

**IN THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF COLUMBIA**

United States of America, *et al.*,

Plaintiffs,

v.

Google LLC,

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

HON. AMIT P. MEHTA

**SEALED VERSION – CONTAINS  
CONFIDENTIAL INFORMATION**

State of Colorado, *et al.*,

Plaintiffs,

v.

Google LLC,

Case No. 1:20-cv-03715-APM

HON. AMIT P. MEHTA

**SEALED VERSION – CONTAINS  
CONFIDENTIAL INFORMATION**

**JOINT STATUS REPORT OF GOOGLE AND MICROSOFT**

Pursuant to the Court’s request for an “update[,]” Minute Order, Aug. 8, 2021, Google and Microsoft Corporation (“Microsoft”) (together with Google, the “Companies”) can report they have narrowed their dispute concerning Microsoft’s collection of documents in response to Google’s April 13 and July 19 subpoenas. On several outstanding searches, and on the timing of substantial completion, however, there is an impasse. Each side’s respective positions on these disputes are set out below.

**I. Google’s Position Statement**

Microsoft is at the heart of this case; there is no other third party that is more deeply enmeshed across all of the Plaintiffs' allegations, nor one whose business units stretch as broadly across the various issues raised in Plaintiffs' Complaints. There is no debate about the relevance of the documents that Google is seeking. The only question is burden, and Microsoft by its conduct has obscured all manner of data that would help answer that question; for example, four months into the subpoena compliance process, Microsoft has provided no firm figures reflecting the duplication within or across the disputed search strings, or across all of the search strings to be applied. A burden objection is Microsoft's to prove, and it has failed to do so. After bending over backwards to compromise—in particular, substantially narrowing the disputed search strings to reduce the volume of documents for Microsoft to review and produce—Google now asks the Court to compel implementation of the eight outstanding search strings proposed by Google, and to require substantial completion of Microsoft's document production by September 30, 2021.

**A. Microsoft's pace of production continues to lag.**

As of the last hearing, Microsoft had produced only about one thousand documents identified as responsive to Google's subpoena. *See* Hr'g Tr. 58 (July 30, 2021). Only a small number of those were custodial documents (*e.g.*, internal emails); most of the documents came instead from non-custodial, "off-the-shelf" sources.

Despite the Court's instruction that it accelerate the production process, Microsoft has proceeded at the same pace since the July status conference. Google received another 4,800 custodial documents on August 5. Notably, none came from the search strings Google had proposed, but instead were identified using search terms Microsoft had negotiated with the

Plaintiffs. Despite months of negotiation, Microsoft essentially ignored the proposed search terms that Google first provided back in May.

Earlier this week Microsoft made another, extremely modest production of 460 more documents. Despite its professed “objective to produce documents as quickly and efficiently as possible,” ECF 172 at 2, this smattering of documents also came from search strings Microsoft had crafted on its own, without input from Google. It was not until yesterday, on the eve of this filing, that Microsoft produced several thousand additional documents, in an apparent effort to appear to have “caught up” ahead of the August 31 hearing; again, however, some of the volume appears to have been driven by documents requested by Plaintiffs or that are publicly available.

Most significantly for purposes of the present dispute, Microsoft still has not made any meaningful effort of the sort Rule 45 requires to substantiate its assertion that the searches Google is requesting impose an undue burden. Months have thus ticked off the discovery schedule with the most important source of third-party documents continuing to stiff-arm Google’s attempts to obtain reasonable compliance with its discovery requests.

**B. Microsoft’s burden objections are unsubstantiated.**

Google and Microsoft have negotiated agreement on 36 search strings from the set that Google proposed on May 27 but remain at loggerheads over the remaining eight strings.<sup>1</sup> Microsoft has not made (nor could it make) any relevance objection to any of these strings. It complains only that the search strings would capture too many documents for it to review and produce.

---

<sup>1</sup> Only one of the 44 search strings derives from the supplemental subpoena that Google served on July 19, 2021 (containing one RFP), relating to a very recently announced agreement between Microsoft and Amazon.

Because Microsoft unilaterally chose to provide a very limited amount and form of deconstructed hit data—the standard means for isolating the particular terms in a given string that are driving hit volume<sup>2</sup>—and because it was willing to do so only for (subsets of) two of the eight disputed search strings, Google has been forced to rely on manually drafted multiple “permutations” of each of its proposed strings in order to attempt, by trial and error, to reduce hit volumes and thereby address Microsoft’s burden objections. ECF 172 at 2. For each of the eight disputed strings, Google then proposed permutations that result in substantially fewer documents than its initial proposal had yielded. But Microsoft has drawn a line in the sand at reviewing the volumes generated even by these compromise proposals.

The eight strings are addressed below. But as an initial matter, the lines that Microsoft is drawing—and its insistence on redrawing search strings so as to reduce the volume of likely responsive documents to be reviewed—are artificial. Several of the disputed strings address more than one request for production; and all of the disputed strings were designed to cover multiple subjects in one shot, and so it is unsurprising that they will have higher volumes.

More importantly, the hit counts that Microsoft *has* reported are substantially inflated because they fail to account for the predictably high level of document duplication—*i.e.*, the fact that many “hits” are the same document found in the files of different custodians or the very same document that hits on multiple search strings. Microsoft has almost universally declined to identify the number of “deduplicated” hits within the documents found from a particular search string, *i.e.*, intra-string deduplication—but the few data points Microsoft has provided suggest

---

<sup>2</sup> By contrast, pursuant to this Court’s August 19 Minute Order, Apple provided Plaintiffs with several deconstructed hit reports, after which Apple and Plaintiffs resolved those disputes. *See* ECF 186 at 2.

intra-string deduplication rates in Microsoft's documents as high as 37%.<sup>3</sup> Thus, the actual number of unique, non-duplicative documents captured by each disputed search string is likely substantially smaller.

Nor has Microsoft provided deduplication figures *between* strings—*i.e.*, inter-string deduplication—an unknown but ascertainable fact given the certainty that the same documents will be retrieved by two or more search strings. Without knowing how much duplication occurs across the search strings, there is no way to judge whether the overall search Google is requesting is burdensome or not. Indeed, a controlled experiment Google conducted to reverse-engineer an approximation of the likely extent of inter-string duplication—using a hit report Microsoft agreed to provide combining several otherwise separate search strings—showed that inter-string duplication is likely about 25%.<sup>4</sup> One can extrapolate from this estimate that the

---

<sup>3</sup> On August 21, Microsoft informed Google that, with respect to a search string that Google proposed for RFP No. 17, the parent-level hit count was 6,968 when deduplicated between ten custodians. Microsoft had, three days earlier, provided Google with a parent-level hit count for six of the ten custodians individually and not deduplicated. Because these six custodians, in total, have 11,067 parent-level hits, the total non-deduplicated parent-level hit count for the ten custodians is necessarily higher than 11,067. Therefore, for this search string, the intra-string deduplication rate is *at least*  $(11,067 - 6,968) / 11,067 = 37\%$ .

<sup>4</sup> On August 16, as part of the permutation testing process, *see* ECF 172, Google provided Microsoft with (1) a list of 25 search strings that correspond to various different RFPs in Google's subpoena, and (2) a consolidated search string that connects with "OR" operators the aforementioned 25 search strings. While a document that is responsive to more than one search string would be counted more than once in the first scenario, it would be counted only once in the second scenario. In the absence of Microsoft's providing information on the inter-string deduplication rate, Google had to conduct this controlled experiment to estimate that rate. Hit counts provided by Microsoft for these search strings reveal that the total parent-level hit count in the first scenario is 222,890, while the total parent-level hit count in the second scenario is 165,460. Therefore, the inter-string deduplication rate (which does not include deduplication between custodians) is estimated to be  $(222,890 - 165,460) / 222,890 = 25.8\%$ .

overall duplication in the subject Microsoft documents is as high as 50%.<sup>5</sup> To put it another way, Google estimates that the actual volume of documents Microsoft would have to review and produce is substantially less than the volume Microsoft is complaining about.

Only Microsoft can determine the actual duplication number or estimate it with further precision. And while Microsoft represented that it would at least “provide to Google regular updates as to the actual de-duplication of processed documents,” ECF 172 at 2, it has not provided any actual deduplicated counts for *any* of the disputed strings. Microsoft instead has claimed—based on its asserted experience with other unspecified litigation and with no data provided to support it—that the duplicates would likely make up only 30% of the “hit” volume. *Id.* But the data points discernible from its disclosures to date confirm that that estimate is substantially understated. Google put all these data points to Microsoft’s counsel, asking for clarification or correction—but received no response on the issue. *See* August 23, 2021 Ltr. to C. Simons (Microsoft’s Counsel) at 1-2.

Microsoft has overstated the burden of compliance in still other ways. For example, Microsoft represented at the July status conference that *its* proposed search parameters (instead of Google’s) would likely yield 1.49 million documents before deduplication. Hr’g Tr. 52, 55 (July 30, 2021). But Microsoft’s calculation assumed that every responsive email collected would, on average, have 3.25 “family members”—*i.e.*, attachments and the like—meaning that Microsoft multiplied the actual number of hits by 3.25. That multiplier is not accurate. A footnote buried in a fifty-page transmission Microsoft sent Google last week acknowledged that a more accurate multiplier was actually only 1.1. To put it another way, the burden assertion

---

<sup>5</sup> Microsoft also has declined to deduplicate its proposed collections against the 400,000 documents it produced during the CID phase.

Microsoft provided the Court at the July status conference was overstated by well over 100%. And that is before applying the deduplication reduction discussed above.

\* \* \*

Google is not insensitive to the burden of document discovery, but Microsoft’s burden objection needs to be put in context: Google has produced more than 3 million documents to the Plaintiffs, culled from a review set involving millions more. During the litigation phase alone, Google has collected documents from more than 120 custodians using more than 100 search strings, in addition to reviewing hundreds of thousands of documents collected from non-custodial repositories requested by Plaintiffs. Google is mindful that one should not “assess the reasonableness of a search query by its output” because “[w]hat matters is the quality of the records produced, not the quantity.” Mar. 23, 2021 Order (ECF No. 120) at 2.<sup>6</sup> But Microsoft has produced virtually none of the documents that Google has requested, and as noted, the pace of its productions suggests that, absent the Court’s intervention, that is not likely to change any time soon.

**C. Microsoft is at the heart of this case.**

Microsoft’s unsubstantiated burden objections would be insufficient even were it an ordinary third party. But it isn’t, and the “facts of the case”—including, for instance, the “need for the documents and the nature and importance of the litigation”—inform “[w]hether a burdensome subpoena is reasonable.” *Linder v. Dep’t of Def.*, 133 F.3d 17, 24 (D.C. Cir. 1998). On that score, the significance of Microsoft discovery is beyond dispute. Microsoft is, as the Court observed at the last hearing, the “first party among third parties.” Hr’g Tr. 65 (July 30,

---

<sup>6</sup> To be sure, had Microsoft made any material custodial productions in response to Google’s subpoena, Google would have been better positioned to even further tailor its search strings based upon the review of actual, responsive documents.

2021). Even the modest discovery Microsoft has made since then has confirmed the accuracy of this observation, [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

[REDACTED] The Federal Trade Commission [REDACTED]

[REDACTED] closed its investigation. [REDACTED]

[REDACTED]

[REDACTED]

[REDACTED]

Plaintiffs in turn have made Microsoft the effective centerpiece of their Complaints. Microsoft’s search tools power many so-called general search engines—not just its own Bing products, but also Yahoo, DuckDuckGo, and Neeva. Nor is the subpoena just about search engines. Plaintiffs suggest that Google’s Chrome browser forces Google Search on consumers, but ignores that Microsoft sets Bing as the default search engine in the company’s Edge browser, just as it did for years before that when Internet Explorer was the dominant browsing tool.<sup>7</sup>

Plaintiffs similarly accuse Google of using the Android operating system to inhibit competition

---

<sup>7</sup> See, e.g., DOJ Am. Compl. ¶ 3; see also *id.* ¶ 137 (“[T]he preinstalled version of Chrome is preset to default to Google search.”); see also Chris Hoffman, *How to Change Microsoft Edge to Search Google Instead of Bing*, HOW-TO GEEK (Jan. 23, 2020), available at <https://www.howtogeek.com/221709/how-to-change-your-default-search-engine-in-microsoft-edge/> (“Microsoft’s new Edge browser uses Bing as its default search engine . . .”).



in mobile-search distribution, even though Microsoft for many years engineered its proprietary Windows Phone mobile operating system to supply search results from Bing.<sup>8</sup> Other examples of square-on overlap between Plaintiffs’ allegations and the Microsoft documents sought by Google’s subpoena abound.<sup>9</sup>

For the reasons stated in Part B above, Microsoft’s proposals for narrowing its discovery obligations—drawing lines that are guaranteed to leave hundreds of thousands of documents undiscovered by Google—could not be justified even were Microsoft an ordinary third party. But because the Complaints position Microsoft squarely at the heart of the case—and plainly did so not just with Microsoft’s acquiescence but at its urging—Microsoft’s burden-balancing objections are particularly meritless, and should be overruled.

---

<sup>8</sup> DOJ Am. Compl. ¶ 147 (alleging that Google agreements “lock up all the main pathways through which consumers access search on Android devices”); *see also Microsoft Unveils Windows Phone 7 Series*, MICROSOFT NEWS (Feb. 15, 2010), *available at* <https://news.microsoft.com/en-au/2010/02/15/microsoftunveilswind/> (“Every Windows Phone 7 Series phone will come with a dedicated hardware button for Bing, providing one-click access to search from anywhere on the phone . . .”).

<sup>9</sup> *Compare, e.g.*, DOJ Am. Compl. ¶¶ 116-17 (alleging that Google has used its Android operating system to almost “completely shut out its competitors from mobile distribution”), *with* Google Subpoena (attached as Exhibit A), Request No. 49 (April 13, 2021) (seeking “documents that relate to any comparison between the quality or performance of the Windows Phone Operating System and Android.”); DOJ Am. Compl. ¶ 130 (alleging that “Amazon . . . agreed to make Microsoft’s Bing the preset default general search engine” on its Fire devices), *with* Google Subpoena, Request No. 5 (seeking “documents that relate to Microsoft’s provision of search results or search advertisements to any contractually designated party,” including Amazon); DOJ Am. Compl. ¶ 155 (referring to “premium placement and default status” as “the most effective forms of distribution”), *with* Google Subpoena, Request No. 6 (April 13, 2021) (seeking documents that “relate to the efficacy of different channels for distributing search engines”); DOJ Am. Compl. ¶ 170 (alleging that Google has pushed organic search results “further down the results page” by showing “more search advertising results”), *with* Google Subpoena, Request No. 61 (April 13, 2021) (seeking “analyses or experiments concerning the number of ad slots on Microsoft’s properties,” including Bing).

**D. The Court should compel compliance with the disputed search strings.**

Each of the eight disputed search strings (listed in Exhibit B) implicate topics bearing on Plaintiffs' allegations:

- ***String 1*** implements RFP No. 1, which seeks “documents that relate to Microsoft’s strategic or business plans regarding search engines, search access points, and search advertising.” This is among the most important document searches that Google has sought from Microsoft. According to Plaintiffs, it is not any strategic misstep by Microsoft—but rather, alleged anticompetitive conduct by Google—that has limited Microsoft’s success in these areas. Google is entitled to test those allegations, including through discovery of documents bearing on Microsoft’s pertinent strategic and business plans. String 1 is intended to address this core fact dispute.
- ***String 2*** implements RFP No. 2, which seeks documents reflecting the cost of or investment in Microsoft’s search engine, browser, voice assistant, and mobile OS businesses, as well as Microsoft’s decision to invest in these products. According to Plaintiffs, Google’s supposedly anticompetitive conduct hamstrung its rivals’ ability to compete. But if Microsoft simply declined to sufficiently invest in its products—notwithstanding its nearly bottomless financial and technical resources—that would severely undermine Plaintiffs’ allegations. And while the request implicates a number of different product lines, each figures prominently into Plaintiffs’ Complaints.
- ***String 5*** relates to RFP Nos. 5, 6, and 12, which address search and search advertising arrangements and related analyses, including documents “involving the promotion, preinstallation, or distribution of any Microsoft search engine, browser, [or] voice assistant” (RFP No. 12), “documents that relate to Microsoft’s provision of search results or search advertisements to any contractually designated party” (RFP No. 5), and “documents that relate to the efficacy of different channels for distributing search engines” (RFP No. 6). Responsive documents bear a clear resemblance to the types of agreements challenged in the present litigation. Microsoft has asserted that third-party search distribution and syndication agreements are not stored in one central location at Microsoft and so should be the subject of a custodial search. As a result, comprehensive search terms are warranted.
- ***String 14*** implements RFP No. 23, which seeks “presentations, including any PowerPoint presentations, and notes or scripts associated with them made to investors or at conferences relating to Bing’s performance, and all documents or data underlying any statements about Bing’s performance.” To date, on this score, Microsoft has produced publicly available earnings call transcripts; the purpose of this custodial search, however, is to target the internal documents behind such earnings calls and other public presentations.
- ***String 15*** implements RFP No. 24, which seeks “documents that relate to Microsoft’s promotion or marketing of any search engine, browser, voice assistant, or other

application with search functionality.” Responsive documents bear on the Complaints’ allegations. For example, Plaintiffs allege that Google’s conduct “denie[s] its competitors . . . vital . . . product recognition,” DOJ Am. Compl. ¶ 6, because, among other things, “the most effective way for general search engines to reach users” is to be preset as a default search engine, *id.* ¶ 51. But if Microsoft managed to drive recognition of its products anyway, that would undermine the Complaints’ allegations—especially if, notwithstanding Microsoft’s promotion, customers still preferred Google products.

- **String 16** relates to Microsoft’s analysis of its and its competitors’ search, web browsing, and search advertising services, and thereby addresses RFP Nos. 25, 26, 29, 30, 67, and 76 in Google’s subpoena. In particular, RFP Nos. 25, 26 and 29 address the quality, performance, relevance, and/or precision of search tools, search advertising, and web browsers; RFP No. 30 relates to user preferences for search engines; RFP No. 67 relates to particular performance metrics for Microsoft advertising; and RFP No. 76 relates to experiments performed on Microsoft’s vertical search service units. These requests all probe how Microsoft’s relevant product lines stack up against others’ (like Google’s) in areas that may matter to their customers. They will thus shed light on whether Microsoft was restrained from competition or simply struggled to deliver products valued as highly as Google’s.
- **String 21** implements RFP Nos. 40 and 41, which seek documents that “relate to or identify the firms and products Microsoft considers its competitors for the provision of search services or search advertising” or that “concern competition between any Microsoft operating system and any operating system that runs on a mobile device.” Microsoft’s competitive landscape matters because, according to Plaintiffs, Microsoft ranks among Google’s most significant competitors.
- **String 38** addresses voice assistants and related issues, including, for instance, “Microsoft’s decision to end voice assistant products” (RFP 81), “customer preferences among voice assistants” (RFP 83), and the extent of consumer demand for “concurrency”—that is, the simultaneous activation of multiple voice assistants on a single device (RFP 84). The Complaints repeatedly address these subjects and allege, for instance, that “Google recognizes that concurrency is a feature that consumers would value.” DOJ Am. Compl. ¶ 163. Microsoft—the maker of the voice assistant, Cortana—doubtless possesses many documents relating to this aspect of Plaintiffs’ allegations.

Despite their manifest significance, Google has narrowed the disputed strings to lessen their burden on Microsoft. At present, Google’s proposed versions of the above eight strings yield, in total, an estimated deduplicated hit count of 1.17 million (using the estimated 50% deduplication rate discussed above, *see supra* § I.B). Compared to earlier Google formulations of these same strings, those figures represent a hit-count reduction of more than 70%. Microsoft,

however, proposes strings that yield far lower hit counts, with an estimated deduplicated hit count of approximately 695,000. And that is because Microsoft’s search strings are unduly narrow:

- For String 1, which concerns “strategic or business plans regarding *search engines*,” Microsoft insisted on deleting the word “Bing” from Google’s proposed search string. But Bing, of course, is Microsoft’s (modern day) search engine. Inclusion of the word “Bing” is not overbroad because of [REDACTED]  
[REDACTED]  
[REDACTED] Notably, Microsoft has not provided Google with a sampling of any documents containing “Bing” [REDACTED]  
[REDACTED] that nevertheless are not responsive to Google’s subpoena; whatever Microsoft’s complaints about the volume of documents to be searched, the volume is simply driven by the fact that there are apparently many documents relating to Microsoft’s strategies for Bing.
- [REDACTED]  
[REDACTED]  
[REDACTED] From a preliminary review, Google believes that Microsoft only produced scarcely two dozen documents, consisting of Microsoft’s quarterly earnings call slides and transcripts, most (if not all of which) appear to be publicly accessible. Microsoft’s preliminary production therefore confirms the need for a comprehensive, custodial search.
- For String 21, which concerns, *inter alia*, “competition between *Google* and Microsoft with regard to search engines, browsers, *voice assistants*, *connected cars*” (RFP 70), [REDACTED]  
[REDACTED] String 21 also concerns “competition between any Microsoft *operating system* and any operating system that runs on a mobile device” (RFP 71); [REDACTED]  
[REDACTED] Again, it has provided no sampling of documents from a search of these terms that are not responsive to Google’s subpoena.
- For String 38, which concerns, *inter alia*, “Microsoft’s decision to *end* [its] voice assistant products” (RFP 81), [REDACTED]  
[REDACTED]  
[REDACTED]<sup>10</sup> String 38 also concerns “customer

<sup>10</sup> See, e.g., Tom Warren, *Microsoft shuts down Cortana on iOS and Android*, THE VERGE (Apr. 1, 2021), <https://www.theverge.com/2021/4/1/22361687/microsoft-cortana-shut-down-ios-android-mobile-app> (“Microsoft has shut down its Cortana app for iOS and Android. It’s the latest in a series of moves to end support for Cortana across multiple devices . . . Microsoft first launched Cortana for iOS and Android in December 2015. The app was originally designed to connect Windows 10 PCs and mobile phones, but failed to gain traction despite a big redesign.

*preference* among voice assistants” (RFP 83), [REDACTED]  
[REDACTED]

- For Strings 2, 5, 15, and 16, Microsoft insisted on using extremely stringent proximity limiters—*i.e.*, requiring that sets of search terms be located within just a few words of each other, substantially narrowing the volume of internal conversations that will be picked up for review. Google has agreed to such narrowing for several strings—and indeed, narrowed its proposals for *these* four strings, too, through the conferral process—but submits that these are sufficiently important to warrant a fulsome search.

Given the importance of the discovery, Microsoft’s vast resources, and the significance of the present litigation, Google respectfully submits that Microsoft’s objections in these respects should be overruled. *Cf.* Mar. 23, 2021 Order (ECF No. 120) at 3 (“In a case of this significance and scope, some deference to Plaintiffs’ expressed need for more records than what Google is offering is warranted.”).

#### **E. Substantial completion.**

Microsoft now seeks to delay substantial completion until October 31, 2021. At the July 30 hearing, Microsoft proposed October 20 (Hr’g Tr. 60 (July 30, 2021)), and since then it has failed to make substantial progress on compliance. Google’s position is that given the centrality of these documents to the fact discovery period, and given that the subpoena has been outstanding since April, substantial phased productions should be made over the next several weeks, with substantial completion by September 30, 2021.

---

Microsoft CEO Satya Nadella recognized the company’s difficulties in competing with other digital assistants a couple of years ago, revealing that Microsoft no longer saw Cortana as a competitor to Alexa and Google Assistant.”).

## **II. Microsoft's Position Statement**

Non-party Microsoft desires to cooperate fully in discovery in these cases. Microsoft recognizes the centrality of its search business—and its experience competing in search and search advertising—to the Parties and the Court's understanding of the claims of restricted competition and consumer harm in this litigation. Moreover, Microsoft appreciates the monumental importance of the litigation to competitors and consumers in digital markets and the world. But the significance of these cases does not give Google the license to act in an abusive manner, and its discovery demands have crossed the line.

### **A. Background and Overview of the Dispute**

Throughout the discovery process, Google has made unreasonable demands of Microsoft. Google issued Microsoft a massive, 120-request subpoena on the last possible day pursuant to the court's discovery schedule, and during negotiations has essentially demanded that Microsoft start from scratch: throwing search terms used for the underlying investigation out the window, rejecting Microsoft's proposals that leverage the lessons learned from the pre-complaint investigations, and insisting on search terms that show little evidence of incorporation of the voluminous information already provided to Google detailing Microsoft's search and advertising businesses.<sup>11</sup> Not only is this approach unreasonable, Google has delayed discovery negotiations and resulting productions from Microsoft by months.

### **Microsoft's Proposal**

---

<sup>11</sup> Google has been in receipt of productions of more than 400,000 Microsoft documents and 180 GBs of Microsoft data beginning February 17, 2021. *See* E-mail from Haley Tynes, Orrick, Herrington & Sutcliffe LLP, to John Schmidlein et al., Williams & Connelly LLP (Feb. 17, 2021) (on file with author and recipient); Joint Status Report, *United States v. Google LLC*, No. 1:20-cv-03010 (D.D.C. Feb. 23, 2021), ECF No. 111; Joint Status Report, No. 1:20-cv-03010, (March 28, 2021), ECF No. 124.

In a good-faith attempt to fully comply with its obligations and advance the process, Microsoft has made successive, meaningful compromises and proposals, culminating in its most recent proposal to run 44 search strings and conduct a search for documents responsive to 9 go-and-gets, totaling a hit count of *over 3.5 million documents*, collected from 45 custodians, and spanning over 21 years of Microsoft’s business (“Microsoft’s Proposal”).<sup>12</sup> Google’s competing proposal insists that Microsoft collect over 1 million more documents.<sup>13</sup> Again, Google has created an impasse preventing the Companies from meaningful progress on document discovery.

At issue before the Court are (1) eight disputed search strings and (2) a dispute over Microsoft’s substantial completion date. *See* Figure 1.

**Figure 1.**

<b>Disputed Search Strings Proposal</b>				
<b>String</b>	<b>RFP Nos.</b>	<b>Microsoft Total Volume</b>	<b>Google Total Volume</b>	<b>Difference</b>
1	1	187,743	441,724	<b>253,981</b>
2	2	131,164	188,698	<b>57,534</b>
5	5, 6, 12	150,425	257,012	<b>106,587</b>
14	23	0 - GNG	8,351	<b>8,351 (GNG v. search)</b>
15	24	114,265	177,049	<b>62,784</b>
16	25, 26, 27, 29, 30, 67, 76	312,612	392,397	<b>79,785</b>
21	40,41	427,110	557,152	<b>130,042</b>
38	81, 82, 84, 84	62,944	308,524	<b>245,580</b>

**Comprehensive Proposal<sup>14</sup>**

<sup>12</sup> The materials that reflect the disputed terms, Microsoft’s proposed go-and-gets, and the agreed-upon terms are attached hereto as Exhibit 1.

<sup>13</sup> Google is not naïve: its strategy in maximizing volume and not engaging meaningfully with Microsoft on substantive discussions as to the relevancy benefits of greater volumes is emblematic of a fishing expedition with the added effects of maximizing burden on Plaintiffs, maximizing burden and cost on a direct competitor, and—in attacking Microsoft in court and in the press—directing the spotlight away from itself.

<sup>14</sup> Microsoft conducts in-place searching which does not provide globally de-duplicated numbers until the search results are collected. These figures represent estimated non-deduplicated,



Microsoft	Google	Difference
3.42 million	4.36 million	~947,000
Substantial Completion Date Proposal		
Microsoft	Google	Difference
October 31, 2021	September 30, 2021	4 weeks

### Meet and Confer Process

Microsoft's Proposal represents significant movement and compromise since its July 21 proposal that was before the Court during the July 30 hearing. Microsoft has met and conferred with Google for months and repeatedly highlighted issues with Google's approach, including:

- the relevance of certain custodians for certain topics, given their business units and roles;
- the overbreadth and improper terms Google incorporated into proposed search strings, which did not align with how Microsoft's business is run and did not leverage the information in the 400,000+ documents already produced;
- the overreach in date range, given the facts and claims in this case; and
- the delay caused by demanding an all-up comprehensive proposal instead of using priority requests and phased negotiations to expedite productions.

At the same time, Microsoft understands the pre-trial schedule in this case and its importance as a non-Party in this case. Therefore, despite its misgivings about Google's approach, Microsoft has actively engaged in good faith with Google and has compromised on as

---

expanded hit counts. Microsoft historically has observed a 3.25 expansion rate (representing total documents collected corresponding to parent level hits) from other litigations but has recently observed an approximate 1.1 expansion rate in this case as documents have been processed. Similarly, Microsoft historically has observed an average de-duplication rate of 30%, and is currently observing a de-duplication rate of between 25%-42%, depending on the search terms and custodians as they are being processed for review and production. By contrast, Google has reported to Microsoft a markedly lower de-duplication rate of 24-27% for its own productions. Even using the largest de-duplication rate of 42% that would represent the most conservative estimate of Microsoft's burden, Microsoft's most recent proposal would still result in **over 2 million documents** that Microsoft would be required to review for responsiveness for this litigation. Microsoft has been providing Google with real-time updates of its observed de-duplication rate as terabytes of data have continued to process.



many searches as it reasonably can, shouldering significant burden. Microsoft accepted the 18 additional custodians<sup>15</sup> Google raised in July. Microsoft has accepted all but 8 of Google's 44 search strings. And despite the fact that Bing was launched in 2009 and Plaintiffs' subpoenas to Microsoft predominantly seek documents from 2010 to present,<sup>16</sup> Microsoft has agreed to search documents from the 2000s, and has further agreed to search from the 2018-2020 timeframe even though Microsoft already produced over 400,000 relevant documents from that period during Plaintiffs' investigations.<sup>17</sup> Recognizing the Court's schedule, Microsoft has been proactively and steadily processing, reviewing, and producing documents responsive to the Parties' Subpoenas, even prior to reaching agreement on search terms.<sup>18</sup>

Google itself has previously represented to this Court that such demands as unreasonable. For instance, when Google objected to Plaintiffs' discovery requests, Google argued that issuing a "massive" discovery request and forcing the responding party to "start from scratch and begin producing everything out of the heavens is not reasonable[.]" and that the appropriate approach is to learn from and leverage the millions of pages produced during pre-complaint investigations,

---

<sup>15</sup> Microsoft agreed despite issues with Google's July 27 filing about the relevance of these 18 custodians, which contained misstatements. For example, Google attributed to one custodian certain statements about Android to substantiate its argument that she had relevant documents, but those "statements" were actually attributable to the headline of an article that she had forwarded, not what she herself had said. *See* Defendant Google LLC's Statement Regarding Disputes Involving Compliance with a Subpoena to Microsoft, *United States v. Google LLC*, No. 1:20-cv-03010 (D.D.C. July 27, 2021), at Appendix A., p. 1, proposed additional custodian no. 3.

<sup>16</sup> *See* Department of Justice's January 26, 2021 Subpoena to Microsoft and the State of Illinois Subpoena to Microsoft.

<sup>17</sup> In an attempt to minimize Microsoft's burden, Google may argue that Microsoft's documents will be de-duplicated against the CID Productions. But even if the de-duplication rate for that universe were 100%—it is not—that would reduce Google's proposal from 4.36 million to a still-extraordinary almost 4 million documents.

<sup>18</sup> To date, Microsoft has made twenty-eight productions in this litigation, in addition to the significant productions of documents and data it has made during the underlying investigation.

and that it would be “ludicrous” to “throw [the search terms used for the investigation] out the window and start from scratch” because it would make discovery “inefficient and not productive[.]”<sup>19</sup> Yet when it comes to discovery of Microsoft—an important non-party, yet still a non-party to these cases—Google has adopted the very discovery approach it previously lambasted as “ludicrous.”<sup>20</sup>

As to Microsoft’s most recent Proposal, Google has refused to meaningfully compromise or even meaningfully negotiate the disputed search terms. Google simply argues in conclusory fashion that the topics are “important” and therefore merit higher volumes. The Companies are therefore at an impasse. Microsoft respectfully requests that, as to the disputed search strings, that the Court rule in favor of Microsoft and further order that Microsoft’s Proposal would satisfy Microsoft’s obligations with respect to the document requests in the Subpoenas.

#### **Substantial Completion Date**

Months of delay have resulted from Google’s boil-the-ocean-before-any-agreement approach. Yet, in acknowledgement of the Court’s schedule, Microsoft has committed to a substantial completion date of responsive, non-privileged documents by October 31, 2021. Despite Microsoft’s best efforts, it acknowledges it has reached impasse with Google on this issue as well because Google demands that Microsoft substantially comply with Google’s 103 documents requests by the impossible deadline of September 30, 2021. Specifically, Google’s proposal ignores wholesale both Microsoft’s burden and the time it takes to process and review millions of documents. Accordingly, Microsoft respectfully requests that the Court order a substantial completion date of October 31, 2021.

---

<sup>19</sup> See Feb. 25, 2021 Status Conf. Tr. 35:21-24; 35:12-15; 34:18-35:6; 35:16-36:2 (attached hereto as Exhibit 2).

<sup>20</sup> See *id.*

**B. Google’s Objective Is to Maximize Document Volume Without Tailoring Search Strings to Yield Relevant, Responsive Documents**

Microsoft’s objections to the eight disputed search strings fall into two broad categories:

(1) overbroad search terms that are untethered to how Microsoft’s business is run, and (2) overly broad proximity connectors.

**1. Overbroad Search Terms**

Instead of incorporating the information provided to date by Microsoft or reviewing the millions of pages of documents Microsoft has produced to date, the disputed search strings in this category make clear that Google has sought to start anew. Google should have tailored its subpoenas and its proposed search terms based on lessons and vocabulary learned through careful review of what Microsoft has already produced.<sup>21</sup> As explained in more detail below, this is not what Google has done.

**String 1 (RFP 1):** “*All documents* that relate to Microsoft’s strategic or business plans regarding search engines, search access points, and search advertising . . . include[ing] *all documents* that discuss or reflect Microsoft’s decision to develop, launch and/or maintain Bing as a search engine. . . . from 2003 to present.”<sup>22</sup>

RFP 1 seeks Microsoft’s strategy documents relating to its search business. Microsoft and

---

<sup>21</sup> See Fed. R. Civ. P. 45(d)(1) and 45(d)(3)(A)(iv); *United States v. AT&T Inc.*, No. 1:11-CV-01560 (ESH), 2011 WL 13273135, at \*3 (D.D.C. Oct. 29, 2011) (with respect to whether a subpoenaed nonparty or the requesting party “bear[s] the obligation of sorting through the 2.2 million pages already produced to determine which portions of the subpoena remain unaddressed by the [nonparty Sprint’s] production” noting that the Special Master is persuaded that “Rule 45’s emphasis on protecting non-parties supports Sprint’s claim that [the requesting party, AT&T] should bear this initial responsibility. Requiring AT&T to make that determination is consistent with the purposes of Rule 45 and with the serving party’s Rule 45(c)(1) duty to avoid imposing unduly on non-parties. Inasmuch as AT&T has had the Sprint/DOJ documents since September 27 [Opp. at 1], it cannot be said that AT&T has not had the opportunity to make such a determination. Indeed, AT&T states in its October 28, 2011, communication with the Special Master that it has already taken significant steps in that direction. Thus, a limited time for further review is warranted.”).

<sup>22</sup> See Non-Party Microsoft Corporation’s Status Report Regarding Third Party Discovery at Exhibit A, No. 1:20-cv-03010 (APM) (July 27, 2021), ECF No. 169.01 (emphasis added).

Google’s dispute rests on the inclusion of “Bing” in the search string, which alone results in over 230,000 documents. Google’s proposal should be rejected because it is apparent that “Bing” is overbroad given its disproportionate impact to the hit count:

Proposal	Search String	Key Difference	Total Documents <sup>23</sup>
Microsoft	[REDACTED]	Exclusion of and “Bing”	187,743
Google	[REDACTED]	Inclusion of “Bing”	441,724

First, with respect to this search string, relevant and responsive documents that discuss Microsoft’s search business would be squarely captured by terms other than “Bing”: a document between Microsoft employees discussing business strategy for search or search advertising is unlikely to reference “Bing” *without also referencing* “SAN,” “MSAN,” or “search advertising,” for example. Indeed, the terms SAN and MSAN, which are acronyms for Search, Ads, News and Microsoft Audience Network, [REDACTED]

[REDACTED]<sup>24</sup> By contrast, “Bing” is Microsoft’s trademark that is used widely by the U.S. market to refer to Microsoft’s

<sup>23</sup> The totals referenced herein are non-de-duplicated.

<sup>24</sup> See generally *U.S. Equal Employ. Opp. Comm’n v. George Wash. Univ.*, No. 17-cv-1978 (CKK/GMH), 2020 WL 3489478, at \*11 (D.D.C. Jun. 26, 2020) (“[A] responding party, not the court or requesting party, is generally best suited to determine and implement appropriate procedures, methodologies, and technologies to identify, preserve, collect, process, analyze, review, and produce relevant and discoverable ESI, and no Federal Rule has given judges the authority ... to dictate to parties how to search their documents.”) (citing *The Sedona Conference, The Sedona Principles, Third Edition: Best Practices, Recommendations & Principles for Addressing Electronic Document Production*, 19 Sedona Conf. J. 1, 118 (2018)).



Microsoft.”<sup>27</sup> This is both unsupported and unavailing. Microsoft’s Proposal should be adopted because Google’s alternative results in a volume of documents that is unduly burdensome and not proportional to the needs of the cases at bar.

**String 14 (RFP 23).** RFP 23 (String 14) seeks investor presentations and underlying materials relating to Bing performance.<sup>28</sup> Microsoft has proposed a Go-and-Get in lieu of a custodial search because, as Microsoft has repeatedly explained to Google, copies of these materials [REDACTED].

Yet Google insists on custodial searches for these documents, proposing custodians who are unlikely to have the material Google seeks in response to this Request.<sup>29</sup> Google’s proposal should be rejected.

Given the nature of the documents sought and the known source where these materials are stored, custodial searches would be ineffectual and it makes no sense for Google to reject Microsoft’s proposal for a go-and-get. Google’s purported concern is that Microsoft would produce only publicly-available materials in response to this RFP. But Microsoft has already represented that it is in the process of conducting a non-custodial search to locate both public investor presentations and the related non-public documents.

**String 21.** Similarly, the Requests that form the basis of String 21 (Request Nos. 40 and 41), broadly seek “[a]ll documents” that “relate to or identify the firms and products Microsoft

---

<sup>27</sup> See Ex. D to Ltr. from Benjamin M. Greenblum, Williams & Connelly LLP, to Caroline Simons, Orrick Herrington & Sutcliffe (Aug. 23, 2021) (Strings 5 and 15) (attached hereto as Exhibit 5); Ltr. from Benjamin Greenblum, Williams & Connelly LLP, to Caroline Simons, Orrick Herrington & Sutcliffe (Aug. 26, 2021) (attached hereto as Exhibit 6).

<sup>28</sup> See Non-Party Microsoft Corporation’s Status Report Regarding Third Party Discovery at Exhibit A, No. 1:20-cv-03010 (APM) (July 27, 2021), ECF No. 169.01.

<sup>29</sup> See Exhibit 5 and Exhibit 6.

considers its competitors for the provision of search services or search advertising” and “[a]ll documents that concern competition” between Microsoft’s operating system and any mobile operating system.<sup>30</sup>

Proposal	Search String	Key Difference	Total Documents
Microsoft	[REDACTED]	Exclusion of “Google,” [REDACTED]	427,110
Google	[REDACTED]	Inclusion of “Google,” [REDACTED]	557,152

Google’s proposed search term, which yields over 550,000 hits, is overbroad because, *inter alia*, it would return documents based solely on some variation of “compete” (e.g., “competition,” “competitor”) within the same paragraph as “Google.”<sup>31</sup> Google states it is “illogical”<sup>32</sup> to remove the term “Google,” but in fact “compete . . . Google” would likely yield numerous non-responsive and irrelevant documents for the simple reason that Microsoft and Google compete on numerous businesses unrelated to search, search advertising, or mobile OS, such as productivity applications (Microsoft Office / Google Workspace) and cloud services.

<sup>30</sup> See Non-Party Microsoft Corporation’s Status Report Regarding Third Party Discovery at Exhibit A, No. 1:20-cv-03010 (APM) (July 27, 2021), ECF No. 169.01.

<sup>31</sup> See Exhibit 5 and Exhibit 6.

<sup>32</sup> See E-Mail from Benjamin M. Greenblum, Williams & Connelly LLP, to Haley P. Tynes et al., Orrick, Herrington & Sutcliffe LLP (Aug. 26, 2021) (on file with author and recipient).

Microsoft’s proposal, which does not contain “Google,” and five other similarly overbroad terms, results in a more reasonable, but still high-volume, hit count of 427,110.<sup>33</sup> Google’s proposal should be rejected.

**String 38 (RFPs 81-84).** The underlying requests for String 38 (Request Nos. 81 to 84) relate to “Microsoft’s decision to end voice assistant products,” the “Voice Interoperability Initiative,” “customer preferences among voice assistants, use of voice assistants to complete actions, use of multiple voice assistants (i.e., multihoming), or the ability to switch between voice assistants,” and “interest in, the effect of, challenges associated with, or customer demand for concurrency.”<sup>34</sup>

The search term proposed by Google results in over 280,000 hits because it includes a series of broad terms including [REDACTED]

[REDACTED]  
[REDACTED]<sup>36</sup>

Given the generic nature of these terms, particularly in light of the fact that Microsoft did not “end” Cortana, their inclusion are likely to bring in substantial numbers of non-responsive documents. Microsoft proposed the permutation drafted by Google that excludes the overbroad terms but is targeted to the documents sought in Google’s request.<sup>37</sup> These more-tailored terms result in a lower hit counts, from 281,374 to 57,405.

---

<sup>33</sup> See Microsoft Corp. Resps. to Google’s Search Term Counterproposal (Aug. 25, 2021) (attached as Exhibit 7).

<sup>34</sup> See Non-Party Microsoft Corporation’s Status Report Regarding Third Party Discovery at Exhibit A, No. 1:20-cv-03010 (APM) (July 27, 2021), ECF No. 169.01.

<sup>35</sup> Exhibit 5 and Exhibit 6.

<sup>36</sup> See *id.*

<sup>37</sup> See Exhibit 7.



## 2. Unreasonably Expansive Proximity Connectors

For the four remaining disputed search strings,<sup>38</sup> the impasse arises from the Companies' disagreement as to the appropriate proximity connectors. For each of these four strings, Microsoft has agreed to both search term permutations and custodians<sup>39</sup> drafted by Google. When Microsoft attempted to object to the appropriateness of terms and custodians, which risked yielding high volumes of non-responsive documents that Microsoft would nevertheless have to review, it was met with Google's unwillingness to narrow or change custodians. Therefore, in an attempt to compromise while still addressing its burden, Microsoft accepted Google's terms and custodians on the condition that more closely-targeted proximity connectors be used to manage Microsoft's burden. More closely-targeted proximity connectors also would be more likely to yield responsive documents. However, Google took the opposite approach and expressly stated it was looking for larger volumes.<sup>40</sup>

String	Google Proximity Connector(s)	Google Total Volume	Microsoft Proximity Connector(s)	Microsoft Total Volume	Difference
2	w/5	188,698	w/3	131,164	57,534
5	w/15 and w/15	257,012	w/10 and w/5	150,425	106,587

<sup>38</sup> String 2 (Request No. 2), String 5 (Request Nos. 5, 6, and 12), String 15 (Request No. 24), and String 16 (Request Nos. 25, 26, 29, 30, 67, and 76).

<sup>39</sup> In an attempt to compromise, Microsoft agreed to run these search strings across the custodians Google requested, despite the fact that, in many instances, the Google-requested custodians were untethered to the business unit, product(s) and subject matter of the request as drafted in the Subpoena. For example, String 15 is targeted to Subpoena Request 24, which seeks marketing and promotional material regarding any Microsoft search engine, browser, voice assistant, or other application with search functionality. Google's proposed custodians include custodians not in marketing or business development roles (██████████). Microsoft agreed to these custodians in the spirit of compromise and reaching an agreement with Google on the substantive terms; a spirit of compromise that Google has not shared with respect to these proximity connectors for Strings 2, 5, 15, and 16.

<sup>40</sup> See Exhibit 7.

15	w/15 and w/15	177,049	w/10 and w/10	114,265	62,784
16	w/5 and w/5	392,497	w/5 and w/3	312,612	79,885

Taken together, Google’s proposed connectors in Strings 2, 5, 15, and 16 increase parent-level search terms hits by **over 306,000** over Microsoft’s more targeted proposal. Google justifies its overbroad connectors with the same, template language: “[i]n light of the importance of this RFP to Plaintiffs’ allegations, and the way the terms can be expected to be used in business settings, Google believes that a proximity connector of [larger] is more appropriate than [smaller],”<sup>41</sup> or simply that Microsoft’s proposal is “unduly narrow.”<sup>42</sup> However, Google does not point to a single document, out of the nearly 440,000 documents that Microsoft has produced, that *Microsoft’s* employees use these terms in the manner Google suggests in *Microsoft’s* business setting. In this instance, Google’s conclusory and repetitive demands fails to adequately explain why these expansive connectors should be used. To be clear, to the extent that Google offers an explanation in this filing beyond the above conclusory statement, it will be the first time Microsoft hears that rationale. Google should be required to negotiate in good faith, and its proposals that simply aim for more volume should be rejected.

### C. Conclusion

Microsoft takes its discovery obligations seriously. Microsoft’s Proposal reflects the company’s good faith offer to shoulder an extraordinary burden in acknowledgement of its obligations and the significance of this litigation. Microsoft’s Proposal reflects a good-faith commitment that is properly designed to produce a significantly large volume of relevant, responsive, and non-privileged documents—while proportionally taking Microsoft’s burden into

---

<sup>41</sup> See Exhibit 5.

<sup>42</sup> See Exhibit 6.

account. The Court should rule in favor of Microsoft on the disputed search strings, order that Microsoft's Proposal be adopted, and order a substantial completion date of October 31, 2021.

Dated: August 27, 2021

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)  
725 12th Street, NW  
Washington, DC 20005  
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*

By: /s/ Caroline Simons

Caroline Simons (*pro hac vice*)  
Orrick, Herrington & Sutcliffe LLP  
22 Berkeley Street  
Suite 2000  
Boston, MA 02116  
Telephone: +1 617 880 1800  
Facsimile: +1 617 880 1801  
csimons@orrick.com

Amy W. Ray  
Haley P. Tynes  
Orrick, Herrington & Sutcliffe LLP  
Columbia Center  
1152 15<sup>th</sup> Street, N.W.  
Washington, D.C. 20005-1706  
Telephone: +1 202-339-8400  
Facsimile: +1 202-339-8500  
amyray@orrick.com  
htynes@orrick.com

*Attorneys for Non-Party  
Microsoft Corporation*