IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

UNITED STATES OF AMERICA, ET AL.,	)	SEALED
	)	
Plaintiffs,	)	
	)	CV No. 20-3010
vs.	)	Washington, D.C.
	)	September 21, 2023
GOOGLE LLC,	)	9:45 a.m.
	)	
Defendant.	)	Day 8
	)	Morning Session

TRANSCRIPT OF BENCH TRIAL PROCEEDINGS
BEFORE THE HONORABLE AMIT P. MEHTA
UNITED STATES DISTRICT JUDGE

APPEARANCES:

For DOJ Plaintiffs:

Kenneth M. Dintzer
U.S. DEPARTMENT OF JUSTICE
1100 L Street, NW
Washington, D.C.
(202) 307-0340
Email:
kenneth.dintzer2@usdoj.gov

Erin Murdock-Park
U.S. DEPARTMENT OF JUSTICE
Antitrust Division
450 5th Street NW
Washington, D.C. 20530
(202) 445-8082
Email:
erin.murdock-park@usdoj.gov

For Plaintiff
State of Colorado:

Jonathan Bruce Sallet
COLORADO DEPARTMENT OF LAW
Consumer Protection Section,
Antitrust Unit
Ralph L. Carr
Colorado Judicial Center
1300 Broadway
Suite 7th Floor
Denver, CO 80203
(720) 508-6000
Email: jon.sallet@coag.gov

William F. Cavanaugh, Jr.
PATTERSON BELKNAP
WEBB & TYLER LLP
1133 Avenue of the Americas
Suite 2200
New York, NY 10036-6710
(212) 335-2793
Email: wfcavanaugh@pbwt.com

APPEARANCES CONTINUED:

For Defendant Google:

John E. Schmidtlein
Kenneth Charles Smurzynski
WILLIAMS & CONNOLLY LLP
725 12th St., NW
Washington, D.C. 20005
(202) 434-5000
Email: jschmidtlein@wc.com

Kenneth Charles Smurzynski
WILLIAMS & CONNOLLY
680 Maine Avenue, SW
Washington, D.C. 20024
(202) 434-5000
Email: ksmurzynski@wc.com

Court Reporter:

William P. Zaremba
Registered Merit Reporter
Certified Realtime Reporter
Official Court Reporter
E. Barrett Prettyman CH
333 Constitution Avenue, NW
Washington, D.C. 20001
(202) 354-3249

Proceedings recorded by mechanical stenography; transcript produced by computer-aided transcription

- - -

WITNESS INDEX

- - -

- - -

INDEX OF EXHIBITS

- - -

- - -

INDEX OF EXHIBITS

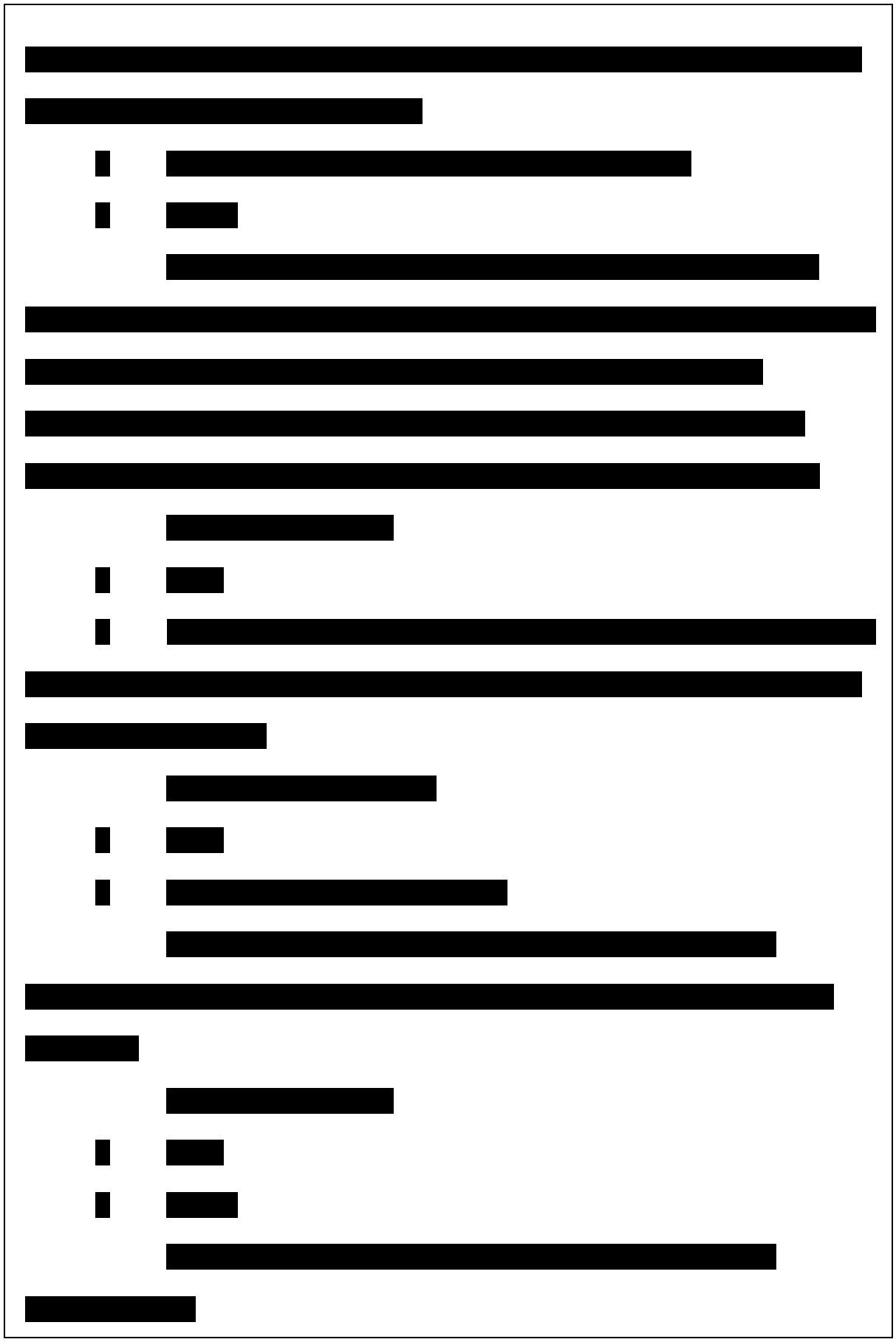
- - -

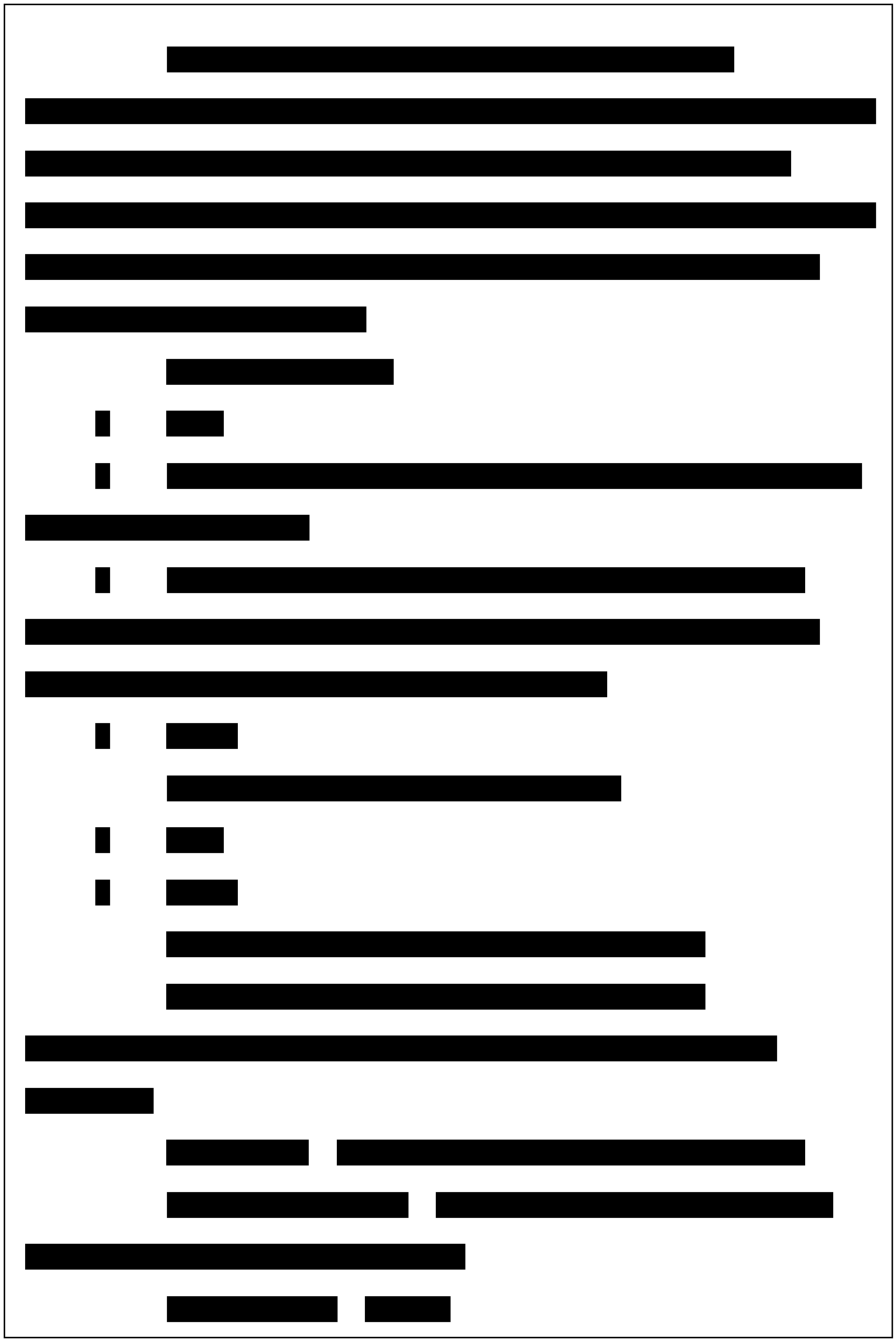
P R O C E E D I N G S

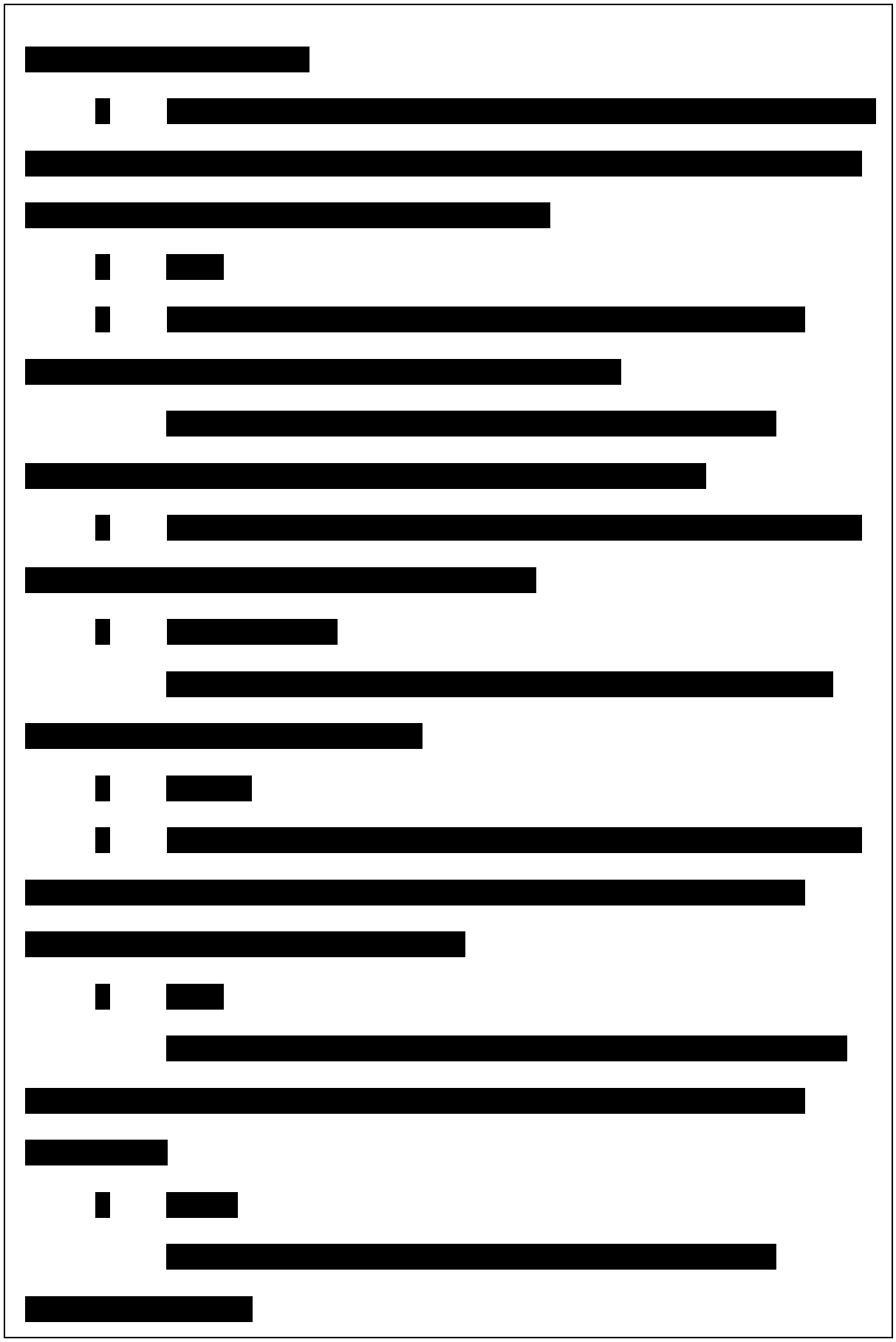
(Sealed closed proceedings)

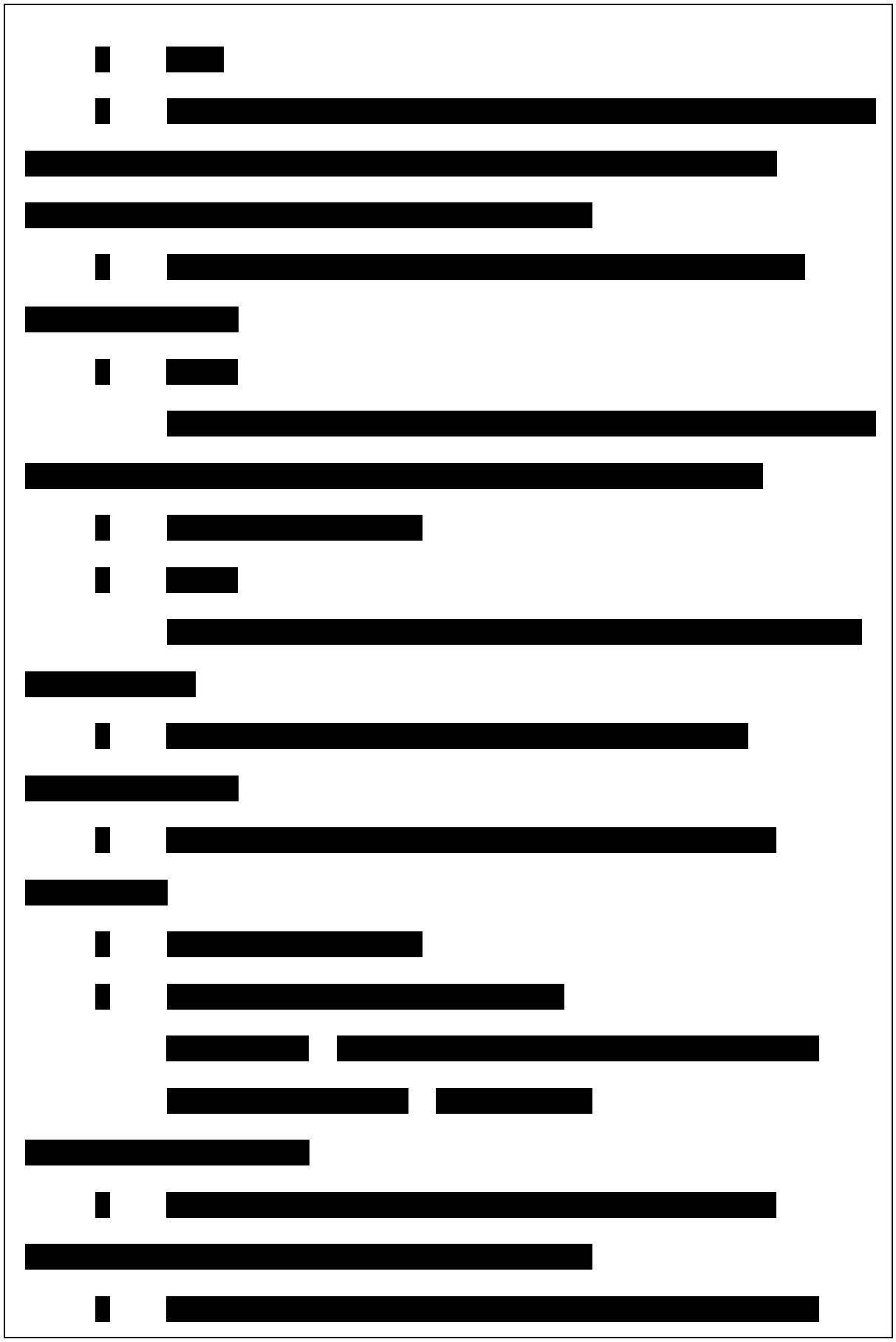
[redacted]


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




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


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


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

[redacted]

(Open court)

BY MR. SMURZYNSKI:

Q Good morning, Dr. Lehman.

Yesterday, the plaintiffs asked you a number of questions about Google's use of clicks and different ranking components. Do you remember that?

A Yes.

Q What is the relationship between, on the one hand, the number of ranking components that utilize clicks in some way and on the other, the volume of clicks those components need?

A So I guess I would make two notes about that. One is maybe sort of obvious, which is that it's not additive. That is, if you have 20 components, you don't need 20 times as many clicks. We've shown the chart showing all the number of clicks for those different components. Again, it's all the same clicks.

There is a more subtle aspect. I think yesterday we talked about the difference between memorization systems

and generalization systems. So now boost would be an example of a memorization system. It remembers for a particular search query and a particular web page user did or did not click.

But the primary effort in use of clicks over the last 10 or 15 years has been on generalization systems. There, the goal is to generalize from situations where we have data to situations where we don't. So these are kind of like filling in the holes where we don't have data.

So I think in terms of how these systems respond to different quantities of available data, a factor to keep in mind is that even though we're adding more systems, many of these are generalization systems. They're systems that are designed to fill in holes in data. So I think that adds a kind of a robustness to the system in that if you reduce data, those systems can patch up the holes.

Some of the -- those older versions of systems themselves need quite a bit of data train, but the new ones require far less.

THE COURT: Can I ask a question?

The systems that you've described as generalized systems, in the development of those systems, was ClicData a component of how those systems were developed?

THE WITNESS: Yes, yes.

Roughly, those systems tried to find patterns

typically in the ClicData so that you can say, Well, I see, for 300 cases like this, people tended to do this kind of click, and hours here's the 301st case which, well, it's sort of like that, so I'm going to bet people click sort of the same way.

And the notion of same varies from one generalization system to another. And gradually, as we sort of move to deep learning, the kinds of patterns these systems could look for to generalize ClicData got more and more complex.

THE COURT: Got it. Thank you.

BY MR. SMURZYNSKI:

Q Dr. Lehman, are you familiar of a principle called diminishing returns?

A Yes.

Q And how, if at all, does that principle impact the value Google obtains from user data?

A So I think there's sort of good empirical evidence that Google's query stream has a certain distribution called Zipfian, Z-i-p-f, and that produces this long tail effect that we talked about previously.

I guess one way to think about it is that you can imagine, you start up your search engine one day and so you start processing queries, and over time, the number of queries that you see grows and grows and grows sort of

linearly.

And then you can say, Well, what is the probability, when I get a new query, that I've seen it before.

And I think a reasonable model, it's pretty well-established over decades, is that it grows logarithmically, that is, sort of a classic diminishing returns curve.

Q And what implications, if any, does that have for the value of additional user data to Google and with respect to the quality of Google's search engine?

A Well, for a system like navboost, as the volume of queries that Google has seen in the past grows, the percentage of situations that we've seen before begins to level off, and so you get diminishing returns.

Q We've talked about navboost, and you just referenced it there.

With respect to queries that Google has seen only a few times, how does navboost treat that relative to queries that it's seen hundreds, thousands, tens of thousands, millions of times?

A Right.

So for many queries in navboost, there's a large number of queries that we've seen a small number of times, a small number of queries that we've seen a large number of

times, and for a query that we've seen a small number of times, we'll have scraps of data. So each click is a little bit of a hint, but they come with caveats. And if you don't have too of many of them, it's hard to draw a strong conclusion.

So, for example, there might be a search query where one person clicked on one result, two people clicked on another result, but, you know, maybe this current user is in Australia and both of those clicks came from people in the United States. So maybe we should trust them, maybe not.

Maybe some other results seem plausible to our other search components, but we're not sure that anybody actually even sort of looked far enough down the page to see them.

So that's sort of the typical picture for one of these queries that -- for which we don't have much data. And so I think navboost does kind of the natural thing, which is, in the face of that kind of uncertainty, you take gentler measures. So you might modify the score of a document but more mildly than if you had more data.

THE COURT: Can I ask another question, I'm sorry.

So the world is ever evolving.

THE WITNESS: Yep.

THE COURT: How do -- how does the search engine

handle new developments in the world?

Here's a terrible example but one that I will understand and some of you may.

A lot of people put Taylor Swift into the search engine, right? So you've already got in navboost lots of connections for Taylor Swift but apparently she's now dating Travis Kelce who's a football player. So if I want to get the latest on that, how does the search engine -- how does it -- how is it able to identify these new developments and get people to those relevant links?

THE WITNESS: Yeah. So this is a complex question within Search, we talked about it as the problem of freshness.

And there's no single simple answer.

Some systems respond very quickly to new developments in the world, while others, others don't.

So one kind of system that actually responds very quickly is the sort of the old school techniques. So if you do Taylor Swift, and I don't know, maybe some other term about her relationships or whatever, then the old systems will still be looking for web pages that talk about Taylor Swift and relationships. Now, they could be dated pages. But at least they'll work.

And I think that's important when something like COVID happens. Like the word "COVID" wasn't even a word so

a lot of our systems didn't know about synonyms for COVID, COVID-19 or whatever. But still these old school systems work, which is why it's kind of nice to keep them around.

There is a special sort of freshness component that tries to look at things like, often news articles will have by-line dates so you can say, well, how old is this article versus this one. This is about Taylor Swift's relationship two years ago, this is one two weeks ago. And so it can sort of boost the more recent ones more aggressively.

Does that answer the question?

THE COURT: I think so.

But it sounds like there are different components that would evaluate these variables. It's not simply navboost or -- there are multiple systems that operate to identify -- or help promote fresh results?

THE WITNESS: Yes.

And Clicks is one of them. You know, like once we start to get Clicks on a topic, there's sort of a fast response system that tries to make use of those as quickly as possible.

But it's something we struggled with for a long time. And, you know, I think one argument was we should try to think about freshness in all our search components. I don't know if we've 100 percent succeeded with that though.

There's some that -- particularly around COVID, there was a crash effort to try to update our systems to ensure that we had relevant information. So if everything were working perfectly, that wouldn't have been necessary but it wasn't.

THE COURT: Got you.

All right. Thank you.

BY MR. SMURZYNSKI:

Q I just want to go back to that concept you were talking about, the weighting of navboost between long-tail queries and what you described yesterday as head queries. Sort of without getting into like the nitty-gritty of the mathematics of it, how does the system do that?

A So qualitatively for these head queries that we've seen many times and for which we have ClicData, we're more confident in that boost data and we can adjust the scores of documents more dramatically or as for queries that are in the long-tail where we have little scraps of data, it's ambiguous, it's harder to figure out what to include. Navboost will adjust the scores of documents more mildly.

Q If you could turn to document UPX228 which is in your binder and we'll just put the cover page on the screen.

A 228? I got it.

Q It should be in the first section, yes.

A Right.

THE COURT: Sorry, Counsel, I was still taking --

what's the number?

MR. SMURZYNSKI: Certainly.

It's 228.

THE COURT: Thank you.

BY MR. SMURZYNSKI:

Q And yesterday, Dr. Lehman, you said there was a somewhat complicated story behind why you gave this presentation. What was that story?

A Right.

So this was an opportunity for me to talk to the whole Search leadership, which is a rare opportunity for me.

And the people to whom I was really targeting this talk were the people who do design of the sort of the user experience of Search, how things look on the screen. And also the more senior Search people who make decisions about how many special features we put on the page and how busy it is.

And the reason that I was speaking to those audiences implicitly is that I was concerned about the connection between the way Search results are visually presented and the ranking of Search results.

And I don't think that connection was widely appreciated and so that's something I was trying to drive home in this presentation.

In particular, my concern was that our Search

results page could get too busy, and sometimes results could be presented in ways that were a little bit hard for people to understand.

And because Clicks are an important signal in ranking, if the page is too complicated, then it's hard to interpret Clicks.

Remember, the point of Clicks is that we're asking people to sort of read by proxy for us. Computers couldn't read. So we have people read the search results and the query and everything else on the page and then we observe their reactions, and those observations are Clicks.

But as the page gets busier and more complicated and if Search result previews aren't really clear, then we observe users doing things but we don't know how to interpret them and extract information from them.

So my message was to people who work in sort of the visual presentation of results and decide the layout of the page, no, you need to think about this. We need to learn from our users in the form of Clicks.

And I think this was not obvious. I mean, Search is a huge organization with thousands of people, and I'm getting up as somebody who's kind of a data wonk who deals in algorithms and that kind of stuff and I have to talk to people from a very different world, the sort of user experience people who are kind of more thinking about human

behavior and art and visual presentation and say, no, we need you.

And so one thing I was trying to do in this talk is motivate those people to help us. I can't just browbeat them.

And so that's why I sort of introduced this notion of Magic. So I wanted to say, if you help us, then we can get Google to do these magical things, these things that seem otherwise impossible. So I'm trying to sort of inspire the people in those organizations to work a little bit differently, work with us so we can make this really cool stuff.

So these are intended to be inspirational examples.

MR. SMURZYNSKI: Your Honor, if I may approach the witness?

BY MR. SMURZYNSKI:

Q Dr. Lehman, I put in front of you a document that has been marked as UPX203 which is in evidence.

And take a look at the cover slide.

What is this presentation?

A So periodically there will be all-hands meetings where everyone from an organization is invited to attend. This looks like it was from the fourth quarter, the end of the year in 2016, it was an all-hands presentation for the

Search organization.

Q All right. And if you go in a page, there is an agenda; is that right?

A Yes. Right.

Q And under "Ranking," there's the name Eric. Was that you?

A Yes.

Q Much less formal at Google than in court?

A Yes.

Q And if you go forward in the deck to the page that ends in 904, and is this the portion of the presentation that you gave at the end of 2016?

A Yes.

Q All right.

And if you go forward one more slide, this is the first slide in your presentation.

What's the topic that you're addressing in your presentation?

A So this is a time when sort of the first wave of deep learning had arrived and we were starting to try to understand, you know, what are the implications for the future.

And I think as we've discussed, you know, the biggest problem in Search is you're given a web page and a query and you need to figure out whether or not that web

page is relevant to the search query. And the challenge has always been, well, computers can't read so we have to do all these workarounds.

But we're starting to wonder if that might possibly change.

And I guess the first line here, "Search is a great place to start understanding language." Maybe there are two sides to that. One is that Search was, I think, a really powerful driving application for this fundamental scientific problem of how you get computers to understand language, how do you get computers to read.

And in this presentation, I also thought Google was a great place to begin exploring that problem, not only because of the motivation but because of the resources. We have access to large amounts of text, and we also had access to lots of user data, which at this time I thought would probably be an important asset.

Q All right. And if you turn the page to 905, what are you saying there?

A Right, so this is, I think, a thing that's come up quite a bit lately.

So with this fundamental problem of determining whether a web page is relevant to a query, as of this presentation, after just the first wave of deep learning systems, we really couldn't read documents. It was -- our

ability was minimal. So we played this game of read by proxy. Show the text to people, we observe their reactions and we adopt them as our own.

And this approach, it has limitations, though, that we've talked about. So, for example, when we show people search results, they don't actually read the whole web page, they just read the preview. They tend to click on results that are lower quality than we'd like to show, and yet, at this time, that was a really important technique. So user data, you know, is an important source of data and use was used in many Google systems.

Q And there you're describing today which is in the context of this document, December of 2016, correct?

A That's right. Sort of after the first wave of deep learning systems.

Q If you turn to the page that ends in 919.

And on the top of the slide you're positing a task; is that right?

A Yes.

Q And what is that task?

A So in this presentation, we're -- I was starting to contemplate the possibility that rather than this whole mechanism of read by proxy, have people read for us and try to learn from Clicks, that we might -- I guess here I'm saying, we might maybe someday, we might get to a point

where -- with a computer we'd be able to understand a document, like a web page or a passage, there are just a few sentences instead of a web page, based just on the content alone. That is, we wouldn't need humans to interpret it for us. We could get computers to directly understand a document or a passage just based on the words.

And, you know, I think this was a little bit of a crazy idea at the time, and I think I felt pretty anxious about presenting this.

You know, the history of artificial intelligence, I guess I would like to say that to sum up the first history, the first 50 or 60 years of AI research, I would say not much happened. And so I think for people in the field, the idea that we could actually do something like AI, it was almost embarrassing to suggest it. It felt like AI was a marketing hype term, not real.

MR. SMURZYNSKI: Your Honor, if I may approach?

THE COURT: I guess before you get to the next document, though.

So finish your thought, Dr. Lehman. This was your -- this was the view of the world at the end of 2016. So how has that changed in the ensuing six years?

THE WITNESS: It's been kind of a complicated journey.

The next big milestone for me was the arrival of

BERT around 2018, and that was just a shock to me on multiple fronts.

One was that -- so BERT was developed in connection with something called WebAnswers. So, you know, when you do a search, sometimes Google will try to guess the answer to your search and put it at the very top. That's WebAnswer. So that's a case where there's just a short passage instead of a full web page involved.

And so BERT was developed, and that was its first application. And there was like a Christmas holiday, and what we found out just before that is BERT beat all of our existing systems combined by a large margin. It just made everything else irrelevant. So that was kind of a milestone moment for me.

And then, you know, in some ways, WebAnswers is sort of like the junior version of Web Search, because you only have to work with just a small passage instead of full web pages.

So I think at that point, it started to become clear that we were looking at a change that would kind of knock all the pieces off the board of search probably at some point within the next few years.

And then things began to pick up speed, so I -- it felt like at first maybe there would be a breakthrough every four years, and then it got to be like, Hey, wow, every year

there's a breakthrough. And then it's just like every few months.

And it's just accelerated and accelerated.

So I was just noticing in the news that the quality of systems that take text as input and produce images that are described by that text has actually jumped significantly since lunch yesterday.

So I think one of the big surprises to me is -- was that -- and I still don't have my head wrapped around it, is that these systems could learn from text alone.

I believe, and I think there may be presentations here, that language acquires meaning because it means something to people. Otherwise, it's just symbols on a page. And so I thought that for machines to understand language in a humanlike way, they would have to learn from seeing people interact with data. And so I thought user data would be essential to helping machines learn language.

But it turned out not to be the case. It turned out that these very large machine learning systems can learn simply from text. So that was another sort of huge milestone.

And I guess when I left Google, things had begun to accelerate so fast, and it looked like the implications were just going to be staggering, and not just for search but for the world at large.

And I did my very best with it, but at some point I feel like I needed to back away and hand it over to someone who's younger and ready to deal with it, I guess.

BY MR. SMURZYNSKI:

Q And, Dr. Lehman, you have in front of you what's been marked as DX237, which is already in evidence.

You talked about BERT in your response to the judge. What was the next sort of milestone in the journey in terms of these large language models?

A Yeah, let me sort of informally say there was sort of these first-generation systems around 2014. There was BERT around 2018 as a sort of second generation. For me, at least, this was sort of the third generation. So this is MUM, which was sort of completed and started to get ready for production use in 2021.

I guess this is a presentation of MUM to the head of search at that time, Prabhakar Raghavan. I think it's kind of a monster document, particularly to give to a senior executive, but I think this one he read through thoroughly.

Q And I used a term there, "large language model." What is a large language model?

A So it's a computational system that tries to, in some way, capture patterns in language.

And the name is maybe a little bit deceptive. Language models could be something as simple as noticing

that like adjectives usually come before nouns, and united is often followed by states, things like that.

And people would try to capture these as models of language, it's sort of systems that reflect the patterns in language.

And then people started building larger language models that could capture more patterns in language, and they got larger and larger and larger.

And then strange things started to happen. It seemed that these systems began -- in their effort to find patterns in language, they began to, in some ways, it appears mimmick the cognitive processes of the people who created that language, you know.

A sort of trivial example is if you give one of these systems a bunch of school worksheets or something like that where it says 3 plus 5 is 8, 4 plus 7 is 11, and so on, then in its search for patterns in language, it will learn to do math, and so it can begin to do arithmetic questions that it's never know seen before.

And that one's fairly clearcut, but then it can do some of the more qualitative cognitive things that people do. And it's limited. It's not AI. Or at least it wasn't then, but it was quite startling to us.

Q This, as you describe, is a long, complicated document, it contains many trade secrets. But if you go to

the third page of the document, there's a section that's marked "TL;DR." Not what those letters mean, but how is an acronym used within Google?

A As sort of executive summary.

Q Okay.

And in the -- about halfway through the paragraph, you write, "These models are significantly more powerful than BERT, demonstrating a large improvement on SuperGLUE benchmarks, and let's just pause there for a second.

What is a SuperGLUE benchmark?

A So as people were trying to make computers understand language, and by people here, I mean people at Google, people in academia, people in other U.S. corporations, corporations around the world, particularly in China, they wanted a way to measure progress that we could all agree upon. And so some folks set up some -- a system called GLUE, and it was a set of several language challenges.

And then whenever someone had a new computer system that they believed could understand language better, they could use this GLUE system to check the progress of their work.

So they would check how well their system worked on each of those challenges. There would be taken -- they'd take an average. And there was sort of a leader board

saying, Who's gotten the highest score across all tasks, who's second highest and third highest. And there were sort of a lot of people competing.

And then BERT came along and just demolished everything. So everything on that leader board became variants of BERT. And the performance of BERT was so strong that for those tasks, it was performing above human level. But the tasks weren't that hard, so that wasn't that impressive.

So the creators of GLUE created a new set of reading comprehension challenges, which they called SuperGLUE that was designed to be the BERT killer. It's even stickier. So they were much harder challenges, and BERT exhibited fairly weak performance. It was scoring about 60, and they speculated this would be much harder for machines.

However, within a matter of, I don't know, about a year or so, a system called T5, which later evolved into MUM, among other systems, achieved essentially human-level performance.

Q And in your answer, you mentioned BERT, T5, and MUM. What was Google's involvement in the creation of each of those systems, if any?

A So BERT was developed by Google researchers working with the WebAnswers team, and then they later

constructed sort of the Version 2 of rank embed, which I think we talked about earlier.

T5 was a significantly more powerful system that got very high scores on the SuperGLUE benchmark. It was also developed by Google researchers. They were sort of off doing their own thing at Google research, but then some of us in search contacted the T5 team and we began working together. And MUM is -- can be thought of as a much scaled up version of the T5 system developed by Google researchers.

And that sentence says -- you continue on, it reads, "And reaching human-level performance for many tasks."

Do you see that.

Yes.

Q Was among those tasks the ones that, in the 2016 document, you had posited maybe someday a machine will be able to do this?

A I don't know if it was exactly the same task. I'm not sure, you know, which one I was thinking of at that time. But it was -- but these are challenging reading comprehension tasks. I think, you know, people find them challenging, they don't score perfectly. They score about the same as MUM.

MR. SMURZYNSKI: Your Honor, this is DX241. And at least one of the plaintiffs, I believe, has lodged a

hearsay objection to it, but I don't know if they continue to maintain it.

MS. MURDOCK-PARK: Your Honor, our understanding is that this is not a document associated with Dr. Lehman.

THE COURT: Well, is there any objection to its admission?

MS. MURDOCK-PARK: I believe there's a hearsay objection, and I don't -- nope, we have no objection.

THE COURT: Okay.

MR. CAVANAUGH: No objection, Your Honor.

THE COURT: So it will be admitted. DX241.

(Defendant's Exhibit DX241 received into evidence.)

BY MR. SMURZYNSKI:

Q Dr. Lehman, if you could turn forward to the slide that begins 448, and this just is to orient you. The slide reads "MUM, a major milestone in language understanding."

MR. SMURZYNSKI: And, Alan, if you could go forward to the next slide.

BY MR. SMURZYNSKI:

Q And we'll go through these, but just at a general level, could you describe for the Court what is being communicated in this slide?

A This looks like a few high-level bullet points about the MUM model, the sort of third-generation language

understanding system.

Q Okay.

And in the first bullet there, what is being conveyed with regard to the comparison between BERT and MUM?

A So when working with language, I think we found that if you need to memorize a lot of facts about the world, you need a lot of memory but not a lot of computation and systems like QBST work well.

But if you want to do reading comprehension and sort of reasoning tasks, you need to do a lot of computation.

And what this is saying is that the amount of computation per unit of text is about a thousand times greater in MUM than BERT, and the consequence is that it should be able to understand language more deeply, perform reasoning more effectively.

Q And in the third bullet, there's a reference to "understand and generate human language."

What data does MUM use in order to achieve that task?

A So MUM was trained on large volumes of text drawn from the Internet.

Q Was MUM trained on click-and-query data that Google had?

A No, and as I mentioned, it was a shock to me that

that's possible, but no, there was -- there are no -- no user data.

Q And the next bullet down refers to "Few-shot learning requiring far fewer data inputs than BERT."

And these data inputs, again, are raw text from the Internet?

A This may be a little bit different.

So with these sort of language systems like BERT and like MUM, typically you train them initially on lots of language, and in the process, they understand the structure of language and acquire some kind of reasoning abilities. It's a little hard to tell what they're doing, they're just giant masses of numbers. And then typically you would apply them to some particular task.

With BERT, the procedure for applying them to a particular task would be a process called fine-tuning, where after they understand the basic structure of language, you give them some inputs and desired outputs, and that modifies the internal structure of the model a little bit so that it gets really good at that one specific task.

So, for example, if BERT were being adapted for WebAnswers, the task that they used for fine-tuning would be, here's a query, this passage is relevant; here's another query, this passage is not relevant. And you would go to a bunch of those examples, and then it would get very good at

that task.

For large language models, they can be adapted to specific tasks in a very different way. These things take text as input and you would just give a few examples written as text as your initial input and then pose your new problem also in text and, based on just those just few examples, in many cases they could just begin to perform that new task.

Q And is that what the reference to few-shot learning is?

A Yes, that's few-shot, each shot is like an example and sometimes, you know, two or three or something are enough.

Q And that's all you need, two or three in order to accomplish this with MUM?

A Yeah. Yeah.

Q If you could go to the next slide.

And to remind you, this is as of July of 2021, we saw on the first slide.

And without reading the number out loud, you see halfway down the page, there's a reference to the number of search related teams that are using MUM to improve existing products. Do you see that?

A Yes.

Q And to your knowledge, that number is accurate as of July of 2021?

A It seems about right, that number started growing very fast and so I don't know exact value, exact dates.

Q Fair enough. In July of 2021, as of the time you left Google, was that number higher or lower?

A It was higher.

Q Dr. Lehman, I put in front of you a document that's been marked as UPX197 which is in evidence?

A Yes.

Q And I'd like to direct your attention to the top of that email. And it's an email you write on December 26th, 2018.

A Yes.

Q And then you write, I'd like to offer a thought for contemplation over the break.

And then you set forth something in bold text there. Do you see that?

A Yes.

Q And what did you write at the end of 2018 to the various folks in search leadership who are identified in the top of this email?

A So this is over the Christmas holiday. And we had just seen the results of BERT for this web-answered problem. We've just seen that it had outperformed everything that we'd done, you know, dozens of engineers over roughly a decade.

And so I sent this email to reflect on what the implications of that were for the future. It was quite a shock. I always felt that Google's -- you know, it's really big advantage is that by focusing on Search for essentially so many years, we developed this deep body of theory and practices and we made a lot of mistakes but we learned over time. And I thought that was kind of our treasure chest.

And then along comes this system and just like, no, never mind, beats everything. And at this point we had just seen that this kind of, sort of junior problem, WebAnswers, had experienced this just dramatic disruption, where machine learning made everything previous kind of irrelevant.

And the thought I'm considering here is that that's going to happen to Web Search, too. I was guessing that these kinds of advances would -- yeah, just kind of clear the table of all past work.

Q And if you could scroll down, Alan, to the paragraph that begins at the bottom of the page, one consideration and then carries over.

BY MR. SMURZYNSKI:

Q And, Dr. Lehman, could you read that paragraph for us, please.

A "One consideration is that such a deep ML system could well be developed outside of Google - at Microsoft,

Baidu, Yandex, Amazon, Apple, or even a startup. My impression is that the translate team experienced this. Deep ML reset the translation game; past advantages were sort of wiped out. Fortunately, Google's huge investment in deep ML largely paid off, and we excelled in this new game. Nevertheless, our new ML-based translator was still beaten on benchmarks by a small startup. The risk that Google could similarly be beaten in relevance by another company is highlighted by a startling conclusion from BERT: Huge amounts of user feedback can be largely replaced by unsupervised learning from raw text. That could have heavy implications for Google."

Q And you wrote this at the end of 2018 with BERT. By November of 2022, with the introduction of MUM and further advances in ML, what's your view?

A I think by 2021, with the arrival of MUM, it seems like this is the path we're going down. And I think, MUM was a big deal in 2021 but now these large language models are everywhere. You can download them from the web and run on a large PC. So I think indeed we have seen such deep ML systems developed outside of Google at pretty much, I think, almost every major tech company in the U.S. and many abroad, and a notable startup like OpenAI.

So, yeah, I think this is largely come to pass and it's a story that's still unfolding, but I think it seems

like this is a road we're still on.

MR. SMURZYNSKI: Thank you, Dr. Lehman. I have no further questions.

THE COURT: Dr. Lehman, just quickly, deep ML system. ML stands for?

THE WITNESS: Machine learning.

THE COURT: Got you. All right. Thank you.

MS. MURDOCK-PARK: If I may have a minute, Your Honor, I will have a few redirect.

THE COURT: Counsel, before you begin your -- I'm sorry, go ahead, Ms. Murdock.

MS. MURDOCK-PARK: No. Please, Your Honor.

THE COURT: No, no. I was going to ask one more question but just whenever you're ready.

MS. MURDOCK-PARK: Happy to get Your Honor's question before I get started. It might have been one of mine.

THE COURT: So, Dr. Lehman, could I ask you to project out five years, ten years from now, what would you guess to be the relationship between these deep ML systems and user data in terms of search?

THE WITNESS: So my guess, and it's really hard to guess --

THE COURT: Fair enough.

THE WITNESS: -- is that search engines will shift

largely from a reliance on user data and all these other tricks we built up over the years to systems that draw upon these deep learning systems largely.

I think there will still be a role for user data but I think it will be much diminished, but I think user data does still bring some things. Maybe the clearest thing would be popularity.

So, for example, you know, I don't know, I search for toothpaste. Well, there are a lot of kinds of toothpaste. And so all those associated web pages are relevant and a machine-learning system can say, oh, they're relevant. Okay, but they're not all likely popular. And I think user data can still help inform something like that.

So I think there will still be a role for user data but, you know, trying to get a high resolution picture of the world from user data requires a vast amount, trying to figure out which is the most popular toothpaste and a bunch of things like that I think will require much, much less user data.

So generally I think a lot of the function of user data will be replaced but probably not all of it. But it's a guess.

THE COURT: Thank you.

- - -

REDIRECT EXAMINATION

BY MS. MURDOCK-PARK:

Q Dr. Lehman, I want to ask you -- we'll start at the last document that your counsel showed you, UPX197.

A Yes.

Q And I want to go to the last page. And the last page is -- this continues the email that you sent, correct?

A Yes.

Well, I think there's several different people contributing text at this point because people were responding.

Q So I want to go to the top paragraph. And you wrote, "And we already know one signal could be more powerful than the whole big system on a given metric. For example, I'm pretty sure that navboost alone was/is more positive on Clicks (and even likely on precision/utility metrics by itself than the rest of ranking). By the way, engineers outside of navboost team used to be also not happy about the power of navboost and the fact it was 'stealing wins')."

You wrote that?

A No.

Q You didn't? This is in the email that you sent.

A I believe that this section was written by Alex

Grushetsky.

If you see on the first page at the very top, where it says from, Alexander Grushetsky. So I believe in this text, his comments are being interleaved with mine and I'm pretty sure that that was his comment.

Q And do you disagree with Mr. Grushetsky?

A Well, he's saying a lot of things here. Parts of it, I agree with. Parts of it, I question what he's talking about, because they don't seem right to me. Alex is a really smart guy. If he were here, I would say, wait, what do you mean, Alex. We could go through it phrase by phrase and I could point out which parts that make sense to me and which don't, if you want.

Q Well, my biggest question is, you said, "I'm pretty sure that navboost alone was/is more positive on Clicks and likely even on precision/utility metrics by itself than the rest of ranking."

A Right.

So there's several assertions there.

So the first part is navboost alone was/is more positive on clicks.

And that is, I think, true but kind of misleading.

So I think at some point we looked at a chart and we talked about how components of ranking could be evaluated using, I call it "live traffic experiments."

And the idea of those experiments is, for a given ranking component, you ask, are the search results at that system boosts getting more or fewer clicks than the search results that would otherwise be there. And so that would be such a system, if such a system -- if a component boosted results that got more clicks than results that were previously there, then we would say it is "positive on clicks."

The navboost system is by far the most positive on clicks; that is, it boosts results that in experiments then get lots of clicks.

But that is somewhat misleading, and so clicks are not usually used to evaluate navboost. And the reason it's misleading is because it's almost circular.

The navboost system, what it does is observe documents that in the past got lots of clicks, and then it boosts those.

So the near circularity is there's some documents or web pages that got clicks in the past, they get boosted up by navboost and then this goes into the evaluation system which says, ah, I see navboost has boosted all these documents and they get lots of clicks. Well, of course they do.

What that measurement of navboost on clicks is really telling you is that people tend to click on the same

results over time. So the past is fairly predictive of the future.

And one way to see why that's kind of a bad type of analysis, I think we've talked about how people often click on low quality results, bad search results, it's not super often but there's a tendency.

So you imagine what's going to happen with those clicks on bad search results. Well, the navboost system is going to observe these bad results, got clicks, it's going to boost them, and then an evaluation system, these bad results will be boosted, and they'll continue to get clicks.

And so navboost will look even more positive on clicks even though it's doing something that's actually damaging search quality.

So, yes, it's true that navboost alone is very, very positive on clicks, but that's deceptive.

Q We can go ahead and put this document aside.

I'd like to go back to UPX203, which your counsel showed you, and specifically the page ending in 907, which is Slide 51.

And this is the portion of UPX203 that you wrote for the ranking section, correct?

A Let's see.

I'm just needing a minute to find it.

Q Of course. It is not in the binder. Your counsel

handed you binder clipped copies. I believe it's the thickest one.

A The big one. Okay.

Okay.

Q And it's page 57, but it ends in Bates 8907.

It's a graphic of some people with arrows to each other.

Do you see that, Dr. Lehman?

A Yes.

Q And the graphic says, "Each searcher benefits from responses of past users and contributes responses that benefit future users," right?

A Yes.

Q And you wrote, "So if you search right now, you'll benefit from the billions of past user reactions we've recorded. And your responses will benefit people who come after you. Search keeps working by induction." Right?

A Yes.

Q Okay.

We can go ahead and put 203 aside.

Now, you talked a little bit about MUM with your counsel. Is another consequence of the amount of computation needed by a system like MUM is that it adds latency?

A That's a little bit of a complicated question.

Certainly computation takes time. There's kind of a whole repertoire of techniques for trying to reduce the time required by these language systems on -- when used in a large-scale application like search.

But, yeah, computation takes time.

Q And time is latency in terms of how long it takes to pull up a search result?

A Yes. And, again, we should be careful because when you do a search, many, many things happen in parallel.

So, for example, the DeepRank system which runs on BERT, I believe that adds no latency to Google's search results because we found a spot where we could run it in parallel with other activities that take more time.

Latency is a challenge, but there are ways to manage it to a degree.

Q Okay.

And another consequence of the amount of computation needed by a system like MUM is expense, right? The systems are expensive?

A Yes, these systems are expensive.

Q When MUM was launched, Google still used more traditional ranking systems, right, with respect to navboost and QBST, certainly?

A Yes, definitely.

Q And you mentioned BERT again. Any competitor can

use BERT, correct?

A Yes.

Q But Google had a nice head-start on using BERT because of the amount of user data available to Google, right?

A I don't think that's true.

I think Google's biggest advantage in using BERT early was that it invented BERT.

The first two applications of BERT were in WebAnswers, where I -- I don't -- I'm not sure if they used user data at all. ClicData isn't very useful for WebAnswers.

And then the second application was DeepRank, where we hustled, we got it launched.

I think Microsoft may have, after we launched and announced it, said in some kind of press release, Well, we did it first, but maybe my memory is not clear.

Q If I could show you UPX193, which is a redacted document, and it's in evidence.

MS. MURDOCK-PARK: If I may approach, Your Honor?

THE COURT: Yes.

BY MS. MURDOCK-PARK:

Q And this is a document, Dr. Lehman, titled "Bullet Points for Presentation to Sundar."

Do you see that?

A Yes.

Q This is a document that was in your files.

If we could highlight Bullet 4 where it says, "Any competitor can use BERT or similar technologies. Fortunately, our training data gives us a nice head-start. We have the opportunity to maintain and extend our lead by fully using the training data with BERT and serving it to our users. This needs multiple orders of magnitude more hardware resources (mainly TPUs)." Right?

A Uh-huh.

MS. MURDOCK-PARK: No further questions, Your Honor.

MR. SMURZYNSKI: Your Honor, since this document was introduced for the first time on redirect, I'd like to ask the witness one or two questions on it.

THE COURT: Could I just ask a question of Dr. Lehman first?

When you say "training data" in this bullet point, what are you referring to?

THE WITNESS: Yeah, I think that's a significant thing.

So this document appears to be talking about the application of BERT to WebAnswers, which was the first application. So, like, I think it says in Bullet Point 2, "Applying BERT to WebAnswers demonstrated this huge coverage

increase."

And then the third bullet point says, "Early experiments with BERT apply to several other areas of search."

So I think it's talking about the use of BERT in WebAnswers primarily and saying that we started exploring its use in other areas.

So I think here, training data refers to the first application, which is in WebAnswers.

So in that case, the training data would be this large bank of query passage pairs, where people employed by vendors hired by Google have marked the answers as good, bad, or somewhere in between, and Google built up that bank of training data over some years. And I think that's what it's referring to here.

I don't think it's referring to ClicData because, for these cases where we show our best attempt at an answer at the top of search results, ClicData is not very helpful because if we show somebody an answer and it answers their question, then they don't need to click.

Click is usually at least a weak indication that they're not satisfied with the answer we provided. So there's a little bit of information.

My understanding of the value of user feedback in this WebAnswer system is that it's much less than a ranking

of search results. One person told me a click is worth about a tenth. Another person disagreed and said it's more. I haven't worked on it and I don't know exactly.

But to answer your question directly, I think this training data refers to human-rated answers.

THE COURT: Do you have any follow-up questions, Counsel?

MS. MURDOCK-PARK: No, Your Honor, but we would object to Google having a re-redirect since we asked questions directly relevant to their cross, used a document to do so.

MR. SMURZYNSKI: Your Honor, you asked the question I was going to ask, so I have no further questions.

THE COURT: Okay.

All right. Dr. Lehman, this means you are done. Thank you very much for your time and your testimony, and safe travels home.

THE WITNESS: Thank you.

THE COURT: Thank you.

All right. Plaintiffs ready with their next witness?

MR. HOFFMAN: Yes. Good morning, Your Honor. Ian Hoffman for the United States of America.

Our next witness is Gabriel Weinberg.

So you know, Your Honor, we anticipate both an

open session and a closed session.

THE COURT: Okay.

MR. HOFFMAN: I've been told the witness is in the hallway, Your Honor.

COURTROOM DEPUTY: Before you have a seat, please raise your right hand.

(Witness is placed under oath.)

COURTROOM DEPUTY: Thank you.

THE COURT: Mr. Weinberg, welcome.

THE WITNESS: Thank you.

MR. HOFFMAN: May I proceed, Your Honor?

THE COURT: You may.

- - -

GABRIEL WEINBERG, WITNESS FOR THE PLAINTIFF, SWORN

DIRECT EXAMINATION

- - -

BY MR. HOFFMAN:

Q Good morning, Mr. Weinberg. My name is Ian Hoffman, and I represent the United States of America in this case.

Could you state your name and spell your name for the Court, please.

A Sure.

Gabriel, G-a-b-r-i-e-l, Weinberg, W-e-i-n-b-e-r-g.

Q And where do you currently work, Mr. Weinberg?

A I work at DuckDuckGo.
Q And what is your role at DuckDuckGo?
A I am the CEO, and I founded the company too.
Q And in what year did you found DuckDuckGo?
A 2008.
Q And where are DuckDuckGo's headquarters?
A Paoli, Pennsylvania.
Q And can you tell the Court what DuckDuckGo is?
A Sure.

So DuckDuckGo is the leading Internet privacy company.

We're most well-known for our private search engine.

Over the last several years, we have expanded the product vision to be more the easy button for privacy is what we call it. Our flagship kind of product experience is our free browser, which is a replacement for Chrome. It has the private search engine built in, along with about 12 other privacy protections that also protect your browsing, email, other online activity.

Q And how many people --

THE COURT: I'm sorry to interrupt you.

I may have missed it. What's the name of the browser?

THE WITNESS: DuckDuckGo. Same name.

THE COURT: Oh, same name, okay. The browser and the search engine are the same name?

THE WITNESS: Yeah. Our basic pitch is to be the easy button. And so not have multiple products, we have one product, it's just like pressing the easy button. You download DuckDuckGo, you get everything.

THE COURT: Got you. Thank you.

BY MR. HOFFMAN:

Q And how many people does DuckDuckGo employ today?

A Right about 250.

Q And how many searches does the DuckDuckGo search engine run every day?

A About a hundred million.

Q One hundred million?

A Yes.

Q Okay.

How does DuckDuckGo generate its results?

A So to kind of appreciate that, it might be useful to give a little context.

Q Sure.

A It -- when search engines started back in the '90s, there was like one index of links, and I think that's how people still think about them.

But since about the mid-2000s, search engines increasingly put these boxes on top of the results.

Sometimes they're on the side, down below, but a lot on top.

I'm going to call these search modules, but other people refer to them as one boxes, sometimes the knowledge graph collectively, instant answers.

So these search modules actually have their own indexes. So -- and there's about -- these are things like maps, directions, weather, news, images, things like that.

I'd say there's about 20 of them that are essential now.

So a search engine today doesn't really just have one index, they're really operating with all of these kind of dozens of indexes.

So when I started the company, this was one of the thesis of the company originally was I thought these search modules were the future of Search and I wanted to focus on that because it was the innovative portion. So I tried to license the traditional links as soon as I could, at the beginning, that was with Yahoo!. Fast-forward to today, it's pretty much the same. We license the traditional links from Microsoft, along with some other content, most notably ads, and then we do most of the modules ourselves, which means indexing, calling, as well as working with other partners like we work with Apple for apps, for example.

Q And when you say you license content from Microsoft, is that from Microsoft's search engine Bing?

A The Bing division within Microsoft.

Q Okay.

And can you explain for the Court how search results on DuckDuckGo differ from search results on Bing?

A Sure.

Another kind of a nonintuitive piece of context may be useful here.

So when you have a search results page, it's intuitive that the further down the page you get, the less things are going to be clicked on. So it might be the case that -- you might expect it to be the case that, you know, things on the top of the first page may be clicked on two or three times more than the bottom. But in reality it's more like 100X.

And the reason for that is the function that we see empirically is about one-half -- things right below the thing above it get about one-half the clicks. So if the top things gets, like, 100 clicks, the next thing will get 50, 25, et cetera.

So when you have -- with that context, if you take two pages that look to a casual observer nearly identical and you just flip the top two things, in practice, that's a much different search experience. And also from the pages that those things are linking to would get much -- a lot different traffic.

So with that in mind, we actually choose a lot of different search modules to put on top of the page so that creates a pretty distinct experience from Bing. And as I just mentioned, most of the search modules, we don't actually get from Bing ourselves so that makes it more different.

And then there's some other differences like the links that we get from Bing and the ads are just sometimes different. So we display less ads on our site than Bing.

Also because we're private, you know, the links themselves aren't changed around based on any profiling so that'll make some differences.

And then finally like visual treatments are different. So fonts, colors, layouts, that kind of stuff. Also subtle effects, there, can have huge differences on what people click on.

Q And with respect to the web index from Bing that DuckDuckGo uses, why is it that DuckDuckGo licenses Bing's web index rather than create its own?

A Well, initially as I mentioned, like we had limited resources and I wanted to focus on the -- what I considered to be the innovative portion which were these search modules.

Today, it's similar. We -- it's a very expensive proposition that we can't really afford.

But in addition, I'd say we get it from Bing with the ads. So in addition to links to monetize a search engine, you need ads. The ads itself have a very strong network effect which is coalesced into two search ad feeds.

And then we kind of get the result -- get those link results with the ads which is kind of further incentive not to invest in that when we can invest in search modules.

Additionally we invest in other privacy technology as well to make this privacy easy button that I was mentioning earlier.

Q And who does DuckDuckGo consider to be its search engine competitors?

A Really most of our users switch from Google. So like far and away, we consider Google just our competitor.

Other search engines that compete with us are Bing, Yahoo!, Ecosia, and Brave are probably the main ones in the U.S.

Q Does DuckDuckGo consider vertical search engines such as Amazon to be competitors that users are likely to switch to or from?

A No, not really.

Q And today, what is DuckDuckGo's percentage share of the overall U.S. search market?

A We're about two and a half percent in the U.S.

Q And which search engine has the largest percentage

of the U.S. search market?

A Google.

Q And does DuckDuckGo try to distinguish its search engine from Google's search engine in any way?

A Yeah.

I mean, in a word, it's privacy.

You want me to expand on that?

Q Go ahead, please.

A Sure.

I mean, so when you go online today, if you don't do anything to protect yourself, it's as if someone is, you know following you around everywhere you go, everything you do and recording that, really, every second.

All this information is then added together into, you know, what I would think of as a huge dossier about you or what we would call an online profile.

So if you use Google, that is, you know, your search history and most of your browsing history based on what you click on, and from their other products, location history, purchase history as well. So these profiles can basically be directly used to cause you harm.

So, for example, the ad targeting that is enabled by them have been tied to discrimination for just access to essential goods and services like jobs, housing, education, healthcare, even financial products.

Additionally, it can be used to charge you different prices as based on your profiles.

And then for Search in particular, I mean, there are many of these harms but relating to Search, as I mentioned earlier from Bing, you know, results may differ based on your profile. So your results may be manipulated, from our perspective at least, based on your user profile without you knowing about it which can bias your kind of conclusions from your search results.

So generally we find that a large percentage of Americans would like to avoid these harms. And so to distinguish that on DuckDuckGo, we don't do any profiling at all. In fact, we don't have any individual user, browser, or search histories. A way to think about it is every time you search on DuckDuckGo, it's like it's your first time or sometimes we say, you know, if 100 people search for cat pictures today, we don't really know whether it's like one person or 100 different people.

And then I would say also besides harms, just privacy gives you benefits. A lot of people from some research we do, do not like ads that follow them around the Internet all the time. And so if you use our browser and search engine, you effectively avoid what we call these kind of creepy ads, and that's just a more pleasant experience overall.

Q And do any of DuckDuckGo's search engine competitors offer by default the level of privacy that DuckDuckGo's search engine offers?

A No, I don't believe so.

I mean, there are some other private search engines to varying degrees, like Brave offers some privacy, Ecosia offers some privacy, but we don't think to the level that we do.

Q And does Google offer by default the level of privacy that DuckDuckGo's search engine offers?

A No.

Q Does Google offer, even if a user selects the most secure privacy options, the level of privacy that DuckDuckGo offers its users by default?

A I don't believe so, no.

Q And why is that?

A Because we've done several studies over the years, for example, this notion of changing results around when, based on your user profile, which we would call the filter bubble, because it kind of puts you in a bubble based on your profile and other results are filtered out -- we did two pretty deep studies of that in 2012 and 2018, and in that study we looked at people who were signed out, in incognito mode on Chrome, and found still individual results they were getting were very similar to the signed-in mode

relative to other people right around them that were not. And so we controlled for things like location, time of day, missing -- all variables we could think of and that was the only thing remaining.

Q So you described some studies that DuckDuckGo has done.

Has DuckDuckGo conducted any studies to evaluate search engine users' level of concern about Internet privacy?

A Yeah.

I mean, when we started, there really wasn't a lot of consumer privacy research. So I'd say just in the last, probably since around 2016 when we got really serious about working on this easy button, we've probably done over 50 large scale surveys of the American population.

Q And before you go on and tell us at a high level what those studies have found, can I ask you to describe how DuckDuckGo performs those studies?

A Sure.

We generally commission them with a survey panel company who then demographically adjusts them for the American population. We don't conduct them ourselves.

And, you know, we have our own user insights team who -- kind of professional survey people. And then sometimes we've also commissioned the entire survey from an

external research firm.

Q Okay.

And now back to the question. At a high level, what have these studies shown?

A So at a high level, it shows about 30 to 40 percent of Americans have a strong preference for search privacy. To give you some color to that, we often ask a question, if a new search engine were to come along, pick the top factor that would motivate you to switch search engines. And if you've already switched before, what was that factor.

So factors include, there's a little more precise language than this in the question, but they're essentially better results; better privacy; fewer ads; better connection to the social values of the company, like climate change; better design; better customization options; and an other category. So you might expect that better results would be the number one thing that people pick. That is not what we find.

Consistently, we find better privacy as the pick, the number one factor, about 40 percent of the time. Better results comes in at number two at about 25 percent. Fewer ads is at 15 percent. And then all the rest of the categories, so that ads up to 80, are about 20. So that's one data point.

The second one that I would give you would be -- because people -- often what we hear from that is, like, what people say or people do online may be different. So we also try to ask very specifically, like, what people are doing and recite back behaviors to us. So we ask things like, do they regularly clear cookies, do they -- have they ever downloaded a privacy app or extension. Have they ever subscribed to, like, a privacy app or service. And we have a number of these.

If people say they've done several of them, we put them in what we consider more our target market which is the care and act on privacy group, so they not only have expressed concern, but they've expressly taken actions that show that concern.

That comes out about right now at least around a third of Americans and it's been steadily climbing since we started doing this about ten years ago. So both of those triangulate around the same, I'd say between 30 to 40 percent of Americans showing a pretty strong privacy preference for search engines.

Q And has DuckDuckGo --

THE COURT: Counsel, I'm sorry, can I help interrupt?

We're a little bit past 11:00, so we're going to take our morning break. It's 11:00 -- a little after 11:05,

so we'll resume a little after 11:20.

COURTROOM DEPUTY: All rise. The court stands in recess.

(Recess from 11:06 a.m. to 11:22 a.m.)

COURTROOM DEPUTY: All rise. The Honorable Amit P. Mehta presiding. Again, in session.

THE COURT: Thank you, all. Please be seated, everyone.

All right. Mr. Hoffman, whenever you're ready.

MR. HOFFMAN: Thank you, Your Honor.

BY MR. HOFFMAN:

Q Mr. Weinberg, when we left off, you were answering the question of whether DuckDuckGo had done any studies about users' concerns about Internet privacy. Had you finished your answer?

A I think so.

I mean, I -- conclusion is about 30 to 40 percent of Americans express a strong preference for search privacy.

Q Okay.

Then let me ask you this: Has DuckDuckGo always been a privacy-focused search engine?

A We became privacy focused in around 2010. So about a year and a half after I started the company.

Q And in 2010 when DuckDuckGo began to emphasize Internet privacy, did that affect the number of searches run

on DuckDuckGo?

A Yeah.

I mean, early on, we had very little searches, just all around. And mainly early adopters are coming to us for a variety of reasons, including like the search modules I was referring to earlier.

Over that time period, privacy became the salient feature and the reason why most people were switching to DuckDuckGo, and it's just been increasing since.

Q Were there any events you observed that made privacy more relevant to consumers?

A Yeah.

I mean, so it was a -- it was more of a hidden issue that people weren't really talking about, I'd say before Snowden in 2013.

2013, that -- the Snowden revelations created an international press moment for -- really for an entire year, and that educated a lot of people about certainly government surveillance going on.

And then, really from then, awareness of privacy has ticked up slowly over time, over that entire decade, 2013 to now.

The other big moment where there was a really increased amount -- I'd say a lot of this was driven by, again, press -- was Cambridge Analytica in 2018. So that

really kicked off almost two years, up until the pandemic started, of nonstop privacy revelations in news -- you know, in the media.

Q And you referenced the Snowden revelations. At a high level, what were those?

A Edward Snowden was a whistleblower who released a lot of secret documents that showed a bunch about government surveillance programs.

One of the programs called -- which is relevant to this, called PRISM, was about the government requesting kind of direct access to big technology companies, including Google, so they could get -- easily search email history, that kind of thing.

Q And you referenced Cambridge Analytica and revelations there. Can you tell the Court what the revelations were in Cambridge Analytica?

A Sure.

Cambridge Analytica was more about Facebook, and some researchers had been using some Facebook data to profile American citizens and for political purpose and, arguably, influence the election.

I think it opened up people's eyes a bit wider, though, to kind of what user data was out there, how companies could get access to it through APIs, how it was kind of floating around.

And so since then, the coverage has been way wider as to what data is out there, how does it get leaked, how does it line up in other places, that kind of thing.

Q So we've talked about the level of privacy the different search engines provide. Let's talk now about the level of privacy that Internet browsers provide.

In your experience, Mr. Weinberg, do leading Internet browsers generally have privacy modes?

A I would say, yeah, most major browsers have a privacy mode. I don't necessarily believe it provides much privacy protection, but they have a mode.

Q And can you explain why you don't believe that most privacy modes offer privacy protection?

A Yeah.

I mean, historically when they started -- this is kind of another -- things were different a long time ago, right?

So when the private mode first kicked off, there wasn't as much online tracking going on. And so the mode was more about wiping your local history. And so if you used the private browsing mode and then closed the window, that history would be wiped such that, you know, your spouse or someone in your household couldn't see your browsing history.

But over time, the real privacy threats were more

online and websites tracking you behind the scenes. The privacy modes did nothing to prevent any of this tracking.

I'd say after many years of complaining about this and at least one class-action lawsuit, major browsers have started to put some privacy protection in privacy mode but definitely not enough to stop the privacy harms I was talking about earlier, and generally no search privacy in the mode.

Q And just for reference, do you know the name that Apple gives the privacy mode in its Safari browser?

A I think they just call it private mode.

Q Okay.

And the same question for Google's Chrome browser?

A Google calls it incognito mode.

Q And do you know the name that Mozilla's Firefox browser uses?

A I'm not sure if they have a name for it. I think it's just like new private tab, new private window.

Q Okay.

And you said earlier that DuckDuckGo also operates a browser. Does DuckDuckGo's browser have a privacy mode?

A Our goal is to offer privacy by default. So all our privacy protections are available in all of our windows.

Q Okay.

And when you say privacy by default, what does

that mean?

A I mean, when you install the product, you don't have to go into a privacy mode or mess around with settings to make the browser private, it just starts with the privacy protections on by default.

Q Now, Mr. Weinberg, is it possible that even if a browser is set to privacy mode, that a search engine could still track a user's private information?

A Yes.

I mean, oftentimes the browser and the search engine are different companies, and so they're not even coordinating. It just -- generally when you go into a private mode, the browsers aren't stopping online tracking. And so they wouldn't be stopping the search engine that you visit, which could be any search engine, from doing anything on their website.

Q And are there ramifications for the privacy of a consumer who pairs a browser set to privacy mode with a search engine that tracks personal information?

A Yes.

I mean, all the harms I was talking about earlier, a lot of them stem from searching and search privacy. So, you know, if you search for something in private mode, it could follow you around. Like we were talking about with ads, it could -- those ads could be discriminatory. Things

could still be added to your profile that could be used for changing your search results, like I was talking about, incognito mode. All of those, tracking could still be happening.

Q Earlier you described some studies that DuckDuckGo has commissioned. Has DuckDuckGo commissioned studies regarding consumer misconceptions about privacy protections offered by browsers set to a privacy mode?

A Yes.

Specifically we went pretty deep into this issue in 2016, and put out a detailed white paper in 2017, that was based on a number of surveys.

In general, we found that usage of these modes was very high. I recall something like 45 percent of people saying they used them at some point.

There was wide misconceptions about what they do. So we kind of asked people to check off what it did, and the majority of people could not identify what it actually did.

A primary misconception was this search privacy issue you're asking about where people thought their searches would be anonymous in private browsing mode, when they're not.

We then went on to ask people -- to tell people what it actually did and then ask people how they felt about it and to identify kind of from a range of emotions, and

people generally, again, the majority, felt kind of a range of negative emotions, confused, misled, that kind of thing.

Q And did DuckDuckGo study whether informed consumers would prefer to pair a browser in privacy mode with a search engine that does not track private information?

A Yes.

I mean, generally -- and that was part of the reason why we're doing this study in particular was, we had a hunch that that was the case. And, you know, we were trying to make the case to browsers to use us in private mode, and so that is what we found.

Q And for a consumer who prefers pairing a browser in privacy mode with a privacy-focused search engine, is it easy for them to switch device defaults to a search engine like DuckDuckGo?

A No. I mean, it's generally impossible. There's no -- there's generally no separate setting for a search engine in private mode versus regular mode.

Q All right. Well, let's talk about potential methods of extending the reach of DuckDuckGo's search engine.

Has DuckDuckGo sought to have its search engine included as an option on any choice screen menus?

A Yes, in Europe.

Q And what happened there?

A There was an antitrust case in 2018, and Google put forth their own remedy of a choice screen, which we -- we have no formal role in this process. We weren't kind of asked about it. The EC Commission didn't really include us in the negotiations. So from our point of view, it was kind of immediately flawed.

The first version, they wanted to charge us money to participate way more than the profits we make per user.

So that and many other things, issues with it, we've been complaining about.

The only thing that's really changed is it became free, but we think it's still not designed very well. But we participate now because it's free.

Q Okay.

Is DuckDuckGo a non-default search option in any browser?

A Yes.

Q And how effective is it for gaining traffic to be a non-default search engine in a browser?

A In general, just being made the option we found not very helpful because users find -- we can't very well instruct people to go to those settings.

There's one exception to that, which is there are browsers where -- and iOS is one of these -- where it is

impossible to set the search engine unless you're an option. And so that gives us some boost because otherwise, it was completely impossible.

But generally, we find just being an option doesn't give us that much.

Q Do you find any roadblocks that consumers have in resetting defaults?

A Yeah.

I mean, I -- so I'd say broadly, the search defaults are kind of the primary barrier for people expressing their search privacy.

They're -- like -- as I was just saying, they're extremely powerful. We find it empirically hard to get people to switch.

Q And could you expand on why it's difficult to get people to switch defaults?

A Sure.

So I think there's a few nuances to this.

First, it just -- as a practical matter, there's too many steps to get a full substitution of DuckDuckGo to Google.

This is because -- Google to DuckDuckGo.

This is because there's just so many devices and search defaults.

So if you think about it, a typical user is going

to have multiple devices they interact with across home, work, family, phone, tablet, desktop, laptop, somewhere between three to seven devices, I would say routinely.

Each one of these devices has multiple search defaults on it, the browser defaults and the operating system defaults, some people use multiple browsers, et cetera, sometimes there's multiple search results per device.

And so if you just switch one default, say, like, the Android widget, that's just a small fraction of the users' queries, that's not a full substitution of Google to DuckDuckGo, because they're still using and interacting with all their other browser and search defaults.

So to really get them to have a real substitution, a real substitute of products, you'd have to switch all the search defaults that they regularly interact with. That's now on the order of like 30 to 50 steps, not like three to five. And it's just too many steps. So, like, that's what we're seeing empirically.

We have survey data, we don't have any user data, but we do these surveys of the American population again -- we do a lot of surveys because we don't have any user data to see what usage DuckDuckGo gets and, like, how people are using it. We can segment those surveys by our most avid users. So these are users who love our search results, they

are recommending us multiple times a month to friends and family. This is like the top end of the range.

These people still aren't switching all their search defaults. They are routinely reporting to us that they're full-time on, say, mobile but not Desktop or they're full-time on one browser but not another. And so that's why I say kind of empirically, it just is not happening.

Two other nuances that might be worth mentioning is -- that we run into is, I think you just said, you know, it resets sometimes. So that definitely happens. So like on browser and operating system updates, the default can reset. Search functionality can also change which is kind of like a de facto reset. There's a new access point added or it's a little different, and so there's a justification to kind of reset it.

What this means essentially is, if you switch some of these defaults, eventually you're going to be forced back on to Google if you do nothing.

And then finally I'd say, it's all just way harder than it needs to be. So on Android, we find it takes currently 15 plus steps, there's more taps if you count all the taps, just to switch everything you can and not everything is switchable like the assistant, for example.

But it could be easy. So it could be the case. Imagine for a second that you had an operating system

setting that switched all the search defaults across the device in actually one step, and now imagine that if you downloaded our app or website, we could take you -- or you downloaded our app or went to our website, we could take you to that setting in one click. That's how easy it could be, but it just doesn't work that way.

Q Is that your choice that it doesn't work that way?

A No.

Q Is there any browser where DuckDuckGo is the default option today?

A I believe we're the default in the Tor browser, which is a privacy focused browser.

THE COURT: I'm sorry, can you spell that browser?

THE WITNESS: T-o-r.

BY MR. HOFFMAN:

Q And what percentage of the U.S. browser market does Tor have today?

A Some immeasurable small amount. I don't know if it's measurable.

Q Well, what have you found to be the most successful method of distributing DuckDuckGo's search engine?

A When we've been the default for brief periods in different times, that's been the best method.

Like in this Tor example, even, we weren't the

default and then we were the default.

Brave made us the default in some countries for a brief period of time.

And so we had kind of a taste of what that looks like. I'd say that's the best way.

Q And is DuckDuckGo search engine the default in the DuckDuckGo browser?

A Yes.

I would say that, as I was mentioning earlier, like, it's an all in one privacy product. So we believe search privacy is essential to get the value proposition of the product.

Q And are there any challenges in trying to distribute the DuckDuckGo browser?

A Yeah.

I mean, so, in fact, we made the browser in part -- in large part because of the friction in search defaults. So believe it or not we find it easier to get people to adopt our browser than to get people to try to convince them to switch to different search defaults.

The browser faces its own issues, though, like, for example, it's hard, in some case, to switch the browser default in a similar way.

Also, we face competition from Google's use of their other properties, like if you go to Gmail or YouTube,

they'll ask you to download Chrome constantly if you're on another browser.

THE COURT: I just want to add, just as a practical question, if I go to an app store and download DuckDuckGo, the DuckDuckGo app is both a browser and search engine or is it just search engine? Can you tell me what gets downloaded?

THE WITNESS: Yeah, I mean, I think that it's a browser, but it has the search engine built in.

I mean, all the browsers since about 2010 have that address bar where you just type, search right into the address bar.

THE COURT: Right.

THE WITNESS: So it's both.

THE COURT: It's both.

THE WITNESS: Yeah.

THE COURT: I wanted to confirm that. Thank you.

BY MR. HOFFMAN:

Q Now, Mr. Weinberg, without naming names and we'll get into that in the closed session later, has DuckDuckGo ever pitched the idea of being the default in a browser's private browsing mode?

A Yes.

I mean, we -- after doing -- we had the hunch about people's expectations there, we confirmed it in our

research. And so we thought it was a great pitch to browsers honestly where -- and the pitch was essentially, hey, we have direct evidence that we think you're misleading your users, like, your users think they're getting search privacy in this mode but they're not, you should really use a private search engine and we really took that pitch pretty far and wide.

Q And when DuckDuckGo made that proposal to different browsers, did it face any obstacles?

A Yes.

I mean, we generally saw a lot of interest. I mean, I think there was a lot of genuine concern about search privacy.

But pretty much with all the major browsers, at least, we hit an obstacle with their people's Google's contracts.

And we ultimately decided, this was after like three years of trying this, that it was a quixotic exercise because of the contracts.

MR. HOFFMAN: Okay, Your Honor, at this time, that concludes the questions that I think I'm allowed to ask in open court. I think my colleague from the states has some open court questions and then we could resume in closed session?

THE COURT: Okay.

MR. CONRAD: Good morning, Your Honor. Joseph Conrad for the State of Nebraska and for the plaintiff States.

THE COURT: Good morning.

BY MR. CONRAD:

Q Good morning, Mr. Weinberg.

First, if you could just satisfy my curiosity, I've been sitting there just curious.

What prompted you to name the company DuckDuckGo?

A Oh.

Unfortunately it was not a great story.

The mascot -- the idea for a mascot came before the name, wanted a cute mascot, tried to think of a name that would go with the cute mascot.

Q Got it. Thank you.

So you talked with my colleague, I think, about what you referred to as search modules. And could you explain where you get the content that fit into those modules?

A Sure.

I mean, as I was saying, there really are many dozens of distinct types of search modules, and each one has its own index, so if you want to get really complicated, set of indexes sometimes.

And so because of their vertical nature they're

going to come, each one is going to be pretty different as to what the content is going to be.

I referenced Apple for maps. You know, we partnered with Tripadvisor, some of them, like, FAQs -- some of them, like, Calculator, we'll just make ourselves. Something like info about people and places, we generally get from Wikipedia but that's from indexing ourselves, not from them directly.

Q And what motivated DuckDuckGo to first develop these features, or excuse me, those modules?

A Well, I actually started myself. The whole reason for starting the company myself was to improve my own Google results initially. And part of the initial motivation was actually going to Wikipedia a lot. So the first code that I ended up writing was crawling and indexing Wikipedia and making search models for Wikipedia and putting them on top of the search results.

Q And how did that strategy develop over time?

A We -- I thought it was the future.

We ended up really making that a core strategy of the company early on from, like, 2000 maybe -- let me get the dates wrong, '11 to '14 maybe. And we actually created a whole open source platform where developers could code all sorts of different search modules for us and the idea was to get pretty niche modules. We ended up with basically

thousands of these in the end.

All sorts of things, like, if you were into a certain programming language, you might have the programming language for that -- modules for that programming language documentation. Or if you were really into Pokémon cards or something, we could have a Pokémon directed search module.

It ended up not being as attractive a way to get people to switch search engines as I would have liked so we stopped focusing on it as much but that was kind of the strategy at that time.

Q And does DuckDuckGo continue to develop these modules?

A Yeah.

We -- I had mentioned earlier there's about 20 that I think are essential so we now focus our efforts on those essential ones and we continue to develop them.

Q And does DuckDuckGo do any testing or experimentation with regards to these modules?

A Yes.

Q And can you just explain for the Court what this testing or experimentation looks like?

A Sure.

I mean, there's not just one type of experimentation or test. We will -- I think the most common would be like an AB test or what they call a split test

where you have two versions of something going on at once and you're looking at what people are clicking on or engaging with on both versions and seeing which one is better by whatever endpoint you want.

We also will do other things, though. We'll do a qualitative test by having people walk through things and tell us what they're thinking or seeing. We'll do survey tests, like I was -- like, surveys like I was talking about before. There's a wide range of methodologies you could use.

Q And does the number of users that DuckDuckGo impact the quality of the testing or experimentation it conducts?

A Yeah.

I mean, we are very limited in the amount of on-site experimentation we can do, which is partly why we use all these other methodologies to try to be a proxy for what we could get if we could run more AB tests live on the site.

Q Thank you. Those are my questions.

THE COURT: Can I ask one question, one additional question, which is -- and you may have answered this, I apologize.

In terms of search quality and the quality of results, what does DuckDuckGo do with the data it receives

from Bing to create its own search results page?

THE WITNESS: So --

THE COURT: And to distinguish its search results page?

THE WITNESS: Yeah.

So we're only getting some of the information that we need to make a search result page from Bing. We're mainly getting the traditional links and ads. So all these search modules we have to do ourselves.

And so we have to develop -- and I'd say a large part of the technology is developing a relevancy engine of where to place what on the page. Like, do you put maps on top of the links here, or should you put it under the third link.

And if you don't put the thing that people want on top, because of the way people click, so much -- and it gave so much more than the thing on top, you have a perceived quality issue.

I would also say, though, that, like, I don't believe there's an overall metric for quality. I think it's -- each query is kind of subjective to some degree, and, you know, each vertical has its own metrics. So it's not like we're -- we can make a topline measurement of it.

THE COURT: So if somebody were to enter a query that would fall outside the modules that you've referred to,

the 20 or so that you focus on, what would the search results be and where would they come from?

THE WITNESS: Yeah.

I mean, I would say they're mainly from Bing at that point.

Most searches, I think, now have modules on them. And so, like, if you -- say my Google and Bing, if you search almost any query, there's going to be other stuff on the page. So that's becoming a more rare experience.

But if you just have traditional links, we would be getting those from Bing.

There are some changes, like I was saying, that might be different. The ones that Bing sends us and what they would display for themselves might be different because of, say, like privacy.

There are also some other differences that there might be, we have some flexibility around that.

THE COURT: And when you say "traditional links," I'm not -- can you expound on what you mean by that?

THE WITNESS: Yeah.

Otherwise referred to as kind of the 10 blue links. I think those are the links that people think about when they think about search engines. They don't often think about the modules, which is why I was talking about them.

And so I just mean like regular links to websites.

THE COURT: Right. What we've been referring to as sort of organic --

THE WITNESS: Organic links. Yeah, that would be a good term, yeah.

THE COURT: All right.

All right. Thank you. I appreciate that.

All right. So why don't we go ahead and then move into closed session, and we'll ask if you are not associated with any of the parties or with DuckDuckGo, to please exit from the courtroom. Thank you.

(Sealed closed session)

THE COURT: All right, Counsel. It's going to take a -- we've got the courtroom connected to a media room downstairs. That's how -- that's about how sophisticated we are with technology. So it's going to take us a few minutes to get that line disconnected before we can proceed.

THE WITNESS: Got it.

(Pause)

THE COURT: We're still waiting. It's going to take a couple minutes.

MR. HOFFMAN: Sure. Thank you, Your Honor.

(Pause)

MR. HOFFMAN: Your Honor, may I approach the witness with a binder?

THE COURT: You may.

MR. HOFFMAN: May I approach the bench, Your Honor?

THE WITNESS: Can I look at this?

THE COURT: He'll direct you to certain documents when he's ready.

(Pause)

THE COURT: Sorry it's taking so long.

We're ready to go.

MR. HOFFMAN: Okay. Thank you, Your Honor.

BY MR. HOFFMAN:

Q Mr. Weinberg, when we left off, I asked you if DuckDuckGo had ever pitched to different browsers that DuckDuckGo could be the default in their private browsing mode.

Now, specifically let me ask you if DuckDuckGo ever proposed to Apple that DuckDuckGo could be the default in Safari's private browsing mode?

A Yes, many, many times.

Q And when did DuckDuckGo first make that pitch to Apple?

A So we had been trying to become a search option in Safari first. That took, itself, several years, and that happened in 2014.

As I mentioned before, if you don't have a

contract with Apple, you can't be a search option. So we got that first search contract in 2014.

Shortly after that, I believe even in 2014, we started pitching this idea for a DuckDuckGo search and private browsing mode.

We did it in 2014, '15. I think in 2016, we got our first kind of response back.

Our take was that they were actually really interested in this.

The people we were talking to were generally DuckDuckGo users themselves interested in privacy.

Q Just so it's clear, the people you were talking to at Apple were --

A At Apple, yeah.

So at that time, this was mainly Rhonda Stratton, who became our partner, manager, at Apple, and Brian Croll, who was an executive in product marketing. Those were our kind of primary high-level contacts.

They both were DuckDuckGo users, deeply interested in privacy, thought this was a good idea.

From those conversations, we kind of identified Craig Federighi was the executive, the main -- highest level project executive.

THE COURT: I'm sorry, what was the name?

THE WITNESS: Craig Federighi.

He's -- I'm not sure exactly their titles. He's like an SVP on their main leadership page.

They -- so we tried to push to get a meeting with him, really to pitch this.

That did occur in 2017.

Q And where did this meeting in 2017 take place?

A It took place -- most all our meetings took place, including this one, at Apple headquarters because, you know, that was before Apple Park so it was like in the main Cupertino place.

Q Okay.

And do you remember what month of 2017 this meeting took place?

A Yeah, it was in October.

Generally we had been coming back once a year around that. So they have a yearly cycle of their -- of releasing iPhones and iOS and Mac iOS. And so it comes out in September. They announce things in June. They think about what they're going to do in the fall. And then they hurriedly program it in the spring.

And so we tried to come back in the fall and kind of make our pitch at the time when they were thinking about roadmapping for the next cycle.

Q So now we're talking about a meeting at Apple headquarters in Cupertino in October 2017. Can you tell me

who you recall notable from Apple that attended this meeting?

A Yeah.

I mean, it was a big meeting. We were in a really big conference room, so we did get the meeting with Craig.

There were other -- a bunch of other executives there. Bud Tribble I remember being there. He's a very longtime Apple person. I think he was on, like, the original Macintosh team. So he's kind of like a senior fellow executive type.

The head of Safari was there at the time, Darin Adler.

Brian Croll was there. Rhonda, who I mentioned.

I believe there were some Siri folks there that we never saw again, so I don't remember who they are.

And some people from privacy, too.

Q And what do you recall the substance of this October 2017 meeting at Apple headquarters to be?

A So --

MR. SAFTY: Your Honor, objection to the extent this is eliciting hearsay testimony.

THE COURT: I think he was just asking generally what the subject was. I don't think he asked what anybody said.

Go ahead.

THE WITNESS: Yeah, so we were -- we essentially wanted to level up our pitch because we hadn't been successful yet, and so we were hoping to make the most compelling pitch we could to Craig.

So we did -- we basically did two things in our pitch that we hadn't done before.

One is, I mentioned -- you had asked me previously about private browsing research we did. So part of the motivation for that was for -- for these and other meetings like this.

And in particular, we segmented the data by Apple users so we could specifically tell Apple what their users thought about private browsing mode, what they expected from it, you know, do they want it, private search for private browsing, et cetera, and that data was pretty compelling from our perspective so we put that up.

The second part of our pitch was, we thought we could, in partnership with them, make DuckDuckGo the best search option on Apple devices for Apple users by integrating their content into search modules. And so, for example, in particular, we had pitched Apple News, Apple Maps, Apple Music, Apple TV.

And if you think about like maps for a second, if you go on Google, you get Google Maps. If you're an Apple user and like Apple Maps, you would prefer to see Apple Maps

for a couple reasons.

One is, when you click into the Apple Maps, it's going to be seamless, it's going to look the same. And it's not just that it looks the same, it's like different directions, points of interest can differ. So it's just a much more seamless appearance if the search modules match the apps you're going to click in.

And same for music, TV, news, et cetera. So our general pitch was, Work with us, let's incorporate all of Apple content into the search engine, and it will be a more compelling search experience overall.

BY MR. HOFFMAN:

Q And what was your impression of how this meeting went?

A My impression of the meeting is it went very well.

I mean, so I've pitched lots of things to Apple over the years. If there's no interest, their move is basically silence. I mean, there's lots of awkward silences. And so you have to, like, you know...

So this case, I actually had to move things along because there was so much engagement. The second half of the pitch about the content, I remember having to, like, let's go talk about this because there was so much talk about the private browsing.

Craig in particular kind of led discussion from

Apple's side, but, yeah, there were lots of discussions around implementation, what implementation could look like, about DuckDuckGo itself that people in the room weren't totally aware of, like some of those questions like the judge was asking me, like how exactly the results worked, the search modules, et cetera.

Q Okay.

And this October 2017 meeting that we were talking about, did it lead to another meeting with Apple to discuss options for Safari's private browsing mode?

A Yeah.

So I would say we -- it kind of accelerated everything.

Our general idea was, you know, we didn't think Craig was going to be involved in everything day to day because he had so much to do, but if he gave kind of blessing to the team to go and figure this out, we could then progress with the rest of the people in the room. That's essentially what happened.

So we came back the next summer and --

Q And when you say "back," is that back to Cupertino?

A Yeah, back to Cupertino.

And there had been talk about implementation. And so we flagged that as like we thought that was going to be a

primary concern of theirs.

So when you say private search or private browsing mode, that could mean a lot of different things the way it looks and how users interact with it.

So we came back with a presentation that offered several alternatives to give them some ideas. Apple, I would say, is also very visual. So they really want to see what it's going to look like. They want to imagine what it looks like in their product, so we tried to design it fully, like take the full screens as to like something they would do.

And so we offered several ideas for this, like a drop-down, like we're the default but there's a drop-down where you can change it, or you could toggle something that said, like, you want to turn on private search or private browsing when you open private browsing mode.

We offered one that was, like, you could activate it. So there was, like, a message about private search educating consumers, and if you click "activate," it would take you to a setting that you could toggle.

So we presented all of this basically at the meeting.

Q And I didn't get a chance to ask you but do you remember any notable Apple executives at this meeting in the summer of 2018?

A So the crowd for this was, again, Rhonda and Brian and Darin, now, kind of picked up as another executive that we worked on this with, also expressed to be a DuckDuckGo user who was running Safari at that time.

Q And when you say Rhonda, that's Rhonda Stratton?

A Rhonda Stratton, yeah.

Q Darin is Darin Adler?

A Darin Adler, Rhonda.

I'd say at all of our meetings, there were other Apple people there often, like, sometimes somebody from privacy, sometimes somebody from legal. Sometimes -- because we also discussed things about our other privacy technologies, so sometimes there's like API or extension people there. Like, I remember John Davis, I think, might have been at this meeting.

Q And we're talking about this summer 2018 meeting. What was your impression of how that meeting went?

A My impression from this point on -- or our impression was that they were really serious about a potentially for the next year's release.

So this is 2018. And part of the reason I think that is we had been going back once a year basically, you know, to Apple headquarters.

But Darin asked us in the meeting to come back basically as soon as possible to brainstorm what privacy

integrations could look like, including this one, kind of a full integration of DuckDuckGo technology into Apple Safari.

And they also made references to roadmaps and things like that.

Q And you mentioned a presentation you prepared for Apple for this summer 2018 meeting. Did DuckDuckGo send that presentation to Apple after the meeting?

A Wouldn't have been me who did it, but that was our protocol, was to send all our materials to the external party after the meeting.

Q And, Mr. Weinberg, I'll direct your attention to the tab UPX0818 in your binder, and I'll ask Mr. Barkey to display UPX0818 on the screen.

MR. HOFFMAN: And, Your Honor, we have DuckDuckGo's 902(11) declaration for this document but I understand Google has a hearsay objection.

THE WITNESS: I don't see it in the binder, but I see it on the screen.

MR. HOFFMAN: Okay.

MR. SAFTY: No objection to admitting this exhibit.

THE COURT: So UPX818 will be admitted.

MR. HOFFMAN: Okay. Thank you, Your Honor.

THE COURT: I'm sorry. Do you not have it in your binder?

THE WITNESS: Would it be the tab?

THE COURT: It should be behind the tab, yes.

THE WITNESS: Yeah. I don't see 818.

(Plaintiff's Exhibit UPX818 received into evidence.)

THE COURT: Can you all just check and make sure he's got the exhibit in front of him. It's on the screen but...

THE WITNESS: Now, I do see it. It's hidden behind another tab. I've got it. My bad.

BY MR. HOFFMAN:

Q Okay, great. And you can use the binder or you can use the monitor. I'm guessing with your profession, you may be more interested in the monitor.

A Yeah.

Q Do you recognize this document?

A Yeah.

Q And can you tell the Court what it is?

A Sure.

This is kind of what he says. This is a summary of our meeting sent by Prakash who was leading this media engagement -- business development engagement at the time.

Q When you say Prakash, what is Prakash's last name?

A Swaminathan, I think, is how you pronounce it.

Q And attached to this email, was there the

presentation that you described you prepared for Apple at the summer 2018 meeting?

A Yes.

Q So now I'll direct your attention to page 3 of this document. It has the Bates number ending in 552.001.

A Yep.

Q Do you recognize this document?

A Yeah. I mean, on the screen is Apple logo and the DuckDuckGo logo. We generally put this up at the beginning of a presentation, you know, to be like let's work together kind of thing.

Q Okay.

And then I'll ask you to flip to the next page which is Bates number ending 552.002. I'll ask you if you recognize this page.

A Yep, same thing. We put this at most of our presentations to just reinforce that we feel similarly about privacy.

Q Okay. Thank you.

And, finally, I'll ask you to turn to page 11 of this exhibit which is 552.009.

And I'll ask you if you recognize these depictions.

A Yeah, these were the different implementations, ideas I was referencing. And then we kind of ended with the

one we thought might be most appealing to them where we -- this is the way their private browsing mode looked at the time. And so they had this explanation of it at the top and so we would, in this idea, we would add this private search piece underneath.

Q Are you talking about the phone screen on the left right now?

A Yes, phone screen on the left, where a user then could understand, try to understand their search is more private than private browsing mode unless they activated private search and private browsing. If they clicked "activate," it would take them over to the OS operating system search setting where there would be this toggle that they could activate it.

Q Can you describe for the Court what the phone screen on the right depicts in this exhibit?

A Yes, that is the Safari search engine setting screen with our addition to it at the bottom where it says private browsing mode, where you would be able to enable private search and private browsing mode.

Q So was this the way Safari appeared at the time of your presentation or was this your proposal to Apple?

A This was a proposal to Apple.

And I would say that we weren't -- we never proposed them to use exactly what we did. I mean, this is a

depiction of what they could potentially do. We're trying to give them visual ideas that they could take back and think about how they wanted to do it.

Q Okay. Thank you. You can put that exhibit aside.

THE COURT: Can I ask you a question?

Why was it designed this way or why did you propose a design to, it looks like it would actually require the user to activate it as opposed to having it already be set as the default in privacy mode?

THE WITNESS: We had already experienced some conversations around some understanding that they wouldn't do a default in private browsing mode because of their Google contract. And so we were already trying to get something that would work.

THE COURT: Thank you.

BY MR. HOFFMAN:

Q So we were just talking about the presentation that was prepared for the summer 2018 meeting. Did that summer 2018 meeting with Apple lead to another meeting with Apple to discuss Safari's private browsing mode?

A Yes.

So I mentioned Darin had asked us to come back pretty much immediately and brainstorm how our integrations could work. And so we did that as soon as we could prepare for it. So we came back about a month and a half later, it

was in September of that year.

Q Okay.

So now we're talking about a September 2018 meeting that, again, is in Cupertino; is that right?

A That's correct.

THE COURT: I'm sorry, '18 or '19?

THE WITNESS: '18.

THE COURT: '18. Okay.

BY MR. HOFFMAN:

Q And what do you recall the substance of this September 2018 meeting at Apple's headquarters to be?

A So we -- it was really about -- you know, they had asked us to give them all their ideas for what -- or give -- present ideas for how DuckDuckGo privacy technology could be integrated into Safari and especially private browsing mode.

So we also produced a lot of other privacy technology that kind of blocks different hidden trackers across the web.

And if you think about when I -- when you originally asked me about how DuckDuckGo works and we have are these different privacy protections and they're on by default, you know, the thought was if they offer this privacy mode, maybe they could offer a lot more privacy protection in addition to private search in this mode.

And what of all of our technology might they

integrate.

So we came with a bunch of ideas, leading with the private browsing one -- private search and private browsing one, I mean, but also thinks about using our tracker blocking list to kind of block hidden trackers on private mode. Increasing encryption is another thing that we do. And so we covered all that.

And there was -- I seem to remember this meeting because it was Darin's meeting and his team is mainly all engineers, and so there were a few non-engineers there but it was stacked with engineers who are just -- who talk very differently. I mean, I'm used to it, but not normally Apple speak, I would say. The more frank, more engineering focused, that kind of stuff.

Q And, again, we're talking about this September 2018 meeting at Apple headquarters.

What was your impression of how this meeting went?

A We were very excited about this meeting.

I mean, they expressed that they were -- it was the first meeting where anyone had expressed they were really considering this for the 2019 release, not just like going to talk about it at a roadmap meeting but they're trying to put it on the roadmap. They're interested, they want to do it, which had never happened before.

So we were coming out very excited.

The -- I'd say they hadn't figured out the design yet so we're still presenting the designs and talking about that.

And they hadn't figured out, even with this design or the designs they wanted to do, which they didn't really tell us, how it would work with their Google contract.

Q Okay.

So you said you were excited. Despite the excitement, did you leave the meeting thinking that Apple had any concerns about implementing your proposal?

A Yeah.

I mean, the explicit next step was, I mentioned about the contract, was they were going to -- they had identified pretty clearly that the people in the room could not resolve this issue around their Google contract and that the next step was they were going to go figure that out.

Q Okay.

And you mentioned the Google contract a few times. What do you understand to be the Google contract?

A I understand what I'm referring to --

THE COURT: I'm sorry. Objection.

Basis?

MR. SAFTY: To the extent it's calling upon the witness to speculate about things that are beyond his personal knowledge.

THE COURT: Well, I think he was asked about what his understanding is. He wasn't asked to speculate beyond that, so...

MR. SCHMIDTLEIN: If he's seen the Google contract, that would be shocking.

THE COURT: Well, I suspect he's not going to quote the text of it, but has a general understanding of what its terms are.

THE WITNESS: I have not seen the Google contract.

What I'm referring to is -- I'm sure the contract is even wider -- but the search portion of the contract and allowing things to be or not be default.

BY MR. HOFFMAN:

Q Did DuckDuckGo -- and, again, we're talking about the September 2018 meeting.

Did DuckDuckGo prepare a write-up of the September 2018 meeting with Apple?

A You mean like internally?

Q Yes.

A Yeah.

I mean, we would typically make a write-up for all major external meetings. So the person who's kind of in charge of the meeting usually -- we have multiple people go in, so somebody there is actually in charge of listening and taking notes, and they would write up what happened, next

steps, all that kind of stuff.

Q Was this a standard practice at DuckDuckGo?

A Yes, standard practice and -- and now.

Q Okay.

And what is the purpose of preparing these write-ups?

A Multiple reasons.

Make sure everyone's on the same page. Make sure we capture all the action items. Make sure people who aren't there can be up-to-date on what happened because a lot of people, like, for example, this, participated in this presentation, couldn't actually go.

Q Okay.

And in what format are these write-ups prepared?

A We use a software, a piece of software called Asana, which is a project management tool that we organize kind of the whole company around.

We don't -- we don't really use email. So anything you think would be in an email would be on Asana instead. It's got threaded conversations. It's not chat, we have that too, but this is more for permanent record and conversation that everyone can see.

We're a distributed company, so it's important that we write everything down, and people are in all different time zones and most things are async in this

company.

THE COURT: When you say you are a distributed company, what do you mean by that?

THE WITNESS: We have been -- we've never -- we have this mailing address, and you could call it our central office, but we've been -- all of our team members have been around the world since day one.

THE COURT: I see. Okay.

BY MR. HOFFMAN:

Q Okay.

Now, Mr. Weinberg, I'd ask you to turn in your binder, or to look on the screen, at UPX0666.

MR. HOFFMAN: And I'll ask Mr. Barkey to display that.

BY MR. HOFFMAN:

Q And I'll ask you, Mr. Weinberg, if you recognize this document.

A Yes, this looks to be the contemporaneous meeting notes we're talking about from that meeting.

Q The September 2018 meeting?

A Yes.

Q Okay.

MR. HOFFMAN: And, Your Honor, we have DuckDuckGo's 902(11) declaration for this document as well. And I understand that Google has an objection.

MR. SAFTY: Your Honor, we object to the introduction to the extent that plaintiffs intend to rely on the embedded hearsay for the truth of the matter; specifically, statements purportedly made by Apple employees.

THE COURT: Okay. Well, do you intend to point him to any out-of-court statements?

MR. HOFFMAN: I do, Your Honor.

At the bottom of page 1, there's a paragraph that begins with "Private browsing opt-in," and the statements in here we're not offering for the truth of the matter but to corroborate the conversations that DuckDuckGo had with Apple, and to give a complete picture of the negotiations, and then to show the effect that they had on Mr. Weinberg and DuckDuckGo going forward.

MR. SAFTY: Your Honor, if I may be heard on that point? They are offering it for the truth of the matter. They're offering it to corroborate the verbal hearsay that he's been testifying about for the last 20 minutes.

THE COURT: Well, he hasn't actually testified about any hearsay, Counsel, let's get that straight. He hasn't offered any statement from out of court that would qualify as hearsay.

It's 12:25. Let's take our lunch break so I can think about this. We'll return at 1:30.

Mr. Weinberg, during the lunch break, I'll ask you not to discuss your testimony with anyone other than speak to your counsel, of course, and we'll see everybody at 1:30.

Thank you.

COURTROOM DEPUTY: All rise.

This Court stands in recess.

(Recess from 12:25 p.m. to 1:30 p.m.)

C E R T I F I C A T E

I, William P. Zaremba, RMR, CRR, certify that the foregoing is a correct transcript from the record of proceedings in the above-titled matter.

Date: September 21, 2023

William P. Zaremba, RMR, CRR