BEFORE THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF COLUMBIA

UNITED STATES OF AMERICA, et al., .
. Case Number 20-cv-3010
Plaintiffs, .
.
vs. .
. Washington, D.C.
GOOGLE LLC, . October 18, 2023
. 9:32 a.m.
Defendant. .
- - - - - - - - - - - - - - - - - -

TRANSCRIPT OF BENCH TRIAL, DAY 24
(MORNING SESSION)
BEFORE THE HONORABLE AMIT P. MEHTA
UNITED STATES DISTRICT JUDGE

APPEARANCES:

For DOJ Plaintiffs:
KENNETH DINTZER, ESQ.
United States Department of Justice
1100 L Street Northwest
Washington, D.C. 20005

DAVID DAHLQUIST, ESQ.
United States Department of Justice
209 South LaSalle Street
Suite 600
Chicago, Illinois 60604

For Plaintiff State of Colorado:
STEVEN KAUFMANN, ESQ.
Colorado Department of Law
Consumer Protection Section
Antitrust Unit
1300 Broadway
Seventh Floor
Denver, Colorado 80203

-- continued --

APPEARANCES (CONTINUED):

For the Defendant: JOHN SCHMIDTLEIN, ESQ.
KENNETH SMURZYNSKI, ESQ.
Williams & Connolly LLP
680 Maine Avenue Southwest
Washington, D.C. 20024

Official Court Reporter: SARA A. WICK, RPR, CRR
333 Constitution Avenue Northwest
Room 4704-B
Washington, D.C. 20001
202-354-3284

Proceedings recorded by stenotype shorthand.
Transcript produced by computer-aided transcription.

C O N T E N T S

TESTIMONY

EXHIBITS RECEIVED

P R O C E E D I N G S

(Call to order of the court.)

COURTROOM DEPUTY: Good morning, Your Honor. This is Civil Action 20-3010, United States of America, et al., versus Google, LLC.

Kenneth Dintzer for the DOJ, William Cavanaugh on behalf of Plaintiff States, John Schmidtlein on behalf of Google.

THE COURT: All right, everyone. Good morning. Nice to be with you all again.

All right. Before we get started, I wanted to just check in and see where the parties were in terms of the submission that *The New York Times* made last evening and your thinking and preparedness to discuss the proposed changes.

MR. SCHMIDTLEIN: Your Honor, for Google, we would like an opportunity to make a written submission in response to that. And I think we will be in a position to file something later today on that, if that would be okay with the Court. But we would like the opportunity to respond in writing to it.

THE COURT: Sure. Of course.

MR. DAHLQUIST: Thank you, Your Honor.

We've reviewed -- we do not feel a need to file a written submission, but we are prepared to address it at your convenience today and go through step by step, if that's your preference.

THE COURT: Okay.

MR. DAHLQUIST: I will say at the outset that we -- our belief is that the order, Your Honor's order as it exists, could stay as it is today and that there's no need for modification.

That being said, if Your Honor is interested in some of the proposed modifications, we're happy to tell you our views on each one.

THE COURT: Okay. And just in terms of timing, Mr. Schmidtlein, when you say you expect to file something today, are we talking about during the business day? Later this evening?

MR. SCHMIDTLEIN: I think it's going to be later this evening, because I've got people here in court that need to --

THE COURT: Fine. Just in terms of notifying counsel for *The New York Times* about when he may need to appear to discuss it.

Let me just -- in the interest of sharing my thoughts on where things stand, I just want to point out a couple of key cases from the D.C. Circuit that I think everybody ought to be aware of, if you're not already, and how it affects some of the requests that have been made by *The New York Times*.

I think most pertinently, everybody ought to take a look at *In re: Reporters Committee for Freedom of the Press*, 773 F.2d 1325 from 1985. It's a decision that Judge -- then-Judge Scalia wrote. And the factual posture of it is interesting, because it

was a case in which the trial court essentially sealed all the trial exhibits and declined to make them available to members of the press until the trial court ruled on post-trial motions and entered judgment. And in fact, the trial court in that case actually entered judgments notwithstanding the verdicts of the jury in favor of the plaintiffs in that case.

The circuit considered what was the request of the reporters to get essentially contemporaneous or nearly immediate access to the exhibits during trial and prior to the entry of judgment. And what the circuit held is that there is no First Amendment right to access of civil records until the entry of judgment. That was the holding of the D.C. Circuit.

And what's notable about the case is a couple of things. One is that you'll see toward the tail end of it, the Court talks both about the First Amendment right and clearly says there's no First Amendment right, there's no -- under the First Amendment test, which is sort of a history and tradition test, the Court says there's no history and tradition of getting immediate access.

The Court then sort of looks a little bit at the common law issue. It's a little unclear on which it's being raised. But the bottom line is that the Court says that with respect to the access under the common law, that what the trial judge did in that case was not problematic. And what the Court essentially does is -- I'll hold my comments to myself -- recognizes the

challenges that a trial court faces in these sort of circumstances.

And in fact, the facts of that case were that the confidentiality issues were limited to, I think, 600 documents that were being considered during the discovery phase. And the question arose as to whether -- because the dissent had taken the position that the documents needed to be made available unless there was a document-by-document justification for the confidentiality designations.

And the circuit said that's not required. And even if it was required, it would not be required for the Court to rule on those designations. And the Court, as appellate courts don't always do, took a very practical approach to this and recognized that there are other demands on parties, lawyers, and judges throughout a trial. And those include, most importantly, running the trial and moving forward with the evidence and making sure that's done in an efficient way.

And almost -- maybe it's prescient, maybe it wasn't, but -- or maybe it's just because Judge Scalia had an affinity for antitrust law. This was not an antitrust case. And I will just note toward the very end of the opinion, in the last paragraph, he writes, "It would have been excessive to expect in addition at that stage of the proceedings the crafting of a document-by-document specification of basis for the claim that would sustain a post-trial legal attack." And then he writes,

"And the difficulties encountered in the present case are as nothing compared with those that a major antitrust trial would present."

So, "In sum, the dissent's made-for-the-occasion categorical rule of instantaneous document-by-document justification is utterly infeasible, and any feasible rule, which would have to accord the district judge a reasonable degree of discretion, could not conceivably have been violated here where the document-by-document justification was required within 30 days after the conclusion of trial."

So everybody ought to be aware of that as we move forward in considering the proposed amendments that *The New York Times* has recommended to the order that I issued mid-trial.

One thing I can say about what *The Times* has requested, and I don't need their counsel to be here, I am now convinced that with respect to their one request, I am correct that I do need to give notice about the closing of the courtroom.

In fairness -- well, not in fairness. This isn't something that happens often, and I will confess that it was not something that -- I appreciate *The Times* bringing that to my attention.

And after I looked at the cases they submitted, I think it is fair to make the request of the parties that for future witnesses going forward, that I get notice by the end of the day to the day the witness is expected to testify, prior day, whether there's expected to be any closed session. And if there

is, I will post that on the docket, and if any interested party -- or, I should say, if there's to be an objection to the closing of the session, we can take that up at 9:30 the following morning. And so we will follow that procedure from here forward.

That's part of the reason we e-mailed everybody last night about today's witness, and the parties have advised that the expectation is that there will be no sealed portion. Certainly, Google has said that it is not intending to examine in its direct examination the witness in a closed courtroom. And I believe plaintiffs are certainly expecting not to have the courtroom closed, but of course, it may depend upon the direct examination.

Mr. Dintzer?

MR. DINTZER: Could I rise to address the Court's point? Thank you, Your Honor.

So just to set the stage, we're moving into a subject that we haven't really dealt with. So we don't have as much experience in exactly what the defendants will say is confidential and how the Court will rule. We're going to be looking at, as I understand it, how Google's search mechanism works, among other things.

With Mr. Nayak being moved up, which of course we were happy to accommodate and there's no issue with that, but it meant that we were somewhat behind the ball where we usually are

in running documents through the confidentiality process. And so we sent them last night the documents for them to take a look at.

What I would propose, Your Honor, so that we save the Court's time and so that we do our best to avoid a closed courtroom that might otherwise be necessary, is that after the defendants are finished with their examination of Mr. Nayak, we take a break and we talk to them and see if there are issues, see if there are issues we need to raise to the Court, and we get -- we work through those before we begin our cross so that we can hope to get as much as or all of in the public sphere as possible.

THE COURT: Okay. I guess the question is whether at this point Google, based upon what you all have sent over, believe that there will be a -- will be requesting a closed session based upon the anticipated cross.

MR. SCHMIDTLEIN: Your Honor, Mr. Smurzynski is going to handle this witness, but they sent us documents at 10:00 p.m. last night. So we have not run those documents through -- they are in violation of the order, the 48-hour order. I'm not sure what Mr. Dintzer is saying when he suggests they just got notice of Mr. Nayak. They've had notice for two weeks that we're bringing this witness on this date.

And so we have not had the opportunity to run this through the various people who would -- they would need to be reviewed

for the sort of confidentiality analysis.

So that's where we are.

MR. DINTZER: Your Honor, first of all, we're in violation of no order, of course, because the Court never ordered us to do that. It was the process whereby we would have things cleared of confidentiality, which we have been doing, and quite honestly, once the documents were selected and analyzed and we decided to use them as a part of the exam, then we sent them to them.

There was no -- which is why I'm asking for this time in between, so that we can go to them, talk about the lines that we're -- and also, it will help us to see, based on what they ask him, what -- there's some terms that -- anyway.

THE COURT: I guess the bottom line is, let's just see where we are. My understanding from Mr. Schmidtlein was that the expectation was that the direct will go through lunch.

Is that right?

MR. SCHMIDTLEIN: I think close.

THE COURT: So we'll likely be at a place where we will have a more extended break than the usual 15 minutes that you all can meet and confer about that, and hopefully, there will be enough time to do so. If not, we will just have to figure out how to proceed.

MR. DINTZER: We appreciate that, Your Honor.

THE COURT: Okay. All right. Great.

I was going to save a lot of this for later, but I feel compelled to say one more thing about not this issue but what I was talking about before. And I think it's underscored in light of the case that I just identified. And that is, I want it to be understood what I have asked the parties to do in this case and what I think is unusual.

And that was, we, at the very outset -- when I say "outset," I mean back in December of 2020 when the protective order was filed in this case -- understood that confidentiality would be an issue throughout the proceedings and would be an issue at trial. The protective order that was put in place understood that and actually planned for it, and it required the parties to make a proposal about how to deal with confidential records and propose something to me in advance of trial.

That happened. In fact, we had extended discussions about it at our monthly status conferences. I gave my thoughts to the parties about how to handle this matter. And it was then entered into an order. And I can identify the number, but it was entered into an order that specified the exact process for the treatment of confidential information.

I'm not being critical of anybody, but there was no objection raised by anyone to that process, and that process was on the public docket, and it was in place for approximately four weeks before we began. And so it should have come as -- it should not have been a surprise the way we have proceeded,

because it was actually spelled out, largely spelled out, in that order.

And what that order said was the following: One is that the parties would work in good faith toward ensuring as much of these proceedings to be open to the public. That's one.

Two, I asked the parties to do something that I thought would facilitate that and, I think, has largely facilitated it. And that was to not just prepare their exhibit lists but to, in advance of a witness coming to the stand, identify portions -- not only the exhibits they intended to use with the witness, but the portions of the exhibits that they actually intended to use to examine the witness. And the reason for that was so that the side that had the confidentiality interests, whether it be Google or a third party, could say in advance whether they had confidentiality concerns about that portion of the record being presented in open court.

The parties did that, and that, I will tell you, is an extra layer that I asked them to do, and they did do, that has required a fair amount of work. It not only requires lawyers to identify the exhibits, but it requires them to think about what they want to present and then do it in a way that's timely to allow the other side notice and give them notice about what they might think is a confidential issue.

And we've operated in that way, and it's largely been successful. And what it's allowed us to do is to put on the

public screen many more exhibits and show them to the public than we have not. So that was the idea, and that's what the embodiment of the process is in that order.

As I said yesterday, to the extent we've had closed sessions, I will be the first to admit -- I will admit two things. One is, I was not -- I did not appreciate that I should have been giving some degree of notice to the public about going into a closed session. That is now being corrected.

And two, that the closed sessions we had during the first two weeks, again in hindsight because hindsight is helpful, I think, has helped us as we've gone forward, and it allowed me to get a better sense of what the parties thought needed to be confidential and what I thought was acceptable to be confidential.

And I think the delta between those two things has been reflected in two things: One, the fact that we've only had, I think, at most 30 minutes of a closed session yesterday within the last three-plus weeks; and two, the transcripts of the sealed proceedings that we have now disclosed and our disclosing on a rolling basis, which was something that had not actually been contemplated by the pretrial order that we entered but is something that I've done and, I think, acknowledges that there was this delta. So anybody that has requested access to those transcripts can see that the vast majority of those transcripts are now available to the public to review.

And so we have done a fair amount to ensure transparency in an open courtroom and in an open process, and I've asked the parties to do it, and they really have, I think, in fairness to them worked very hard toward that. Has it been perfect? No. It's a complicated trial. Could I have done better? Yes. It's a complicated trial. But we're all working towards the same goal.

I just wanted to say that at the outset. And we will get to the specifics that have been requested by *The New York Times* later today or after we have an opportunity to hear from Google in a written filing, and then we can consider those. But I did think it was important to share those thoughts with everyone before we got started today.

Mr. Dintzer?

MR. DINTZER: I rise to say, Your Honor, and I try never to speak on behalf of the other side, but I think I can speak on behalf of the other side in saying that the process from our end has worked. I mean, it is imperfect, and there are documents I'm going to need to talk to them about to make sure we're all on the same page before I use them in the court.

But given the complexity, from our point of view, it has worked and allowed us to do effectively what we can in the public sphere as much as possible.

MR. SCHMIDTLEIN: The order I think you're referring to is 647, and it was entered on August the 15th after an

extensive amount of work, really a lot of collaborative work between the parties to try to come to a solution.

Google absolutely supports the notion of keeping as much of this trial in the public as possible. It has required, I can speak on behalf of my client, an extraordinary amount of work. But we understand that that was important work that needed to be done to facilitate this process that Your Honor has put in place. And from our perspective, we've been very grateful for all of the guidance you've given us, and we think it's worked very, very well to date.

And as I said, in 30 years, nothing ever goes perfectly in trials. We always have a few little bumps. But this one actually has worked, I think, extraordinarily well, given the sensitivity of information, given the complexity of all the third-party information that also has to be taken into account.

So I echo and agree with Mr. Dintzer's comments, and we're very grateful for all of your careful consideration of this.

THE COURT: Look, you know, it is entirely appropriate for people to have different views about what we've tried to accomplish and whether it's been successful or not. I appreciate that. I did think it important to frame it in a way that I thought provided a more fulsome explanation, let me just put it that way, of the process that we've put in place and how we've gotten to that process.

No one should be under the impression that we didn't think

about this and that we didn't give it a lot of thought and that there hasn't been a lot of thought and effort in how to do it.

Again, has it been imperfect? Sure. But we've done, I think, overall a pretty good job, and we're continuing to strive to do better.

All right. With that, why don't we turn to -- we'll turn to Google, and we'll hear from your witness.

MR. SMURZYNSKI: Thank you, Your Honor. Ken Smurzynski for Google, and we call Dr. Pandu Nayak.

THE COURT: Just so the record is clear, we're obviously taking Dr. Nayak -- come on up, Dr. Nayak -- out of turn to facilitate his schedule. And so none of this should be considered as evidence that's being presented in the plaintiffs' case. This is Google's case.

P. PANDURANG NAYAK, WITNESS FOR THE DEFENDANT, SWORN

THE COURT: Dr. Nayak, welcome. Thank you for being with us.

MR. SMURZYNSKI: Your Honor, before we start rolling, let me hand out some binders. Your Honor, if I may approach the witness.

DIRECT EXAMINATION

BY MR. SMURZYNSKI:

Q. Good morning, Dr. Nayak. Would you please state and spell your name for the record.

A. It's P. Pandurang Nayak is the name. It's P and then

P-a-n-d-u-r-a-n-g, and Nayak is N-a-y-a-k.

Q. What is your business address?

A. It's 1600 Amphitheatre Parkway, Mountain View, California 94043.

THE COURT: Dr. Nayak, if I can ask you to keep your voice up. That's a mic there, and it will help amplify your voice.

THE WITNESS: Is that better?

THE COURT: Yes. Thank you.

BY MR. SMURZYNSKI:

Q. Would you please tell the Court where you grew up.

A. I grew up in India, primarily in Mumbai.

Q. Where did you go to college?

A. I went to college at the Indian Institute of Technology in Bombay.

Q. Did you receive any awards while there?

A. I did. I received the President of India Gold Medal.

Q. And what does that represent?

A. That's an award given to the graduating student with the highest GPA.

Q. After that, what did you do in terms of education?

A. I came to Stanford University to get a Ph.D. in computer science.

Q. What was your dissertation on?

A. So my dissertation was on artificial intelligence, and I

was interested in the question of how you model systems in a way that it's sort of detailed enough to support the task you want to solve but not so detailed that you get lost in the details.

Q. After graduating from Stanford -- first, what year was that?

A. I graduated in 1992.

Q. After you graduated from Stanford with your Ph.D., what did you do next?

A. I started work at NASA at their Ames Research Center.

Q. Is there any particular area that the Ames Research Center was focused on?

A. So I particularly joined the lab that did artificial intelligence research.

Q. Are there any particular projects that stand out in your mind from your time working at NASA?

A. Yeah. I had perhaps the most exciting project I have ever done that I did at NASA. We built an intelligent spacecraft, a software system called the Remote Agent that provided high-level autonomous control to a spacecraft, and we actually flew it on a real spacecraft for about a week as an experiment. So that was super exciting.

Q. What did you do next after NASA?

A. I joined a start-up that was founded by some friends.

Q. And what was the nature of the work you did at that start-up?

A. So the start-up -- so the technology that we had developed was around text understanding, so things like classification, clustering, this kind of thing. And it started out as a consumer product. So the company was named Purple Yogi at the time, and it was meant to be a personalized news service.

But of course, being a start-up, we had to change things. We changed the name to something less exciting, let's say. Stratify was the new name. And we made it into an enterprise portal. And we had to change once more until we finally got to a successful product, which was a legal discovery system.

Q. There's a lot of call for that, I assume.

Have you ever taught at the university level?

A. Yes, I have.

Q. And where?

A. I taught at Stanford in the Computer Science Department.

Q. What are the classes that you taught in the Computer Science Department at Stanford?

A. So I taught two classes. One was a class called Reasoning Methods in Artificial Intelligence, and this was a class that I developed based on the research area that I was working on. And I taught that for several years.

And then subsequently, I co-taught a class with Professor Chris Manning and with Prabhakar Raghavan on Information Retrieval, which is the science and engineering of search. And I taught that for many years.

Q. You mentioned Prabhakar Raghavan. Who is Prabhakar Raghavan?
A. Prabhakar Raghavan is the SVP of Knowledge and Information at Google. He's currently my boss, but at the time we thought he was not my boss. He was merely a colleague.
Q. And this -- these classes you taught, were they at the graduate school level, undergraduate level, or something else?
A. The classes themselves were graduate classes, but advanced undergraduates, of course, took the class also.
Q. How did you come to join Google?
A. So, I had a lot of friends who were already at Google, and one of them approached me and suggested that I consider joining Google. I spent a fair amount of time talking to them, and after, you know, a fair amount of time, I decided this was the right thing to do. And so I joined Google.
Q. Let's just date that. What year did you join Google?
A. I joined Google in 2004, in November 2004.
Q. What drew you to work at Google?
A. So there were a number of reasons that Google seemed right and has seemed right over the last, whatever, 18, 19 years that I've been there.

The first and, I think, one of the most important reasons was the people. I already knew a lot of the people there, and I had high respect for their capabilities, their creativity, and things. And there were people that were really fun to work

with. And that's actually held through these last 19 years, and that's a really crucial part of what makes Google a wonderful place to be.

The second, just as important, is Google's mission to organize the world's information and make it accessible and useful, and this is a mission that I think really inspires me, and it gives me purpose in my work life at least, and that's really, really important.

And the third thing is, at the end of the day, Google is a technology company, and they really value, say, the skills that I have been trained to and I possess. And I think it's a really great place for people like me to work, and I think I'm incredibly lucky to work in a company that values my skills and gives me the opportunity to have this kind of positive impact on the world through its mission.

Q. When you started at Google, what was your initial role?

A. So when I started at Google, the first project I took on was -- I joined in search, and the first project I took on had sort of this character. Search involves people coming to Google with search queries. And then we try to serve those queries. But it happens sometimes that the query that the user issues has a problem in it. There's something wrong with the query.

The most obvious kind of problem, of course, is like a spell correction. And Google had already developed lots of techniques to help users with spell correction.

But there's another kind of problem that happens, which is that the user has a misconception about what it is that they're looking for.

So the motivating example I remember from the time was this query "NBC survivor." That was the query. Survivor, of course, is a very popular TV show. The only catch is it was actually on CBS, not on NBC. And the user had this misconception that it was on NBC, and that was what the query was.

So we had developed a technique to, in situations like this, to observe maybe the user really meant CBS Survivor and not NBC Survivor. We would issue a second query to get the results for this alternate query and insert those results into the result set with an appropriate user interface around it so that if they indeed meant CBS Survivor, that they would get those results. And not surprisingly, users really loved that, because they really did mean the thing we thought they meant.

Q. Could you sketch for the Court your roles at Google from that 2004 time through today.

A. So after working as an individual computer -- individual contributor or software engineer, I started taking on sort of more management leadership roles. I started with managing the spelling team. So I ran that for a little while. That expanded into running teams that were looking more broadly at query understanding, and that slowly grew until I started being responsible for the whole ranking team. And that's what I've

been doing. We call ourselves the search quality team. And so I've been leading the search quality team now for many years.

Q. What is your title at Google today?

A. So my title today is VP of search. I'm one of the four VPs of search.

Q. What areas do you manage at Google today?

A. So my primary focus is in this area that's called search quality, and that involves all kinds of sort of quality-related things, primarily starting with the ranking of web results, but also ranking of the home page, things like spelling correction, things like featured snippets. There's a whole lot of things, all of which are really sort of quality-oriented things. Those are sort of a part of the team that I lead.

Q. We're going to turn to a demonstrative now. It's DXD17. Let me start with page 002. It should be on the screen in front of you, Dr. Nayak.

What does the outer circle in this demonstrative represent?

A. So the outer circle in this represents essentially the web and the trillions of documents that are on the web. It's the corpus that Google seeks to search.

Q. And what are the implications for search that there are trillions of pages on the web?

A. Well, to start with, it's a lot of pages, and it's a lot of pages that need to be searched and indexed. And so it's a real challenge to figure out how you start by building an index of

that web.

Q. And the next circle in the demonstrative in yellow is information index. How does Google go about creating an index?

A. Yeah, so the index is a core piece of the search, of the search system, and an index is just like an index at the back of a book. There's a word and all the pages that it occurs on. So you need a clear index of the web.

Like I said, trillions of pages is a lot of pages. So it's a little difficult to get an index of the whole web. It's not even clear you want an index of the whole web, because the web has a lot of spam in it. So you want an index of sort of the useful parts of the web that would help users.

And so we go out using a process of crawling the web and other mechanisms for acquiring the content, and we create an index of hundreds of billions of documents that we hope is comprehensive in terms of the kinds of queries and questions that users come to us with.

Q. What is the implication for search of a document being either in the index or not in the index?

A. It has a very significant implication, which is if it's not in the index, we can't serve it to users. It's as simple as that. So making sure that when users come to us with queries, we want to make sure that we've indexed enough of the web so we can serve those queries. And so that's why the index is such a crucial piece of the puzzle.

Q. Does creating an index for purposes of search require any judgment or decisions on the part of Google?

A. Yeah, there's actually a lot that goes on to make sure you have a good index. Like I said, one big piece of it is that there's tremendous amounts of spam out on the web. It's frankly quite shocking. And so being able to make sure that -- if you fill your index with spam, then you're not really going to help your users. And so making sure that you get sort of the right subset of the web, that's one piece of it.

There are other things that we need to do, things like the freshness. So it's not as if you go to the web, crawl it once, and you're done. The web changes all the time. Pages change all the time. Some of them change very quickly. The home page of CNN changes, you know, every 15 minutes as they add new pages. Other pages might change more slowly. A Wikipedia page might change slowly as people make edits to it. And some pages, like say the blog post I wrote back in 2004, may never change.

But pages change at different rates. And if you want to search the web effectively, you need to keep your index up-to-date. So you need to estimate which pages are changing quickly, which pages are changing slowly, and so forth.

Q. Is there a cost associated with creating an index of this size?

A. Oh, there's a very significant cost associated with it, because you need machines to create the index. You need the

network bandwidth to go out and fetch the documents. You need the storage to create the index. So there is a fairly big investment that goes into creating this index, yes.

Q. Now, Dr. Nayak, has there ever come a time when the size of Google's index as measured by the number of documents in it ever decreased?

A. Yes, there have been times when the size of the index has decreased in terms of the number of documents.

Q. And why is that?

A. So we start out with some set of resources in terms of storage, let's say, and that hosts some number of documents. But over time at various times, the average size of documents has gone up for whatever reason. Webmasters have been creating larger and larger documents in various ways. And so for the same size of storage, you can index fewer documents, because each document has now become larger.

In addition to that, our understanding of documents has also improved over time. So when we get these documents, not only do we create an index, we create a bunch of metadata associated with the document which reflects our understanding of the document. And that has also grown over time. And so that also takes space in the index. And as a result, that results in the number of documents that you can index in a fixed size of storage to go down.

Q. Have those changes you've just described in the index

compromised the quality of Google's search?
A. No. At an aggregate level, it does not compromise the quality of search.
Q. Does Google from time to time compare its index to that of Microsoft?
A. Yes, we do that every so often, yes.
Q. Dr. Nayak, you have a binder in front of you that has a white cover. And I would like to direct your attention to UPX268A, which is an evidence. And the A just simply reflects that it's a native version.
And, Your Honor, there are pieces of this that are confidential. So we will be following along.
First, what is this document in front of you?
A. It looks like a document that talks about various competitive facts.
Q. Okay. And if I could ask you to please turn to the document that ends in .033.
A. Yes.
Q. What is being shown -- and I ask you to be careful not to identify any particular numbers. But what is being shown on this page?
A. So this is an attempt to measure the relative coverage of the Google index against the Bing index.
Q. And what is shown by the blue line on that page in front of you, .033?

A. So the strategy that was adopted to do this measurement was as follows: We started with a large sample of queries that were representative of the query stream. And for those queries, we looked at what results Bing returned for those queries. And then we looked to see what fraction of those results were in our index. All right? So that's the blue line.

And you can see over time most of the results that were in the Bing results were in the Google index. You can ignore that little sharp line that goes down. That's just a data error at that point, and that can be ignored. But you can see that the blue line suggests that most of the results that were in the Bing results were in the Google index.

The red line is the opposite of that. So we looked at all the results that were showing up for Google, and we asked ourselves which of these results were in the Bing index. And as you can see, there's sort of a meaningful gap in the index coverage there.

Q. Does that gap in coverage have any implications for the differences in quality between Google and Bing?

A. I think in general, it goes back to the point that if it's not in the index, you can't serve it. And so if you have useful results that are not in the index, as is the case here, then that has a direct impact on quality.

Q. Let's turn to the next page, 034, which is also redacted. What are we seeing here?

A. Here, we have two charts that are created in the same way as that -- the top-level chart that we talked about in the past. The top chart here is on a subset of queries that are identified as long-tail queries. So these are queries that occur quite infrequently in the query stream. And you can see that the gap between the Google line and the Bing line is actually larger here. The bottom line is on popular queries. These are more common queries that occur more frequently, and you can see that the gap is smaller here.

So this suggests that the index coverage is poorer, or the index coverage gap is larger in long-tail queries.

Q. Dr. Nayak, what are the implications for the differences in search quality between Bing and Google on popular and long-tail queries with respect to this chart?

A. So what the charts would suggest is that the quality gap between Google and Bing is larger on long-tail queries than on popular queries.

Q. What, if anything, does that have to do with the index?

A. The thing with long-tail queries is that this is where you really need to have a more comprehensive index, because the user is asking for something very specific. It's not a very common document. And if you don't have that document in the index, you can't serve that long-tail intent that the user came to you with.

So it's the index -- comprehensiveness of the index is

crucial to being able to serve long-tail queries like this.

Q. I would like to turn now to the topics of mobile and desktop.

THE COURT: Sorry. If I could just follow up with a question.

Dr. Nayak, to what -- let me back up.

Can you just share with me what -- in your estimation, what are the most important variables in creating a quality index?

THE WITNESS: I think the first thing you need to create a quality index is to know which parts of the web to crawl. All right? So you have to make sure that you don't fill your index with spam or low-quality content like that. You need to know where the high-quality pages are that people will find useful. So that's a big part of it.

You need to invest enough in the infrastructure for crawling, and that's another important part.

There's another sort of a subtler point here, which is when you crawl the web, you actually use resources of the website, because you're actually, you know, issuing requests to the website to serve this thing. So a really important part of your crawling system is to make sure you don't overwhelm those websites. You need to be very respectful of websites when you do this. So there's a lot of subtle, careful algorithms that are built to allow you to do this in a way that allows you to keep things fresh, while not overwhelming these websites in this

way.

And so then I touched upon the freshness issue. You need to know how to recrawl them to make sure that you do at all times have a reasonably fresh copy of the web that you are looking at.

THE COURT: And does user interaction data play any role in determining, for example -- well, in terms of index quality? Let me just ask generally.

THE WITNESS: I mean, I think certainly pages that users have found useful in the past you want to make sure continue to be in the index. So certainly, there is a role for user interactions in that way.

But there is clearly more to it than that, because there's a ton of pages on the web that have no user interactions, but you still want to make sure they're in the index, particularly for long-tail queries, which by their definition are unlikely to have many or any user interactions.

So yes, you will certainly use user interactions to make sure some set of the web that has interactions is in your index, but that's by no means the whole story.

THE COURT: Thank you.

BY MR. SMURZYNSKI:

Q. Dr. Nayak, focusing for a moment on mobile and desktop, what, in your view, is the primary difference between how Google serves results on mobile versus desktop?

A. I think the most salient difference between mobile and desktop is in the user experience. Right? And there's a very good reason and obvious reason for that. The mobile device has very limited real estate. There's just sort of one column of information. Whereas, the desktop device, of course, has a lot of real estate to provide your search experience. There's multiple columns. There's more information vertically that you can see. It's just a very different experience.

So to me, that is sort of the primary difference between mobile and desktop, is just the form factor.

Another sort of difference, which hopefully is becoming less true today but is still, I think, there, is that mobile devices tend to be on mobile networks, which are just not as fast or reliable as WiFi. That is, desktop devices tend to be on WiFi. And this difference is also something that you need to take into account when you build your mobile experiences.

Q. What does the term "search feature" mean within Google?

A. Search feature for us, so the core of search, of course, is the web results. But then we add to the web results various features. They might be like the "did you mean" feature of spell correction that we talked about. Or it might be a knowledge panel about a particular entity of some sort, President of the United States or something like that, or it might be a certain experience for sports, let's say, to highlight what's happened to your sports team or a sports league

that you're interested in, or weather experience.

So these are all features that we add to make search be more helpful to users.

Q. And how do search features relate to differences between mobile and desktop, if they do?

A. So again, the first implication is that of real estate. So to give you sort of a concrete example, when we first put out knowledge panels, it was done on the desktop, and the knowledge panel was in the right-hand column, and the main web results were down the center column.

But of course, when we brought it to mobile, there was no right-hand column to put the knowledge panel in. And so the knowledge panel had to be bordered onto the center panel, and then there was a question of how you rank the knowledge panel against all of the other results. So a number of technical challenges had to be solved in that regard.

So this is something that is something you have to look at carefully, as you add these features in, where on the page do you add them, because there's only one vertical column of information to be shown on a mobile experience.

Q. As it relates to the area that you have focused most on at Google, ranking, what differences, if any, do you see between mobile and desktop?

A. So I think there's two things to highlight here, one of which I think does affect ranking and one that affects it less.

MR. DINTZER: Your Honor, I hate to interrupt, but if the witness is testifying -- the question was sort of broader than Google. If the witness is testifying about Google, we don't have any problem with him, of course, testifying about Google. If he's testifying broader than that, then we would need a foundation laid, and we might have some questions about that.

So if we could ask counsel to focus his questions about what Google is doing now, that would clean this up.

THE COURT: Okay. I understand him to be drawing on his experience at Google in talking about what Google was doing.

MR. DINTZER: As long as that's understood.

THE COURT: Am I wrong about that?

MR. SMURZYNSKI: Your Honor, he's obviously talking about his experience at Google. He's been in the industry for 19 years.

THE COURT: Right.

MR. SMURZYNSKI: I think the objection, if it is, that these observations can't have any import beyond Google is -- we can all argue about that later, but he's obviously testifying based on his experience and 19 years at Google and the like.

MR. DINTZER: Your Honor, this is not a minor distinction. If Google is offering him to testify about things that are beyond the four walls of Google, then we would ask the Court to hear an objection of ours.

THE COURT: Well, I guess I'm not clear on the distinction in the sense that it's not clear to me that, for example, when he's -- a lot of what he's been talking about is at a very high level. It's not clear to me that that's different from one company to the next.

So if, Counsel, you have a sense that something is specific to Google, if you could embed that in the question so that that's clear. And it may not always be.

MR. SMURZYNSKI: I understand, Your Honor. I think -- I'm obviously asking him for his actual experience as somebody at Google. And we can argue later about what the implications of that are for anything else, but what I'm eliciting is based on his experience at Google.

MR. DINTZER: And Your Honor, I appreciate what counsel has said. I just want to be very clear about our potential objection, because they're fuzzing the line a little bit.

If the witness is talking from personal experience about what he's doing at Google, we're good. If he is extrapolating from what he has done at Google, talking about what other search engines might be doing inside their four walls, thinking, doing, then that enters the realm of expert testimony, for which he is not tendered, not qualified, and we have not had a chance -- obviously, it opens a bunch of doors that we haven't wanted to raise, and as long as he talks about Google, then we don't need

to raise.

But this is not something that we can kick down the road and argue about down the road. If Google is tendering him to talk about something other than what happens in the walls of Google or what's on Google's page, obviously, if he sees something on Bing's page, we didn't object when they run tests against Bing. We don't have a problem with that, because that's happening within Google. But if he's making broader sweeping statements about the industry, about what people do, then that's entering expert testimony that he hasn't been qualified for, and we have concerns.

THE COURT: Okay. Look, from what I've heard, I've understood him to be limiting his testimony to his experience at Google and speaking to that.

Dr. Nayak, if you are going beyond that, you will let us know.

But my understanding is so far he's stayed within the walls you've suggested.

MR. DINTZER: We appreciate that, Your Honor. Thank you.

THE COURT: Why don't we revert to that last question, which I think had to do with ranking between mobile and desktop.

BY MR. SMURZYNSKI:

Q. So I think the question was, and you may have started to answer, I don't know if you had completed, as it relates to your

area within Google, ranking, what differences, if any, do you see between mobile and desktop?

A. So we observed two types of differences between mobile and desktop. But before I get to the differences, we also observed that there was a lot of similarity also.

So on the first point, on the differences, one difference is that the distribution of queries on mobile and desktop, we found differences. There tended to be more location-specific queries on mobile. Those same queries did occur on desktop. So it wasn't like they didn't occur at all on desktop. They just occurred less frequently. So the distribution was sort of skewed towards a little bit more local queries.

Similarly, on desktop, you found more queries which were more research-oriented, because, you know, maybe people like to do research with -- more realistic like this. That doesn't mean those queries didn't occur on mobile. It's just that the distribution was such.

So that was, I think, one difference. But what was interesting was, a very large fraction of the queries were really the same on both sides. So it wasn't like the difference was completely different.

The second interesting difference was even for queries that occurred on both mobile and desktop, we noticed that in some cases, again, by no means all, by no means even a large fraction, but in some cases, the intents on mobile were slightly

different.

A good example of that is, let's say you issued the query "Bank of America" for your bank. Then on desktop, chances are you want to go to the online home page of Bank of America to do online banking. Whereas, on mobile, chances are you were looking for the locations of the ATMs for the bank, Bank of America.

So the intent was slightly different. It doesn't mean that the intents -- that on mobile you couldn't be going to the home page, because you might be wanting to do online banking there also, and it doesn't mean that on desktop you didn't want the locations; you would want that. But it felt like there was a bias on the mobile side to have this more location-focused intents. And so there was somewhat of a difference in that. Again, not for all queries, but there were queries where there was this distinction.

Q. How does Google Search handle a situation like that where there are potential multiple intents for a query, whether it's on desktop or mobile?

A. So one of the signals that does go into Google Search is, you know, is it a desktop query or is it a mobile query. In most cases, that signal doesn't make a difference, but in some cases, you might recognize the difference in intent. And for like that Bank of America query, the result might be that on mobile you might promote the local block that shows the map

where the ATMs are and make the home page of Bank of America be the second result.

Whereas, on desktop, those two would be switched. The home page would be the first, because maybe that's the primary intent, and the map with the locations near you would be the second result.

Q. Are the interactions -- the user interactions on desktop have any relevance to how Google presents results on mobile?

A. So I think I gave you some examples where queries are different. But in a large fraction of queries, they're really the same. The intents are the same. And so the user interactions on desktop are really no different from the user interactions on mobile.

In that subset of cases where there's a difference, you will see the difference in user interactions.

Q. I'm going to move on from mobile to another topic, Dr. Nayak.

THE COURT: Let me just ask you, can you estimate just roughly what you would consider to be in that subset, that is, the number -- the percentage of queries where the intent may be slightly different depending upon the device you're using?

THE WITNESS: So I don't know what it is today, honestly. I remember a study that was done some time ago, and it seemed like there were about of the order of 6 percent of queries which were sort of more mobile-focused and 6 percent

that were more desktop-focused, and the rest, whatever, 86, 88 percent, whatever, were roughly the same.

So that's the kind of qualitative feel I have that says a lot of queries are really, really the same, and then there are some that are more sort of mobile-focused and some that are more desktop-focused.

THE COURT: Thank you.

BY MR. SMURZYNSKI:

Q. Dr. Nayak, you have in your binder a document that's identified as DX227. And this again is a document that there's some partial redactions on.

But to start, what is this document?

A. So this is a presentation that we gave at the search quality all-hands meeting in December of 2020.

Q. Okay. And if you could turn to the page that ends in .011.

A. Yeah.

Q. And without identifying any of the particulars, what is set forth in this page?

A. So this -- remember, this is December 2020. So this is a slide that was talking about the work that we had planned to do in 2021, in the new year. So this was a strategy document, or a summary of the strategy anyway, of the primary pieces of work that we were planning to do in 2021.

MR. SMURZYNSKI: I would just note this document is in evidence. Thank you.

BY MR. SMURZYNSKI:

Q. If you will turn to the second -- or the next page, at .012.

Again, what is being set forth, without getting into the exact detail, with respect to these points?

A. Yeah. So on the previous slide, where we had talked about the different -- the strategy for 2021, we had broken up the work into different buckets of work. One of the most important buckets of work was that first bucket, which is highlighted on this slide here, which is to maintain industry-leading search quality. This is an important objective, important goal for us every year, is to do this investment.

And what this slide is describing is the specific metrics we were using and the specific key results, so the specific goals we were trying to achieve to allow us to continue to maintain industry-leading search quality.

And there were two metrics that we were measuring ourselves with. One is our sort of top-level information satisfaction metric. That's IS. It's a measure of overall quality of results. And the other is the PQ, the page quality metric. It's a measure of the reliability of search results.

And ever since the beginning of Google, when they introduced page rank as a measure of reliability, our goal has always been to surface results that are relevant and, whenever possible, from reliable sources. So those are sort of the two

key elements of great search results, and information satisfaction and page quality are the two metrics we use to track that.

And so we set ourselves goals every year for improving IS and PQ, and that's what this was capturing.

Q. It's hard to read, but within the IS portion under "key goals," there's a reference to "NBU goal."

What does NBU mean?

A. NBU stands for next billion users, and it represents various emerging markets where we operate, things like India, Indonesia, Brazil, Nigeria. We seek to serve users all over the world. So it's important for us to have search quality improvements everywhere. And so one of the key things we had called out there was to set ourselves a goal of improving search in these NBU markets.

Q. Now, this document we just reviewed was for 2021. Are there similar documents and goals for other years at Google?

A. Yeah. This is a part of our annual planning. Every year, we set ourselves goals that -- OKRs, you may have heard of those, objectives and key results. And we set up objectives along these lines.

In search quality, we always have an objective like this one here to improve search quality according to the metrics that we measure here. And we do this every year, yes.

Q. Dr. Nayak, does Google only seek to make improvements in

its quality when it senses some competition?
A. No, not at all. This is something we do all the time. This is -- this goes to the core of serving the mission that we have, is to constantly improve search quality.
Q. Does Google conduct any measurements of its quality in the ordinary course of its search engine?
A. Yes, we do.
Q. And why does it do that?
A. Well, there's a saying which is sometimes attributed to Lord Kelvin, maybe misattributed. It says you can't improve what you don't measure. And even if the attribution isn't accurate, it makes an important point. If we are seeking to improve search quality, then we better have a way of measuring whether it actually is making improvements or not.
Q. What metrics does Google use to measure its search quality?
A. So we use a number of different things of course, but there is one measure that is sort of, I think, the most important, and that is the information satisfaction measure that we touched upon briefly.
Q. Are there any subcomponents of that IS score that you've mentioned?
A. So the IS score is meant to capture what makes great search results. And really, there are, as I mentioned, really two elements to that: Are other search results relevant, and are they from reliable sources or not?

And so those are sort of the two major pieces of the IS metric. And we get raters to help us make these judgments that then get put together into this metric.

Q. On what scale is that aggregate metric portrayed on?

A. The aggregate metric is on a 0 to 100 scale.

Q. And can you please give the Court some sense of what one IS point or some portion of an IS point means in terms of quality.

A. So not surprisingly, it's a little tricky to give a very clear definition of what one IS point would mean. So we came up with the following way of thinking about it: Wikipedia is a really important source on the web, lots of great information. People like it a lot. If we took Wikipedia out of our index, completely out of our index, then that would lead to an IS loss of roughly about a half point.

So that gives you a sense for what a point of IS is. A half point is a pretty significant difference if it represents the whole Wikipedia wealth of information there.

So that's how we've been thinking about it.

Q. And how does Google go about scoring -- or generating the data to create these IS scores?

A. So we have a whole system of evaluation to produce these IS scores. The core of it is we have a whole lot of raters. These are people that are hired from all over the world. We have about 16,000 or more of them around the world. And we essentially ask them to take a look at a sample of queries and

the results for them and to provide judgments, ratings for how well the results match the queries.

Now, to do this, to help them do this, we have actually developed a very detailed search quality rater guideline document. It's a 160-page document. It's a public document. You can actually search for it on the web. And if you don't find it, you can complain to me. But it is a public document that anyone can look at.

And raters are expected to read this document, understand it, internalize it, and take a test on it. And having done that, they are then expected to interpret what this document is saying in terms of the specific queries and results that we show.

And the document itself, as I mentioned, what search wants to be is to produce relevant results from reliable sources whenever possible. And so this document goes into a lot of detail on what does it mean to be relevant, what does it mean to be reliable. Right? And so that gives raters guidance on making those judgments.

So they give those judgments on these queries, query result pairs. We aggregate those judgments up to the query level and then aggregate it up to the query set level. And so we get an overall metric for how we're doing for that particular sample of queries that represents our query stream.

Q. You may have already mentioned this, and I apologize if you

did. Approximately how long in terms of pages is this rater guideline?

A. It's 160 pages.

Q. And you mentioned it's public. Why does Google publish its rater guidelines?

A. I think a big part of it is transparency. In a sense, the search rater guidelines is our product spec. It's what search wants to be. It's what we try to accomplish. And we want to be transparent about what it is that search is trying to do here, and the search rater guidelines is a key part of that transparency. It also gives webmasters a lot of guidance on what we think is pages that will serve our users and what are the characteristics of those.

So those are sort of the reasons why we've made this document public.

Q. Allen, if you could bring up the demonstrative in slide 3.

Dr. Nayak, what is shown on the demonstrative in front of you, slide 003?

A. So this -- these charts show the different kinds of experiments that we do as a part of search quality. And you can see we do -- we certainly do live experiments, which are shown in red there. But a big part of the experimentation we do is around the human rater testing evals that we do.

Q. And why does Google choose to use all these human rater tests when it already has available live traffic?

A. I think the reason for that is live traffic is certainly helpful. That's why we do live experiments like this. Right? But they can also be extremely misleading, particularly in issues of sort of page quality-type things.

So I think everyone is probably familiar with click bait, for example. These are -- tend to be low-quality results, but that when you look at the headline for it, you know, you really do want to click on it to find out what's going on.

And so users will do that, and if we just follow what the live experiment says, we're likely to promote click bait. And so while live experiments are useful, they're not a great sort of top-line metric for what we do.

Instead, we want to have the human raters, human rater program with the search quality rater guidelines so that we can have a very clear sense for this is what we think great search is.

Q. And you mentioned page quality. What is the relationship between page quality on the one hand and clicks?

A. I mean, the -- when looked sort of at an aggregate level, what we have noticed is that page quality is a little anticorrelated with clicks.

Q. When you say "anticorrelated," what do you mean by that?

A. It means that in cases where we improve page quality on the margins, not sort of at the -- at the big level, I think page quality is a good thing in the long term. But on the margins,

whenever we improve page quality, what we've noticed is that our live experiments come out not so great. So they move in sort of opposite directions.

And I think part of it is things that I mentioned around things like click bait and variations thereof. If we demote the click bait, users may still seek it out, and that looks like a loss on the live experiment when really it's actually a good thing to do that.

Q. All right. We can take that demonstrative down, and let's turn to a document in your binder, UPX1082.

Your Honor, this was on the government's exhibit list. I believe we at one point posed an objection, but we're obviously withdrawing that objection at this point.

THE COURT: Okay.

BY MR. SMURZYNSKI:

Q. Dr. Nayak, what is UPX1082?

A. This is -- I believe this is a deck that describes some experiments that we did to connect improvements in search quality to growth in search usage.

Q. And why did Google perform a study such as this?

A. So for the longest time, we have believed that improving metrics like IS -- in the past, we have had other variations of IS. They were called different things. But it was the same basic idea.

We have long believed almost as an article of faith that

improving metrics like that leads to a more helpful search for users. And some years ago, actually a little before when this deck was done, we decided well, let's make sure it's not just an article of faith, it really is the case that we are building a more helpful Google.

And the way we wanted to say that we're being more helpful is to see if users would use search more. We used growth essentially or usage of search as a proxy for helpfulness, the idea being that if users find it helpful, they're going to use it more. They're going to use it more frequently. They're going to use it for more tasks that they're looking for, but in general just using it more.

And so we were trying to sort of connect these two things. And so we wanted to see whether making improvements in IS was actually correlated with greater usage of search, so that it wasn't just an article of faith but we could say we've got data that shows this. And that's what this deck was describing.

Q. All right. And this deck describes a degradation of one IS point. We've redacted the actual effect from that, but what conclusions did you draw from the change in user behavior with a one IS point degradation?

A. I mean, the conclusion I draw is I find the study to be remarkable in a sense, that we make these changes in search quality which nobody really notices obviously, you know. You're changing the web results around. No one knows you launched

anything. No one knows you made an improvement. And yet, users subliminally notice that search got better. And they feel like they need to use search more. I think it's just a remarkable thing that this shows.

And to me, that was the big takeaway, that it matters what we do in improving search quality. People notice it, even if subliminally.

Q. If you could turn to the fifth page of that document, under "caveats." And in the bottom paragraph there, there's a statement about the relationship between IS and growth and conclusions one can draw for different or not draw for different IS points.

What is being conveyed there?

A. So the document itself or the studies that they did seemed to imply that there was a linear relationship between IS and growth. Okay? So that was -- these were sort of good data scientists who did it, and they wanted to make sure that we understood the limitations of the study.

And what they were pointing out is they did these experiments in a certain range of IS changes where there was this linear relationship that they observed, and they wanted to point out that if you made much larger IS changes, the relationship might not stay linear. It might become nonlinear. There might be inflection points where if you make search much worse, for example, you might actually lose a lot more traffic

than one might imagine with this.

So this was just a caveat, that it might not be linear outside because we had not investigated that.

MR. SMURZYNSKI: Your Honor, I'm about to change topics. If this is a convenient time for our morning break, it would be useful in the course of this exam.

THE COURT: Sure. All right. It's a little bit before 11:00. We will resume at 11:15.

Dr. Nayak, I will ask you not to discuss your testimony during the break. Thank you, sir.

THE WITNESS: Thank you.

(Recess taken from 10:56 a.m. to 11:15 a.m.)

(Call to order of the court.)

THE COURT: Please be seated. Thank you, everyone. Counsel, whenever you're ready.

MR. SMURZYNSKI: Thank you, Your Honor.

If we could put up the demonstrative again.

BY MR. SMURZYNSKI:

Q. Dr. Nayak, is the quality of Google's search engine determined by the volume of user interaction data it has?

A. There's actually a lot that goes into the quality of search results at Google. User interaction is one piece of it, but by no means the only piece. There's many, many other important pieces in there.

Q. Using this demonstrative, could you start to explain that?

I realize it's a lengthy and complicated topic.

A. So as we discussed earlier, we have hundreds of billions of documents in the index. And certainly, being in the index is a crucial part of search quality, because if it's not in the index you can't serve it. But even once it's in the index, there's a lot of work that needs to happen.

And given a query, the first step you have to do is to retrieve documents that match the query. A typical query might have millions of documents on the web that match it, but there's no way that in the fraction of a second that we need to do all this in we can look at a million or millions of documents and retrieve them.

So instead, what we do is we have a retrieval process that gets us of the order of tens of thousands of documents from the index that you can actually look at. And here, too, you have to do a good job of deciding which tens of thousands of documents to get, because if you don't get the relevant or the important documents in there, again, you've lost it. It's like it's not in the index. So the retrieval step is a crucial one there.

And then given the retrieval step, given that you've retrieved tens of thousands of documents, you then need to go in and start deciding which ones you're going to really spend a lot of time scoring so that you get down to like several hundred documents. You do some lightweight scoring at the beginning to get down to several hundred documents that you're actually going

to do a detailed ranking with your best scoring functions. And then you bring it all the way down to, say, the ten documents that you're actually showing.

So there's a culling process starting even with crawling from the trillions of documents to the hundreds of billions, from the hundreds of billions to the tens of thousands for retrieval, and then to the hundreds and then down to the ten documents.

There's a culling process that goes on, and in each step, you have to be careful to make sure that the relevant documents are returned.

Q. And does all that have to happen before you can even have a click?

A. Yeah. I mean, I think it's important to realize that the only documents that get clicks are the ones that we surface to the user. So you have to work hard to make sure you surface good documents to the user before they can even consider clicking on them.

Q. At a high level, how has Google gone about developing its systems to surface documents in the first instance?

A. I mean, we use a variety of signals. We have several hundred signals that we use that work together to give us the experience that is search today.

The signals vary on a number of different dimensions. It starts with the most basic and in some ways the most important

signal, which is just the words on the page. The words on the page are actually kind of crucial, and that's where the index comes in. Where the words occur, is it in the title or is it in some metadata or is it in the body, these kind of signals are very important.

Another very important signal is the links between pages. That proves to be another extremely valuable signal. When Google first started, they introduced this notion of page rank, which is really a business of processing these links between pages, and it continues to be an important signal today.

There's a variety of other signals that we --

THE COURT: I'm sorry. That topic has come up a couple of times. Can you explain what you mean by links between pages? Do you mean hyperlinks between websites? Is that what you're referring to?

THE WITNESS: Yeah, it's exactly the hyperlinks between pages on the web. And the important thing here is, each hyperlink has some text associated with it, the linked text associated with it. The linked text is actually about what the target page is about, not what the source page is about.

So the linked text is a very valuable clue in deciding what the target page is relevant to. So that's what makes it such a powerful signal.

THE COURT: And I'm sorry. I interrupted you. There was more to your answer.

THE WITNESS: Then there are other signals that we use. We talked a little about freshness as a signal. Freshness is important even as a notion of relevance. For example, if you wanted to find out something about your favorite sports team, you want the pages that were published maybe this morning or yesterday, not the ones that were published a year ago, even though they might be relevant in that sense, but they're not really relevant because they're not the information you're seeking.

Similarly, if you're looking for a new laptop, maybe you don't want the page that was published today, but you want laptop reviews from 2023, because those are the laptops you will be looking at, not the laptop reviews in 2022.

On the other hand, if you're planning your Thanksgiving meal and you want a turkey recipe, then maybe the recipe from ten years ago is actually better than the recipes from today.

So the notion of freshness and deciding whether to use it or not is a crucial element here.

Similarly, another important signal is location. We touched upon it a little with mobile. But location is an important one. You can search for pizza. You want the pizza restaurants near you, because you actually want to eat the pizza -- not maybe the Gino's East or Renaldi's Pizza in Chicago, which everyone tells me is the worst crust greatest pizza. Right? So that's not as relevant to you. And the thing

that gives you that signal is location.

And then we have a whole lot of signals that we broadly think of as language understanding. And language understanding is sort of a crucial element here, because you need to understand the language of queries, the language of documents, and the match between them.

Q. You made reference in your answer to freshness. How do the existence of a collection of clicks in the logs interact with this concept of freshness?

A. The challenge with freshness and clicks is that clicks accrete over time, which means older pages, potentially stale pages, tend to have more clicks than fresh pages which may start out with no clicks at all, but even if they start acquiring clicks still will have fewer clicks than sort of the pages that have been around for a while.

And so if you want to have a good fresh set of results, you really have to take into account the fact that clicks tend to create staleness, and you need to compensate for that in some way.

Q. And when you were talking about those links between pages, is there another term that's used at Google to describe that sort of text in the links?

A. We call it link text is the text on the link. Anchors is another phrase that we use. Anchors is also links, yes.

Q. And all of these things you've described, some of which use

click data and some of them which don't, does Google just mine the clicks and create a table and serve results, or is there other stuff that's gone on over the last 20 years at Google while you've been involved with search ranking?

A. No, I think -- I mean, certainly, we use clicks and clicks are important. There's no question about that. But you have to do a lot more. And one whole area that I didn't mention and even the previous answer was this notion of page quality, which is largely nothing to do with clicks and, in fact, as we noted, can be anticorrelated with clicks.

Page quality signals are tremendously important, because we want to have the sort of authoritative, reliable information being surfaced in Google.

And so there's just a lot of work that goes on in all these different areas. Language understanding today is, perhaps, the most exciting area of work going on.

Q. We'll get to that in some more detail in a little bit.

We talked earlier about long-tail queries. What's the role, if any, of user interaction data in responding to long-tail queries?

A. So long-tail queries, by their very nature, occur infrequently. That's almost by definition is what happens. And as a result, they have few, if any, clicks for them. And even the clicks that they do have, again by the nature of being clicks, can be noisy. And when there's few of them, the noise

becomes meaningful. When there's a lot of them, then you can say that the noise can be modulated.

And so for ranking of long-tail queries, language understanding becomes sort of the crucial element in ranking long-tail queries.

Q. Google has a large collection of sessions logs. Does each click, each piece of data have the same value to Google?

A. So there's a broader notion of the law of diminishing returns, which very much applies for us in our use of logs. And the idea is very simple: When you start out and start getting some data, there's a lot of value to it. Right? And so there's significant increase in quality as you get more data.

But after a while, the value you get from every additional piece of data starts falling, starts diminishing, and it starts sort of flattening out. So you get this sort of law of diminishing returns.

Now, of course, it continues to increase a little bit. So more is better than less at that level. But even here, the story is more nuanced. Because as you get more data, it's more expensive to process. So the cost of processing the data goes up if we're talking about large amounts of data. So the cost of processing it goes up. The time to process it goes up. It has implications on the number of experiments you can run, because each time you have to make a change you have to wait a while for the model to be built and so forth.

And so there is this trade-off that we have in terms of amount of data that you use, the diminishing returns of the data, and the cost of processing the data. And so usually, there's a sweet spot along the way where the value has started diminishing, the costs have gone up, and that's where you would stop.

Q. Dr. Nayak, in your binder, there is a document DX108. And the DOJ or maybe both plaintiffs have a relevance objection to this, but I don't know if they're standing on it.

MR. DINTZER: No objection, Your Honor.

MR. SMURZYNSKI: Your Honor, we ask that DX108 be admitted.

THE COURT: It is admitted.

(Exhibit DX108 received into evidence.)

BY MR. SMURZYNSKI:

Q. Dr. Nayak, this is a confidential document. So it will be in the binder but not on the screen.

What is DX108?

A. This is discussing a launch report that I guess was in 2017 where we were looking at decreasing the amount of data used for one of our systems by about a third.

Q. Okay. And we can identify the name of the system. What was the system?

A. This was for the Navboost system, which used session logs, click inquiry data on session logs, and it was looking at

decreasing the use of that data.
Q. And what did Google find when it made the determination to reduce the Navboost data in the respect that's shown in DX108?
A. I mean, what we found was that there was no meaningful change in search quality with this decrease, which was really great, because we could then process the data more quickly because there was a third less data to process.
THE COURT: Did Google implement that change?
THE WITNESS: Oh, yes. This was approved, and we moved ahead with this, yes.
BY MR. SMURZYNSKI:
Q. As head of search quality, do you have any involvement with Google's machine learning systems?
A. Yes, I do.
Q. Earlier this morning, you mentioned that you received your Ph.D. in artificial intelligence in 1992. At a high level, what is artificial intelligence?
A. Artificial intelligence is the science and engineering of getting machines, typically computer programs, to exhibit intelligent behavior. It's a bit of a circular definition, but that's as close I think as you're going to get.
Q. And without tracing the entire history of it, has artificial intelligence work changed since 1992?
MR. DINTZER: Objection, Your Honor. Again, if he's talking about Google and how it's changed Google, we have no

problem. This question is about as broad as you can ask for, and that floats into expert territory, which again we would have some problems with. So if he wants to ask about artificial intelligence work at Google, then we have no problem at all.

THE COURT: Well, I think this is a predicate to getting there.

MR. SMURZYNSKI: I'm not going to focus on exactly how AI worked in 1992, Your Honor.

THE COURT: I understand this to be just a general background question that he's more than qualified to answer, and we will move from there.

MR. DINTZER: Thank you, Your Honor.

THE COURT: So the objection is overruled.

BY MR. SMURZYNSKI:

Q. Do you recall the question?

A. Yeah, can you just repeat it?

Q. Certainly. At a high level, how has AI changed from the time you received your Ph.D. in 1992 to today?

A. So when I did my Ph.D. and my own work, for that matter, a lot of the work was focused on directly developing algorithms to exhibit the kind of intelligent behaviors that we were interested in: Planning, diagnostic reasoning, this kind of thing. And a lot of the work had that character to it.

There was work in machine learning even at the time, and machine learning is a very different way of approaching the same

problem, which is, instead of trying to directly develop algorithms to exhibit the behavior of interest, you start with a corpus of data that describes the phenomenon that you're interested in modeling. And then you use machine-learning algorithms to induce patterns from that data so that your program can then exhibit the behavior of interest. So there were machine learning systems there.

What has happened over the years is in the 21st century, particularly with the rise of deep learning as a very powerful machine-learning approach, AI has essentially moved to the business of machine learning. So there is very little work as far as I know anywhere. Certainly, all the excitement is around machine learning with deep learning to exhibit the kinds of intelligent behaviors that we're talking about here.

Q. Turning now to your time at Google, in the first decade or so of your time at Google Search, so that's 2004 to 2014, how did Google Search use machine learning?

A. So in those early years at Google, we did not use machine learning very much. There was a deeply philosophical position that we held that effectively said that it's very important that we understand in detail how our systems work. Right? And so we would develop ranking functions by hand, which we understood the properties of. And the reason for wanting this understandability was that when things went wrong, which they reliably did, you wanted to go back and understand what about

your system led to that failure. And by understanding the system, we felt that you could actually do a really good job of fixing it and improving your system.

And this philosophical position held us in very good stead for many years. We ventured into forms of machine learning where we would learn some of the parameters of the functions along the way, but still, the systems were still understandable in this way.

This changed in 2015 when we moved to using machine learning a lot more, but this was sort of the position we had for the first many years.

Q. And Allen, if you could put up in the demonstrative the timeline.

Dr. Nayak, at a high level, what was your involvement in creating this timeline?

A. I worked with you and various others to produce this timeline.

Q. And what is the relationship or the significance of some items being on top of the line and some items being below the line?

A. So the items on the top of the line is work, either systems or papers, done by members of the Google research team. So these are all various research advances that our research teams did.

The items below the line are applications of this research

that our team in search and search quality did to use these advances in search.

Q. Let's start with the 2011 entry for Google Brain.

What is Google Brain?

A. So Google Brain is, I think, the first large-scale machine learning system that was built. It was done in collaboration between Professor Andrew Ng, who was an expert from Stanford on machine learning, on deep learning, and he was spending some time at Google, and he collaborated with Jeff Dean, one of our foremost engineers and an expert on distributed systems.

And so the idea in Google Brain was to see if you could train massive neural networks on a distributed platform with lots of computers, because with these very large networks of neural net, they won't fit on one machine. So you need a distributed setup with many different computers talking to each other.

And the result was Google Brain, and it started doing some pretty amazing things. The scale and the size of these things and the way to train them was all quite remarkable and has really set off the explosion of interest that we're seeing in machine learning, large-scale machine learning today.

Q. And you mentioned distributed computing. How does that relate in any way to the work that Google Search had done up to that date?

A. So distributed computing is absolutely central to

everything Google did from the beginning, for the simple reason that the only way you get scale, that is, you can operate with large amounts of queries coming in, large indices that you have to create, et cetera, none of these things can be done on a single machine. You need to have a collection of machines working together and all the challenges that that poses in terms of the reliability of those machines and so forth.

So distributed computing is one of the core technologies that people like Jeff Dean and others developed over the years to make Google what it is.

Q. All right. I see the next entry on the timeline is 2013, Word2vec.

If you could advance the slide deck, Allen, to 06.

What was Word2vec?

A. So when you start -- neural nets essentially operate on numbers, and it's computing some sort of a big function with numbers. So if you want to model text in a neural net, you have to find a way to convert text or words into numbers.

And so the very clever idea that was brought up here was this notion of embedding vectors. And the idea is as follows: You take every word and you map it into a high-dimensional space. A high-dimensional space, 128 dimensions, 256 dimensions, something of that sort, so some high-dimensional space. So every word is a point or a vector in that space.

And you want this embedding, this mapping, to have the

property that if two words sort of mean the same thing or are close in meaning, then they map to points that are close to each other in this space. Right?

Now, we've tried to show that in this demonstrative here. It's in three-dimension, because nobody has yet figured out how to visualize a 128-dimension space. But it illustrates the idea.

As you can see, the word "porpoise" and "dolphin," they're sort of related, very closely related items. So they both get mapped to or they get embedded to points or vectors that are very close to each other. SeaWorld is quite correlated with dolphins, Shamu and so forth. Right? So the embedding or the mapping of SeaWorld is also close, though not nearly as close as porpoise, as close to dolphin. But something like Paris is actually quite far away.

And so this is the kind of property you want from embeddings, is this notion of semantic similarity being proximity in this Cartesian space. Word2vec was a way of automatically creating these mappings by processing data and looking for word occurrences in a big corpus of data, like a Wikipedia corpus, or nowadays, it's done with even larger corpora like the web corpora.

And you look at this co-occurrence of words. And Word2vec was one of the first really powerful ways of creating these embeddings that have this property and lots of other interesting

properties here.

And so in that sense, it was sort of a central element in being able to apply neural networks for text.

THE COURT: Can you just explain what you mean by vectors? I think of a vector as a line, but that's not what you mean.

THE WITNESS: Actually, the way that -- a vector is simply the line from the origin to this point. So when I have these two points, the porpoise and the dolphin point, they're actually two vectors, one from here to here and one from here to here. So the two are interchangeable.

THE COURT: So the point is representative of the proximity or closeness of the words?

THE WITNESS: Yes, exactly. When those two points are close to each other, the implication -- what you want is that should mean that the two words sort of mean the same thing or are close in meaning or are related in some way.

BY MR. SMURZYNSKI:

Q. Allen, if you could advance the slide -- sorry. Go ahead.

A. These are not mappings you can do by hand. So you have to induce them from corpora of data, and that is what Word2vec did, as sort of a concrete example.

Q. Below the line on the timeline in 2015, we have RankBrain. What is RankBrain?

A. So RankBrain is the first application of Google Brain, this

sort of massively -- this massively parallel neural net that they had built, but applying it specifically to the problem of search. So it essentially looked at a query in a document, representations of those in terms of text, created the network and, you know, generated a score that said how relevant is the document to the query.

Q. And I don't know if in 128 dimensions or more there's a notion of closeness, but is that the idea that's going on, that you have a query and you have a document and you're trying to find those that are close in 128-dimensional space?

A. So that's an important point, but that actually is done by one of the later systems.

Q. I've jumped ahead. I'm sorry.

THE COURT: You're ahead of your time.

THE WITNESS: But the -- what this is doing is it's inducing from data this notion of relevance between queries and documents. And the way this proximity comes in is in the generalization step.

So if you have two queries, and suppose for the first query you know this document is relevant, and the second query defers from the first query in only, let's say, one word, which is a synonym word. Then you would expect that the embeddings for those two queries would be sort of close to each other, because they're synonyms in this sense, which means if the first query is determined to be relevant to the -- the document is relevant

to the first query, then an embedding that is sort of close by will also be relevant, the document will be relevant to that query.

So in a sense, you've generalized from one query document pair to a different query document pair in this way because of this notion of proximity. And that's one of the ways that these neural nets generalize. They do it in other ways also, but that's where the real power of these neural nets comes in.

Q. Did the launch of RankBrain have an impact on search quality at Google?

A. Yeah. RankBrain was the single biggest launch for any single launch. It was the biggest search launch that we've had in improving search quality since I can remember, at any rate. Maybe before that there were bigger ones, but certainly, it was a very significant improvement.

Q. All right. And above the line in 2017, there's a reference to the paper "Attention is All You Need."

Who published "Attention is All You Need"?

A. This was a paper again published by researchers in Google Research.

Q. And what was the point being explained in "Attention is All You Need"?

A. So this is a really, really interesting series of ideas. The first thing to note is that language and the meanings of words is incredibly context-dependent. All right? There's this

famous linguist from the 1950s, John Rupert Firth, and he had this beautiful line, which was, "You shall know a word by the company it keeps." And the observation being that words derive their meaning from the context in which they're used. Right? So that is sort of one key point.

Now, in Google Brain and in RankBrain, we took queries or any sort of sequences of words and broke them up into pieces and created a bag of words, individual words as such.

And what that meant was a query or a sentence like "dog bites man" would look very similar to a sentence like "man bites dog," because they have the same set of words in them. So when you think of them as bags of words, those two look the same. But of course, one of them generates a news article and the other doesn't. Right? So they mean very different things. And what makes the difference in meaning is the actual sequence.

So "Attention is All You Need" was an attempt at understanding words in sequence, not as bags of words as was done in Google Brain or in RankBrain, but as in this sequence manner.

And what the paper did was it introduced a component called a transformer, which is a collection of neural net elements put together into a unit like that. And the transformer looked at all the context of words around a given word to try and understand its meaning.

But it did one other very important thing, which is it said

not all the words in the context are important. There's only some words that are important. You need to pay attention to only some words.

And so this transformer element both gave you the context but also gave you a mechanism to focus on only certain parts of the context that were relevant that you paid attention to. And they showed that using this notion of a transformer, you could get a lot of good stuff for understanding words in sequence.

Q. And in this context of machine learning, what is a transformer? What's its nature?

A. So the transformer, as I said, it's a network of neural elements. It's a software component. And you can build larger networks out of it. In fact, all of the excitement around large language models that we have today are built out of transformers. So you might have heard of ChatGPT. The T stands for transformers. So they're all based off of this article here.

Q. Okay. And then the next article in the timeline, 2018, refers to BERT.

What is that article about?

A. So BERT was really quite a landmark publication, again by people at Google, that essentially took the transformers from the 2017 paper, and it created an architecture that -- and they introduced a particular way of training this model that proved to be incredibly helpful in solving just about every language

understanding problem that the research community was looking at.

Q. And if you will advance the slide deck, Allen.

What's depicted here, Dr. Nayak?

A. So this is what BERT looks like. It's made up of a set of layers, up to 24 layers. So it's a very deep network. Each layer consists of a set of transformers. There's -- and what each transformer is doing is it's taking one of those initial word encodings, encodings or embeddings that we talked about. It looks at the context in which that word occurs, and it creates an output encoding and sort of sends it through. So there's one sort of transformer stack for each word.

And I think it's easiest to describe this in the context of an example. Consider the sentence "the animal didn't cross the road because it was tired," and consider the encoding or the embedding of the word "it" and the word "animal." In a context-independent fashion, which is what the Word2vec and what we had before, you would expect that they would get embedded quite far apart. But of course, we know that in this context "it" refers to the animal.

So what BERT was doing is, as it was looking at the context in which it is used, it is creating these contextual word encodings at the bottom. And in the contextual word encodings, the word "it" in this context will map close to the word "animal," because in fact they refer to the same thing.

And in terms of attention, notice that it probably paid attention to the words "was tired," because that's the thing that clues it in to the fact that it refers to animal. If the sentence had been "the animal didn't cross the road because it was too wide," then "it" would refer to the road, and it was paying attention to the "too wide" is what told it that it's the road.

And so that's what the word is doing, is it's developing this contextual understanding of words in the sequence. And it turns out that this contextual understanding becomes incredibly powerful to help you solve all sorts of natural language problems, question answering, and a part of speech tagging and all kinds of further such things.

Q. How is BERT trained, Dr. Nayak?

A. BERT is trained -- in the BERT paper, they trained it on a large corpus of text. I think it may have been Wikipedia text, but it might have been something more than that. But it was trained on text.

Q. In your binder, if you would turn to Exhibit DX134, which is in evidence. We'll start with the cover slide, "zero to BERT in 60 minutes, 2019 Rankapalooza."

What is Rankapalooza at Google?

A. It's an event that we have for our whole search quality team. We try to bring them together. We build community. We have educational events. We take it as a place for people to

work together. We find it to be incredibly helpful in enabling our team to do good things.

Q. If you will turn to page that ends of .030, and it's a slide that has the title "happy things."

That first bullet reads, "BERT emerged from a collaboration between web answers and research, extending a line of internal and external research."

Can you just explain what's being conveyed there?

A. Yeah, so we described what BERT was, this pretty remarkable landmark publication. It didn't just come out of nowhere. The researchers, Jacob Devlin and others who worked on it worked very closely with our web answers team. This is a part of our search quality team. The web answers team is focused on extracting the right passage from a document that is most relevant to your query.

So it's the featured snippets you sometimes see at the top of the search results. So it's like a question answering-type task.

So the research team was working closely with us to see how they could help us improve their answers. And as a part of that collaboration, BERT emerged as sort of the general research idea that helped us improve, among other things, web answers.

Q. And the next bullet point on page 030 reads, "Not a lucky coincidence."

What's being conveyed there?

A. I think what's being conveyed there is that innovations like BERT again don't just happen. It's because -- they happen because Google has set up an environment that enables innovations like this.

So what are some of the elements of that environment? One is Google understood early on that machine learning was going to be extremely useful. And so they made a lot of investments in machine learning hardware. They developed TPUs, Tensor Processing Units, for example. They invested in enabling researchers and product teams like ours to use these TPUs in various ways. And so that was a key element that allowed us to develop BERT at all.

The other thing is relationships. That is, these problems that our researchers were working on, they were not working on it sort of in an abstract ivory tower fashion. They were working closely with us on the product teams to say how can we advance the research to enable more effective products, and that's what's led to these very impactful things here.

Q. Allen, if you could go back to the timeline, please.

How has BERT been applied by Search into the search product?

A. So BERT is applied in a number of ways, but I think one of the most interesting ways was for our core ranking work. We launched a system called DeepRank, which is -- essentially uses the technology of BERT to significantly improve language

understanding that we need for ranking.

And DeepRank essentially was the largest single change we made, improvement we made to search since RankBrain. So big advance again.

Q. All right. Allen, if you could advance the demonstrative.

Briefly, Dr. Nayak, what is shown on slide 010?

A. So this is an example from the DeepRank launch report, one of the queries in the valuation that we had there, and it illustrates some of the more subtle nuances that DeepRank seemed to understand about language that we didn't capture before.

So here, the query is "can you get medicine for someone pharmacy." Before DeepRank, we surfaced a pretty good result about filling prescriptions, but it seems to miss the nuance that it's not just about filling prescriptions. You want to know if someone else can pick up your prescription at the pharmacy. So it missed that nuance.

After DeepRank, we seem to have captured that particular nuance, and we surfaced another good result that stemmed from HHS which is specifically about whether a patient can have a friend or family member pick up the prescription for them.

So what we saw in that evaluation is that language understanding that BERT provided DeepRank really manifested itself in us being able to understand more of the nuance in queries like this, these longer, more complex queries.

Q. The other search product that you identified on the

timeline derived from BERT was Deep Embed BERT --

A. RankEmbed.

Q. Excuse me. RankEmbed BERT. What was RankEmbed BERT?

A. RankEmbed BERT is essentially the idea you were talking about earlier. Remember we had this notion of embeddings for words.

This thing takes it one step further. It says why don't we embed queries into that space, and why don't we embed documents into that space, and let's do it in a way that if a document is close to a query, then the document is relevant to that query, so let's create the same mapping in a manner similar to what we had done before.

So we've got this mapping here, and now if you hide -- if you embedded all these queries in documents in this way, if you're given a new query, you embedded into the space and you look in the neighborhood around it for documents that are close by and you retrieve those documents. So you can augment your retrieval that we needed to do, and it turns out that this thing retrieves some really great documents, particularly for long-tail queries.

Q. You mentioned long-tail queries. Overall, what has been the impact of RankEmbed BERT on search?

A. RankEmbed BERT was again one of those very strong impact things, and it particularly helped with long-tail queries where language understanding is that much more important.

Q. All right. Continuing on, in 2021, there's a reference to MUM.

What was MUM, Dr. Nayak?

A. MUM was a -- one of the first generation of the so-called large language models that we developed particularly for search, and it proved to be incredibly valuable for many, many different aspects of search.

Q. And I see above the line there's another article with regard to the T5.

What was that?

A. So T5 was again a paper published by researchers at Google. They introduced a particular neural network architecture that we found to be very useful that we used in MUM.

Q. All right. Allen, if you could advance the slide.

Dr. Nayak, under the heading "multitask unified model," or MUM, there are a number of bullet points here. The first refers to "1,000 times more powerful than BERT."

A. Yeah.

Q. What does that mean?

A. So this is a fairly simple idea. The network on the right is the T5 network, and it's got a number of parameters, a very large number of parameters. The "thousand times more powerful than BERT" simply says there were a thousand times more parameters than the BERT model had. So it was a much larger model, and as a result, it was able to exhibit much more

powerful, more capable behaviors.

Q. And then going down those bullets, in the third bullet, there's the statement "understand and generate human language." Let's start with the first half of that first. Actually, let's do both parts.

How do you know that MUM understands and can generate human language?

A. So in the public domain, there are various metrics that have been developed. The most recent one that people use are these SuperGLUE metrics, which is a measure of how well these big, large language models are doing for tasks like understanding and generating human language.

MUM scored above 90 on the SuperGLUE scale. Just to give a sense, they believe that human performance on that particular collection of problems gives you a SuperGLUE score of about 90. So at least on the problems described there, MUM was sort of at human performance.

Q. What is MUM trained on, Dr. Nayak?

A. MUM is trained on a high-quality subset of the web corpus. So it's part -- it's trained on that.

Q. Let's go to the next slide, please.

Here, we have a query on the right side. Could you explain what's going on there?

A. So the context of this slide is that we used MUM to improve many, many different aspects of search. So you can see we

powered some 90 launches in the past year using MUM.

One of the launches that's sort of highlighted here is in the web answer space, and this is where we actually try to identify the particular passage in the document that answers the question that you're looking for.

So in this case, the question is "can worms have seizures." And before the MUM-powered version of the system, we sort of didn't quite get this right. We surfaced a passage about whether worms can cause seizures in humans. Right? But of course, we know that that's not what the query is asking for. They're asking if worms can actually have seizures.

And after MUM, I think we understood that, and we were able to surface a passage that indeed they do get some epileptic seizures in these tiny round worms that sort of look like seizures, which is quite fascinating.

Q. And on the left, there are a number of observations about MUM in search. The first refers to "improvements in IS."

A. Yeah.

Q. Did MUM -- go ahead.

A. We used MUM on some of our systems that improved IS and RankEmbed in particular, and it led to very significant improvements for all the reasons that we've described. It was just a more powerful way of understanding language.

Q. And the next piece of this slide reads "most wins in understanding long-tail complex queries."

First of all, what does "wins" mean there?
A. Wins is where using the system that was powered by MUM, using that led to improvements in search in some way compared to the old system. We call those wins.
Q. And why is it that MUM in particular had this effect, this most wins with respect to long-tail complex queries?
A. I think it gets back to the point we've made, which is on long-tail and complex queries, language understanding plays a crucial role. And as machine learning has advanced, as the technology has advanced to understand language better, we're able to handle long-tail and complex questions that much better.
Q. All right. Then on the bottom, it indicates "powered 90 launches in past year."
What's being conveyed by that?
A. I think the main point conveyed there is that MUM wasn't some one singular launch. Rather, it was used in a variety of different ways. Different projects used it in different ways, but it is used in a variety of different ways to improve many, many different aspects of search, so not just the ranking of results, but many of the features and so forth.
In particular, it powered like 90 different launches in this way.
Q. Allen, let's advance this slide, if we could.
What is this, Dr. Nayak?
A. So this is just a small sample of those 90.

MR. DINTZER: Your Honor, I have to object. If this was powered launches in the last year, that means this powered launch is without any depositions or documents related to this.

COURT REPORTER: Counsel, if you could turn on the mic.

MR. DINTZER: That means this is addressing material that was not turned over to us in discovery, we don't have the information on, and to the extent that the witness is relying on this to talk about the importance of MUM and how powerful it is and great, we don't have that, and we don't have the ability to challenge him about the accuracy or the completeness of his testimony, given that.

And so he's welcome to testify about all the stuff, we haven't challenged any of the stuff that he's done so far, but this is expressly stuff that has been created outside of the discovery window that we had the ability to contest and ask.

I can go further, that they're using him as an undisclosed expert on this area, but we believe that that's enough to keep this out.

THE COURT: Let's at least start with the question of when these were -- when these launches were released. That might be a helpful way to suss this out.

MR. SMURZYNSKI: Your Honor, if I could just make an observation. We heard from Dr. Ramaswamy about events that occurred in 2023.

THE COURT: I know.

MR. SMURZYNSKI: We've heard from others about events that have occurred since the close of discovery. I'm not going into great depth on this. But I will elicit the information you are interested in, Your Honor.

BY MR. SMURZYNSKI:

Q. First, Dr. Nayak, when was MUM launched by Google?

A. We can look at the timeline to make sure, but I think it's 2021.

Q. Okay.

A. We announced this at Google I/O.

Q. Looking at that slide 014, in the upper left-hand corner, there is a reference to a launch with regard to helping users in crisis.

When was MUM first used with respect to queries of that nature?

A. I think this was early last year, early in 20- -- yeah, for MUM, this was early last year, I think.

Q. Okay. And what is being done in that launch?

A. What we did here, so we had -- we sometimes get queries from users who are in crisis of one sort or the other. Sometimes, they're suicidal, and they come to ask us questions about their state.

And in those cases, when we detect that a query refers to someone who is in a crisis, who is suicidal, we believe it is

useful to surface a one box, an experience that points them towards help lines to help them in those crisis situations.

So we had launched this earlier. And a key part of doing that is to make sure that the query really refers to a user in crisis. So you don't want to trigger this experience if the query doesn't refer to someone in a crisis.

So if a user comes to us with a query Suicide Girls, for example, that's actually a music band, and surfacing this would not be helpful in that case.

So you need a classifier that says oh, look, this feels like a crisis query, please trigger this.

We had launched something with an earlier classifier, but with MUM, we were able to do a much better job of detecting whether a query was or was not a user in crisis like this. We were able to double the coverage in seven languages and launch this experience in 16 different languages all told. So a very focused, small-use case, you might say, but in some ways, an incredibly important use case for something like MUM.

Q. Let's go back to the timeline. Dr. Nayak, there are references here to LaMDA, PaLM, and PaLM2.

At a high level, what are these?

A. So LaMDA itself was also put out in 2021 in Google I/O. And it was one of these large language models that was particularly focused on conversation. It was very good at conversation.

And PaLM and PaLM2 have built on that. They've become larger. They have more capabilities than LaMDA. And these are all systems that our research teams have built.

Q. And then below the line, Dr. Nayak, there's a reference to "2023 SGE." Just at a very high level, what is SGE?

A. SGE stands for search generative experience, and it's an experience that we've built that takes all this excitement around generative AI and brings it to search.

Q. And are there others within the search organization who have focused more on SGE than you?

A. Yeah, one of the other VPs in search, Liz Reid, she's the one primarily responsible for SGE.

Q. Okay. Now, you've touched on this briefly as we've gone through the timeline, but I would like to talk about overall the data that's being used in these models, these applications. And let's focus first on the top, above the line, the research applications.

Is there any search user click and query data that is used to train those models?

A. No. They usually work off of open web corpus or other corpora that have been developed outside. They do not use search data.

Q. Okay. And I would like now to talk about the applications below the line, those in search. If we could advance the demonstrative to 016.

And some of these numbers, Your Honor, are redacted, but they should be on the deck, and you should have them in front of you.

Could you walk through sort of the evolution of the use of click and query data with respect to these applications that use these large language models and ML that we've been talking about this morning?

A. Yes. All of these, at least the first three classes of models, they use both search logs data, which is click and query data, but they also use IS scores, the human-rated scores that we had generated. So two sources of data are used there. RankBrain uses a few months' worth of search logs. DeepRank uses a little bit more than half of what RankBrain does. And RankEmbed uses a small fraction of the data that DeepRank uses, a small percentage of what DeepRank uses.

Q. And then let's continue on with regard to MUM.

How is click and query data used, if at all, in MUM?

A. MUM itself, as we have discussed before, is pretrained on a corpus of web data. It's a subset of the web corpus. The specific applications was 90-plus applications that we mentioned. They bring some amount of training data for their specific tasks. Some of them may be click data. Some of them may not be click data, like the one in the example we talked about. They had a small training set of queries that represented users in crisis, and there was no click and query

data for that. But there were applications that also used it, but it was the applications that used it, not MUM.

Q. What has been the trend over time in terms of the amount of click and query data and its significance in these models?

A. I think the trend is clear, is that you need less and less click and query data, that what these models are doing is they're very powerful at generalizing, and they're generating very good language understanding. And as a result, you need less and less query data to get that additional benefit from them.

THE COURT: Can I ask what is undoubtedly an oversimplification in a simple question: Can you explain to me how these models interact when it comes to a search? Is it that one model fully replaces the other? Do they do it on top of one another? How would one conceptualize the relationship among these models and affecting search quality?

THE WITNESS: That's a great question. This really goes back to that philosophical point that we had made earlier about how we build search ranking, which is we want to understand it.

Now, when you get to these deep learning systems, they're much harder to understand. But we want to use them, because they're very powerful. So what we have done is we don't turn over the ranking as a whole to these large models. Rather, we still have an infrastructure that we understand, a series of

ranking functions with signals coming in. And the outputs of these models are additional signals into that ranking function.

And so they're used as additional signals, maybe very powerful signals, maybe very informative signals, but nonetheless additional signals that get balanced both against each other as well as against other signals, like our page quality signals and so forth.

So there is no sense in which we have turned over our ranking to these systems. We still exercise a modicum of control over what is happening and an understandability there.

Q. I'm going to switch topics now. Are you familiar with Microsoft's Bing?

A. Yes.

Q. What is Bing?

A. Bing is a search engine.

Q. Does Google Search ever conduct comparisons of itself to Bing?

A. Yes, we do.

Q. And how does Google go about doing that?

A. It's very much like how we evaluated the quality of Google Search itself. We start with a sample of queries, and we see what results Google and Bing generate for those, and we get our raters to rate them. And so we get an IS score for Google and an IS score for Bing, and that's how we do the comparison.

Q. Why does Google compare itself to Bing in particular in the

search organization?
A. The -- we actually do more than just comparison to Bing. We have regular comparisons to Bing, because they're relatively easy to do, because the way Bing operates is very similar to the way Google operates. But we do comparisons with other places that users look for information. We just don't do it as frequently, because it's not as easy to do it. But we get insights, more qualitative insights in those cases on how people are looking for information.

So, for example, recently, we've been doing comparisons with TikTok, where young people particularly are increasingly turning to TikTok for their information needs, and we want to understand what is it that they're doing there, what are they finding useful, what should we do with Google to address that.

MR. DINTZER: Objection, Your Honor. To the extent he's going to go any further about what they've recently done with TikTok, it hasn't been turned over to us and given a chance to examine. I think at this level, it's fine, but anything more than that, we would have a significant objection.

THE COURT: Let's see where the questioning goes.

BY MR. SMURZYNSKI:

Q. The question simply is, why is it harder to compare something like Google Search to whether it's TikTok or Facebook before it or Amazon, a system such as that as compared to Bing?
A. I think the fact that Bing looks so much like Google and we

have all the infrastructure set up for evaluating Google makes it easier to make those comparisons.

The others, they look different in various ways. And so you have to do more qualitative analysis, and it's not something that you can automate in quite the same way.

Q. And over those 19 years in which Google's been doing those comparisons to Microsoft's search product and other search engines of the same ilk, what has Google seen in terms of the quality differences?

A. I think we've seen a fairly meaningful difference in quality. I would guess in the range of three to four points of IS at various points is the gap we've seen.

Q. Dr. Nayak, it's been suggested by plaintiffs that perhaps Google Search took the approach of being good enough.

Was it ever the case in your years at Google that its approach to search quality was to be good enough?

A. No, not at all.

Q. And what is Google's culture?

A. We very much have a culture of trying to improve search for our users. We are consumed by this. We spend all our time doing this. We set ourselves goals to continually improve search.

And frankly, there's a lot of work to be done for all sorts of reasons. The most interesting of them is every time we improve search, users ask us harder questions. There's always

this sort of boundary of questions that we don't do a good job on. And so there's always lots of work to be done, and we continually expand the boundaries as we go. So there's no sense in which search has ever been good enough.

MR. SMURZYNSKI: Your Honor, I have no more questions for the witness.

I'd like to move into evidence the demonstrative in the same fashion that we have been doing that for others, DXD17.

THE COURT: Okay.

MR. DINTZER: As long as it's coming in as a demonstrative, we don't have any objection, Your Honor.

MR. SMURZYNSKI: In the same fashion that we've been doing this with the others.

THE COURT: So it will be accepted for that purpose.

(Exhibit DXD17 received into evidence.)

THE COURT: Okay. So this is an opportunity -- so you're finished with your direct examination?

MR. SMURZYNSKI: Yes, Your Honor.

THE COURT: Terrific. It's a little after 12:25. We will resume at 1:30. We will obviously take our lunch break.

Dr. Nayak, the same instruction as before, please do not discuss your testimony during the break. Thank you.

THE WITNESS: Thank you.

(Recess taken at 12:26 p.m.)

CERTIFICATE OF OFFICIAL COURT REPORTER

I, Sara A. Wick, certify that the foregoing is a correct transcript from the record of proceedings in the above-entitled matter.

/s/ Sara A. Wick — October 18, 2023

SIGNATURE OF COURT REPORTER — DATE

B

D

E

J

K

N

Q

R

U

V

W