010) (rejecting movants’ assertion that non-movants were “required to disclose privileged documents because they have previously disclosed other non-privileged documents that deal with the same subject matter”).

Plaintiffs insinuate that privilege must attach to at least one, if not both, of the documents that Google produced in response to Plaintiffs’ questions. *See* Pls.’ Mem. at 6–7. In particular, Plaintiffs assert that “[b]y Google’s own telling,” the European data-reduction experiment “was done at the direction of counsel and is, therefore, privileged.” *Id.* at 6. That is a non-sequitur. There are myriad examples of processes that are directed by counsel and involve a significant number of privileged communications but that lead to a non-privileged output. *See supra* note 1. Plaintiffs do not suggest that any of those circumstances would result in a privilege waiver over every privileged communication related to those processes, and they offer no reasonable basis for concluding that a waiver has occurred here. Consistent with this, neither document ever appeared on Google’s privilege log, and Plaintiffs never explain why either (or both) would be privileged.

Plaintiffs further contend that the current state of affairs “is untenable because Google asserts that every other document relating to [the European data-reduction] experiment is withheld on claims of privilege.” Pls.’ Mem. at 7 (italics omitted). There is nothing unusual, however, about an expert reviewing a final, non-privileged document that summarizes work performed at the direction of counsel. For example, experts frequently review or rely upon another expert’s final report or a pleading filed with a court. That review does not trigger a waiver of the privilege that otherwise applies to underlying materials developed to create those documents.

In any event, the corpus of relevant documents at issue is much smaller than Plaintiffs’ argument implies. As discussed, Dr. Fox reviewed two documents reflecting the output of the experiment (which Google has already produced), and a reasonable search of the millions of documents collected and reviewed pursuant to the parameters negotiated by the parties during fact discovery turned up only four additional documents related to the experiment that were properly deemed privileged and logged.²

As for the four withheld documents, there is no dispute that they were properly withheld for privilege in the first instance. Three of the documents reflect a request for or the provision of legal advice in connection with regulatory submissions. That makes them privileged because “[d]rafts of documents that are prepared with the assistance of counsel for release to a third party are protected under attorney-client privilege.” *Alexander v. FBI*, 198 F.R.D. 306, 312 (D.D.C.

² Google notes that there is a typographical error in Dr. Fox’s report. The European data-reduction experiment occurred in 2021, not 2020. A relatively small number of documents relating to the experiment were captured by the “refresh” collections that the parties negotiated in late 2021 and early 2022, but the experiment had not occurred when the bulk of the document collections occurred in 2020 (in response to Plaintiffs’ CIDs) and early 2021 (in response to Plaintiffs’ RFPs).

2000). The remaining document contains deliberation regarding the design of a technical analysis meant to facilitate the provision of legal advice. It, too, is privileged because it reflects the transmission of information in order to fulfill a request from counsel needed to provide legal advice. *See, e.g., FTC v. Boehringer Ingelheim Pharms., Inc.*, 180 F. Supp. 3d 1, 31, 33 (D.D.C. 2016) (finding that largely factual “PowerPoint presentations, charts, graphs, and tables” were privileged where “one of their primary purposes was to enable [the company’s] counsel to advise it on how to settle” pending litigation), *aff’d*, 892 F.3d 1264 (D.C. Cir. 2018). Thus, even had Plaintiffs challenged the applicability of attorney-client privilege to these four documents in the first instance—and, again, they did not—that argument could not succeed. For the foregoing reasons, Plaintiffs’ motion should be denied.

Plaintiffs’ request to expand the scope of the re-opened depositions of Prabhakar Raghavan, Philipp Schindler, and Hal Varian, *see* Pls.’ Mem. 8, fails for all the reasons discussed above, as well as additional ones. Each has already been deposed for two days during fact discovery, and the prior rationale for reopening their depositions (documents deprivileged after their depositions) has nothing to do with the contents of Dr. Fox’s report. If Plaintiffs believed that Dr. Fox’s report in some way warranted reopening fact witnesses’ deposition, the time to say so would have been the July 14 status conference, when the Court resolved the parties’ disputes regarding the scope of any reopened depositions. Although Plaintiffs alluded at that hearing to an “emerging issue regarding a subject-matter waiver,” July 14, 2022 Hr’g Tr. at 4, they in no way suggested that it should affect the scope of deposition questioning or the timing of any depositions. To the contrary, Plaintiffs assured the Court that the reopened depositions would not “go past the documents and obvious questions that follow the documents on the related subject matters” following Google’s production of certain deprivileged documents. *Id.* at 30.

The Court should reject Plaintiffs' effort to retroactively broaden their request to reopen the depositions of Mr. Raghavan, Mr. Schindler, and Mr. Varian.³

II. GOOGLE HAS EXCEEDED ITS PRODUCTION OBLIGATIONS WITH RESPECT TO THE PANELS DATA CITED IN DR. ISRAEL'S REPORT

The second half of Plaintiffs' motion seeks an order compelling production of certain "raw Google Panels data" and a corresponding "data dictionary." Pls.' Mem. at 9. Although the motion ostensibly relates to the expert report of Dr. Mark Israel and borrows certain phrases that appear in the CMO, the relief that Plaintiffs seek is untethered to the terms of the CMO or any other disclosure requirement. As described below, Dr. Israel timely produced all of the data that he and his support staff relied upon together with all of the other materials specified in the CMO, including "all programs and codes necessary to recreate the calculations" in the report "from the initial ('raw') data files" that Dr. Israel and his support staff accessed. ECF No. 108-1 ¶ 24(e). Through the motion to compel, Plaintiffs seek discovery of different data that they did not request during fact discovery and that was not seen by or made available to Dr. Israel because it was not part of his analysis or his opinions. There is no basis for Plaintiffs' request and, to the extent that the Court considers it on the merits, Google respectfully submits that it should be denied.

Before the Court even reaches that question, however, Google respectfully submits that the Court should deny the motion as moot. Although Plaintiffs are not entitled to any of the additional discovery they seek, in the interest of ensuring that Plaintiffs' unnecessary request

³ Plaintiffs' requested relief is also too broad insofar as it seeks all previously withheld and/or redacted documents linked to the four documents identified on the privilege log. Pls.' Mem. at 2. Again, Plaintiffs have not challenged the claim of privilege with respect to any of those documents, and they have not offered any basis for requiring production of documents that do not address the "subject matter" of the purported "waiver," which Plaintiffs acknowledge is limited to the European data-reduction experiment. *See, e.g., id.* at 5.

does not interfere with the case schedule or further burden the Court, Google has worked expeditiously to collect the data at issue since Plaintiffs first requested it on July 19. Before Plaintiffs filed their motion, Google agreed that it would produce the requested data in a form that would obviate the need for a time-intensive process of hashing or redacting certain search queries to protect user privacy. Google completed that production on August 5. After Plaintiffs filed the instant motion, Google separately began the work of hashing and redacting the unproduced search queries in a manner consistent with the protocol that the parties agreed to last year in connection with the production of certain other user query-related data. Google produced that data on August 15, along with the data dictionary requested by Plaintiffs' motion. With that, Google submits that the motion can be denied as moot.⁴

A. Background

Google, like many other companies, convenes “panels” of volunteers who agree to share certain information about how they interact with technology and media over time in exchange for compensation. Google panels are similar in some respects to the well-known Nielsen panels, which include a network of households that have agreed to install a device that measures which television programs they watch. In Google's case, the volunteer panelists may be asked to share information that relates to how they use smartphone applications and websites, including search engines. Information about Google panels is available online,⁵ and Plaintiffs obtained additional

⁴ There is no reason for Plaintiffs to reject the same privacy safeguards that were previously applied, particularly given that Google's own expert did not review the user-level or query-level data that Plaintiffs have since requested. If Plaintiffs have any issue with the previously-agreed privacy safeguards, Google respectfully requests an opportunity to submit a supplemental brief addressing whatever objection they may raise.

⁵ See, e.g., *About Google and Panel Research* (<https://www.google.com/landing/panelresearch/>); *Google Panel Privacy Policy* (<https://support.google.com/screenwise-meter/answer/9744317?hl=en>); *Ipsos Screenwise Panel for Google* (<https://screenwisepanel.com/>).

information about the panels during fact discovery, including through documents and deposition testimony.⁶

The expert report of Dr. Mark Israel, served on June 6, uses certain aggregated data from a Google panel to analyze which websites consumers use to search for information regarding flights, hotels, shopping, automobiles, and banking. According to the data that Dr. Israel analyzed and relied upon, a majority of the panel searching for those kinds of information alternated between searching on general search engines and specialized vertical providers, while a smaller number of searchers only used sites that fell into one of those two categories. Dr. Israel also considered aggregated data from the panel to evaluate the frequency with which users visit certain other sites or mobile applications popular with advertisers, such as TikTok and Instagram, within a given period of time. He used this aggregate data as part of an analysis of the options available to advertisers to reach consumers who search on Google.

The Google panels data reviewed by Dr. Israel and the analysts working at his direction was aggregated. Dr. Israel and his team did not look at whether one particular individual panel participant navigated to Amazon's website or searched for "running shoes" on Google. This user-level or query-level data was not part of Dr. Israel's analysis because he was analyzing how a statistically significant group of consumers used websites and mobile applications overall, as opposed to analyzing the individual searches of any particular volunteer panelist at any particular time. Because there was no need for Dr. Israel or his team to review the data in its most granular form, Google compiled the particular data relevant to the analysis, and Dr. Israel used only that aggregated data in performing the calculations in his report. On June 13, *i.e.* the date specified in

⁶ See, e.g., Jan. 28, 2022 Rule 30(b)(6) Dep. Tr. at 18:16–21, 63:14–64:19 (describing the Google Panels program and citing examples of panels that have been convened to study volunteers' use of Google Search and other websites); Mar. 23, 2022 Dep. Tr. at 40:16–41:15.

the CMO, Dr. Israel produced to Plaintiffs all of the Google panels data files that Dr. Israel and his team relied upon. ECF No. 108-1 ¶ 24. On that same day, he also produced all of the programs and codes necessary to recreate the calculations appearing in his report from the data files he had been provided. *See id.* ¶ 24(e).

There is nothing unusual about using and producing data in an aggregated form, rather than at the most granular level available. For example, financial data relating to revenue and expenses was produced throughout discovery in monthly or quarterly increments, rather than at the level of individual ledger entries reflecting each credit and debit. The same is true of much of the search-related data analyzed by Google in the ordinary course of business or produced in this case. In this instance and many others, the individual search query or other user-specific information is irrelevant to the analysis at hand.

On July 6, Plaintiffs noted that Dr. Israel's backup data relating to the panel analysis was in aggregated form, and asked for confirmation that the data was provided to Dr. Israel in the same aggregated form. Google confirmed that it was, and further confirmed that "there were no materials used to interpret the data other than what [was] provided in the report itself."

On July 19, *i.e.*, more than five weeks after Plaintiffs received all of the back-up materials for Dr. Israel's report, Plaintiffs asserted for the first time that Google should collect and produce, among other things, "[d]ata dictionaries defining each column in the files" produced with Dr. Israel's report as well as "[c]omplete Google Panels Datasets." Although Plaintiffs did not detail what they envisioned receiving in response to that request or how it could possibly be justified at this stage of the case, Google nevertheless provided three additional categories of material, all of which were beyond the scope of what is required by the CMO.

First, at Plaintiffs' request, Dr. Israel's team prepared a data dictionary further detailing the meaning of the fields in the aggregated panels data that Dr. Israel had used as the basis for the calculations in his report. Even though Plaintiffs had not asked any specific questions about how to interpret the data or identified any ambiguities in the code produced along with the data, Dr. Israel's team provided the requested data dictionary in order to ensure a shared understanding of the terminology.

Second, Google produced the computer code that had been used to extract the panels data provided to Dr. Israel and his team. This code differs from the code referenced in Paragraph 24(e) of the CMO, which requires production of "programs and codes necessary to recreate the calculations from the initial ('raw') data files." Dr. Israel had already produced on June 13 the code that he used to analyze the panels data and to perform the calculations reflected in his report. The code that Google subsequently produced at Plaintiffs' request was used to perform the back-end computational task of compiling the "initial ('raw') data files" that Dr. Israel considered and relied upon in conducting his analysis.

Third, Google confirmed on July 29 that it would collect and produce the user-level panels data reflecting, among other things, the webpages visited by a panel participant, the category of search conducted on Google (*e.g.*, shopping or hotels), and the corresponding time stamp for each such site visit or search. As indicated, this data had not previously been collected because it was not requested at any point prior to July 19, and neither Dr. Israel nor anyone working at his direction had ever viewed it, let alone relied upon it. Google completed that production on August 5.

In the meantime, Plaintiffs filed the instant motion seeking "(1) the raw underlying Google Panel data, including all query data as well as site visit data, and (2) a corresponding data

dictionary.” Pls.’ Mem. at 12. As Plaintiffs know from having requested different types of user-level or query-level data during fact discovery, there are a number of privacy-related issues that must be addressed before production can occur. Instead of explaining the relevance of the query data sought or discussing a privacy-preserving solution, Plaintiffs filed the instant motion. Nevertheless, Google proceeded with collecting the requested data and implementing privacy measures similar to those that Plaintiffs agreed to last year in the context of a different request directed to certain query data. Google completed the production of that data, along with the requested data dictionary, on August 15.

B. Legal Argument

Google satisfied its disclosure obligations relating to Dr. Israel’s opening report on June 13, when it produced all of the materials described in Paragraph 24 of the CMO. Plaintiffs’ argument to the contrary is predicated on a misinterpretation of Paragraph 24(e), which provides that “for all calculations appearing in the [expert’s] report,” the parties will disclose “all data and programs underlying the calculations, including all programs and codes necessary to recreate the calculations from the initial (‘raw’) data files, and including the intermediate working-data files that are generated from the raw data files and used in performing the calculations appearing in the report.” In this instance, as in others throughout the parties’ reports, “the initial (‘raw’) data files” consist of data in an aggregated form that was previously extracted from other data. Neither the CMO nor any other authority imposes an obligation to take the “initial (‘raw’) data files” used by the expert, trace them back to the data set(s) where they originated, and then produce those data sets in their entirety.

The starting point of the analysis performed by Dr. Israel and his support staff was an aggregated record of many panelists’ visits to particular websites. Dr. Israel and the team working at his direction wrote code to perform certain calculations using that “initial (‘raw’) data

file” and included those calculations in a pair of tables in his report. Dr. Israel then produced on June 13 the data that he used as the starting point for his analysis together with the code used to perform the calculations using that data. That is precisely what the CMO requires in calling for the “data and programs underlying the calculations” that appear in the expert report. Plaintiffs seek to append a new requirement by seeking not only data “underlying the calculations” in the report, but *also all other data underlying that data*.

It is noteworthy that Plaintiffs have not applied this additional requirement to the data relied upon by their own experts, even though much of the data produced in this case is in aggregated form. For example, some third parties have produced data reflecting the number of unique visitors to a web site during a given month or the number of times a mobile application was downloaded by users in the U.S. in a given year. Plaintiffs have not suggested that this form of aggregated data is “curated and altered” because it is not accompanied by the web server logs that contain the raw data reflecting each visit to a website. *See* Pls. Mem. at 9. And they have not suggested that this aggregated data was subject to some form of “manipulation” because it was not accompanied by the user’s IP address and any other identifiers recorded by the web server. *See* Pls.’ Mem. at 11. There is no basis for Plaintiffs’ suggestion that the CMO requires production of the most granular data that is potentially available. That is not what the text of Paragraph 24 requires, and it is not consistent with ordinary practice in this case or others.

Instead of explaining why they want the query-level data at issue, Plaintiffs attack Google’s motives in a way that is both baseless and improper. Although no explanation should be necessary, the data collection process that led to Plaintiffs’ motion is entirely unremarkable, and a common part of a litigation that features data-driven expert testimony. The analysis that Dr. Israel performed generally involved a consideration of which websites a large group of

volunteer panelists collectively visited, and the category of information the panelists collectively were searching for on Google (*e.g.*, flights or hotels). The specific query that a particular panelist entered was not part of the analysis. There was no need for Dr. Israel or his team to access query-level data, which Google maintains with the utmost regard for user privacy, including in the context of a panel study where participants have expressly volunteered to share the information with Google. Dr. Israel and his team therefore received an aggregated data set that included the information pertinent to their analysis, but none of the other data that was both beyond the scope of their analysis and the subject of additional privacy considerations.

To resolve the privacy issues with this sort of data, the parties agreed more than a year ago: “[I]n the course of this litigation and the investigation preceding it, Google’s litigation and investigation teams (including any experts) have not used, and will not use, any data substantially similar to the data that Google refuses to produce to Plaintiffs. This includes unhashed queries, whether for the days Plaintiffs have requested or for other times, to the extent Google hashes or redacts those queries (or substantially similar queries) in the data provided to Plaintiffs.” Google and its experts followed that agreement here.

Accordingly, if Plaintiffs maintain that the August 15 production does not moot their motion, the motion should be denied because the relief sought is not supported by the CMO.

Dated: August 15, 2022

Respectfully submitted,

WILLIAMS & CONNOLLY LLP

By: /s/ John E. Schmidlein
John E. Schmidlein (D.C. Bar No. 441261)
Benjamin M. Greenblum (D.C. Bar No. 979786)
Colette T. Connor (D.C. Bar No. 991533)
680 Maine Avenue, SW
Washington, DC 20024
Tel: 202-434-5000
jschmidlein@wc.com
bgreenblum@wc.com
cconnor@wc.com

WILSON SONSINI GOODRICH & ROSATI P.C.
Susan A. Creighton (D.C. Bar No. 978486)
Franklin M. Rubinstein (D.C. Bar No. 476674)
1700 K Street, NW
Washington, DC 20006
Tel: 202-973-8800
screighton@wsgr.com
frubinstein@wsgr.com

ROPES & GRAY LLP
Mark S. Popofsky (D.C. Bar No. 454213)
2099 Pennsylvania Avenue, NW
Washington, DC 20006
Tel: 202-508-4624
Mark.Popofsky@ropesgray.com

Counsel for Defendant Google LLC