08:06:45 1 IN THE UNITED STATES DISTRICT COURT

2 FOR THE DISTRICT OF DELAWARE

3

4 VB ASSETS, LLC,) VOLUME 2
)
5 Plaintiff,)
) C.A. No. 19-1410(MN)
6 v.)
)
7 AMAZON.COM SERVICES LLC,)
)
8 Defendant.)

9

10 Tuesday, November 3, 2023
8:35 a.m.
11 Jury Trial

12
844 King Street
13 Wilmington, Delaware

14

15 BEFORE: THE HONORABLE MARYELLEN NOREIKA
United States District Court Judge
16

17

18
APPEARANCES:
19

20 WILSON SONSINI GOODRICH & ROSATI PC
BY: NEAL C. BELGAM, ESQ.
21 BY: JAMES C. YOON, ESQ.
BY: MATTHEW A. MACDONALD, ESQ.
22 BY: RYAN R. SMITH, ESQ.
BY: BRAD TENNIS, ESQ.
23 BY: JAMIE Y. OTTO, ESQ.

24
Counsel for the Plaintiff
25

APPEARANCES CONTINUED:

ASHBY & GEDDES
BY: STEVEN J. BALICK, ESQ.

-and-

FENWICK & WEST LLP
BY: J. DAVID HADDEN, ESQ.
BY: SAINA S. SHAMILOV, ESQ.
BY: RAVI R. RANGANATH, ESQ.
BY: VIGEN SALMASTLIAN, ESQ.
BY: MELANIE MAYER, ESQ.
BY: JEFFREY WARE, ESQ.

Counsel for the Defendant

- - - - - - - - - - -

08:22:54 THE COURT: All right, everyone. Please be
08:34:36 seated. So we have some objections. I know you guys gave
08:34:41 some stuff to Mr. Buckson, but I haven't looked at it, you
08:34:46 can do that all on the clock. Okay. What do we have?
08:34:50 MR. YOON: Your Honor, there are some issues
08:34:52 that have been made, but we did want to report to the court
08:34:55 that the parties did work through all the issues with
08:34:58 Dr. Polish.
08:34:58 THE COURT: We're wasting the jury's time, I do

08:35:02 1 appreciate that and thank you, but let's not waste time.
08:35:06 2 MR. SALMASTLIAN: Good morning, Your Honor,
08:35:09 3 Vigen Salmastlian. There is an issue with a source code
08:35:14 4 exhibit that --
08:35:15 5 THE COURT: I have this, this is the only
08:35:16 6 document that I have right now. So do you have an objection
08:35:19 7 to this?
08:35:21 8 MR. WARE: Yes.
08:35:22 9 THE COURT: Let's deal with it.
08:35:24 10 That was the one that was actually timely made,
08:35:29 11 so let's deal with that first.
08:35:30 12 MR. WARE: Jeffrey Ware for defendant.
08:35:33 13 And here Amazon is using -- sorry, let me just
08:35:39 14 put this -- Your Honor, may I put this on the Elmo?
08:35:43 15 THE COURT: Yes.
08:35:46 16 MR. WARE: So we are re-raising Amazon's Daubert
08:35:51 17 to the portion of the Daubert motion that was denied with
08:35:56 18 leave to re-raise at trial. It may be appropriate, or more
08:36:02 19 appropriate for this Court to consider this issue after the
08:36:06 20 damages experts have testified and the Court has gotten more
08:36:11 21 information regarding the specifics of the methodology
08:36:14 22 Amazon is objecting to, and so for the appreciating the
08:36:21 23 issue of time here, Amazon simply re-raises--
08:36:25 24 THE COURT: No. No. You're not re-raising
08:36:28 25 something. Tell me what the objection is.

08:36:31 MR. WARE: Okay.

08:36:31 THE COURT: All you're going to do is say I
08:36:37 re-raise what I did, you know, weeks ago or even a week ago,
08:36:40 I have no idea what that is, okay? So if you want to make
08:36:44 an objection, you're -- it's worth your time and you make
08:36:48 it. Because I don't know what the problem is with it.

08:36:51 MR. WARE: Your Honor, so in the -- in Your
08:36:53 Honor's Daubert order, you granted the motion partially with
08:36:58 respect to Mr. Reed's math error with respect to his growth
08:37:05 input that he used as the basis for his hundred -- at that
08:37:10 time about $150 million damages opinion. Now it is down to
08:37:15 46.7 million. Amazon objects, though, to -- and that was
08:37:21 part of his cost savings methodology or so-called cost
08:37:25 savings methodology.

08:37:27 Amazon still objects to the underlying cost
08:37:31 savings methodology as raised in the motion and specifically
08:37:35 --

08:37:35 THE COURT: You have to tell me what it is, I
08:37:37 don't remember. I have no idea what that basis was. So
08:37:41 rather than recap and waste that time, how about you just
08:37:45 start a new -- what did he do, what's objectionable?

08:37:51 MR. WARE: Your Honor, his cost saving
08:37:53 methodology compares Alexa's turn rate, versus for dialogs,
08:38:03 a turn rate is you speak to Alexa, Alexa provides a
08:38:07 response, that's one turn. He took Alexa's turn rate and

08:38:13 1 compared it -- which right now they accuse the '681 patent
08:38:19 2 of infringing, and then he took Alexa Shopping's turn rate
08:38:26 3 which plaintiff accuses the remaining patents of infringe --
08:38:32 4 or Amazon of infringing the remaining patents, and because
08:38:36 5 at one point in time in 2018, there was a 1.5 turn
08:38:41 6 difference, and he further opines that there was a certain
08:38:47 7 thing that happened with Amazon Shopping, a simple CX
08:38:52 8 program that, at that time, made it so that he could support
08:38:57 9 the valuation for his 1.5 turns benefit analysis he compared
08:39:05 10 Alexa Shopping with Alexa active customers and so the basis
08:39:09 11 of his calculation if you look here on this -- on the
08:39:15 12 demonstrative is that he said because Alexa Shopping had
08:39:23 13 poor customer experience prior to this time, you can see
08:39:27 14 when Alexa Shopping, which is in red, goes way up in late
08:39:32 15 2018, that the benefit of 1.5 turns could be calculated
08:39:37 16 based on the difference here pre the September 18 of Alexa
08:39:43 17 and Alexa Shopping user growth.
08:39:46 18 So Amazon's first objection to cost saving
08:39:52 19 methodologies is no cost saving methodology at all except by
08:39:58 20 the Federal Circuit, because here at this time, plaintiffs
08:40:02 21 have accused Amazon of infringing both Alexa, and Alexa
08:40:08 22 Shopping, you can't take the difference of growth rates in
08:40:12 23 both to determine the cost savings of using either the '681
08:40:18 24 patent for Alexa, or the remaining asserted patents for
08:40:22 25 Alexa Shopping.

08:40:23 1 So that is the first basis of Amazon's objection
08:40:29 2 that cost savings is supposed to be as recognized by the
08:40:32 3 Federal Circuit, the difference between infringing and a
08:40:40 4 non-infringing use, usually a non-infringing alternative or
08:40:44 5 sometimes the alternative of actually doing nothing. In
08:40:48 6 this case, they're comparing two infringing products --
08:40:51 7 THE COURT: But infringing of different patents,
08:40:53 8 so it's an infringing of the '681, and non-infringing of the
08:40:57 9 '681, but infringing of the others?
08:41:00 10 MR. WARE: Well they accused '681 actually of
08:41:04 11 infringing all -- or they accused both Alexa and Alexa
08:41:08 12 Shopping of infringing all of the asserted patents, but only
08:41:13 13 specifically Alexa Shopping for infringing the other
08:41:16 14 asserted patents.
08:41:22 15 THE COURT: So they're comparing the red line is
08:41:25 16 alleged infringing '681, and the blue line is a different
08:41:29 17 product that also is alleged to infringe the '681?
08:41:33 18 MR. WARE: Yes, it's the Alexa Shopping growth
08:41:36 19 rate. I'm sorry, the blue is Alexa, and red is Alexa
08:41:41 20 Shopping.
08:41:43 21 THE COURT: Okay. What's your next basis?
08:41:46 22 MR. WARE: The next basis is it's an apples to
08:41:50 23 oranges comparison, and in order to effectively assert that
08:41:56 24 because Alexa worked well at the time because of the
08:42:00 25 patents, and even though Alexa Shopping wasn't working well

08:42:06 1 before this change in 2018, even though again it was using
08:42:10 2 -- it was practicing the other patents, that the 1.5 turn
08:42:14 3 rate --
08:42:15 4 THE COURT: You keep saying the other patents.
08:42:17 5 So --
08:42:22 6 MR. WARE: So the voice ad patents and the '703
08:42:26 7 patent.
08:42:26 8 THE COURT: Yes, I know. Where I'm getting
08:42:29 9 confused is what is -- you just -- when I thought that's
08:42:33 10 what you were talking about before, '681 versus the other
08:42:36 11 patents, you said no, no, everything infringes, is alleged
08:42:41 12 the infringe the '681, so now we're talking about something
08:42:44 13 different, right? Now you're saying one thing is the '681
08:42:47 14 and another thing is the other patents?
08:42:49 15 MR. WARE: Well, Your Honor, the '681 patent
08:42:53 16 claims the more -- just most generally may call it
08:43:00 17 contextual understanding using short and long-term
08:43:03 18 knowledge. And so, they assert that all Alexa interactions
08:43:09 19 infringe the '681 based on that, whether it's in the Alexa
08:43:14 20 Shopping domain, or in any other Alexa domain.
08:43:17 21 With respect to the voice ad patents and the
08:43:20 22 '703 patent, they only accuse Alexa Shopping functionality
08:43:27 23 of infringing those patents.
08:43:30 24 THE COURT: Okay. So with that understanding,
08:43:32 25 can you tell me your second basis again?

08:43:35 MR. WARE: So the second basis is that to use
08:43:40 the different turns between Alexa and Alexa Shopping, is
08:43:46 completely baseless and erroneous because the context for
08:43:50 Alexa and using Alexa more generally and Alexa Shopping are
08:43:55 very different, or quite often very different. 95 percent
08:43:58 of Alexa interactions are just one shot interactions, where
08:44:03 you may ask her "Alexa, turn on the lights," or "please play
08:44:10 me my play list," or please do any number of things, where
08:44:16 it doesn't require an extra dialogue turn to do anything
08:44:19 because it's Alexa just automatically does it. Where Alexa
08:44:25 Shopping is often a more by-design, involves multiple turns.
08:44:30 So you may ask for a product and then you may then, want to
08:44:36 put that product in the shopping cart and then ultimately
08:44:39 you may want to purchase that product. And you can do that
08:44:42 using voice or any number of -- or screen or anything else,
08:44:45 but that can be a multiple turn dialogue. And really by
08:44:51 design by just the specifics of the whole purpose of the
08:44:57 Alexa Shopping app.
08:44:59 So to compare turn differences and attribute
08:45:04 that to the benefits of the patent while ignoring the true
08:45:08 difference of Alexa and Alexa Shopping, is an erroneous
08:45:16 comparison and it's comparing apples and oranges, not
08:45:20 actually comparing anything that reflects the value of the
08:45:22 patents to the accused product.
08:45:25 THE COURT: Okay. Any other grounds?

08:45:27 1 MR. WARE: And the third ground is that he is
08:45:30 2 pointing to a single Alexa Shopping document, OP1 and the
08:45:35 3 purported improvement of Alexa Shopping after certain
08:45:40 4 improvements were made, but he's applying the value of those
08:45:45 5 improvements to Alexa overall. Where again, specifics of
08:45:51 6 Alexa Shopping and Alexa are very different. And on top of
08:45:58 7 that, he fails to show that even when you see the red line
08:46:03 8 moving up, he fails to show that the red line moving up is
08:46:07 9 actually caused by reduction of turns in the patent, or
08:46:10 10 reduction of turns in dialogue which is the purported
08:46:15 11 benefit that Mr. Reed claims, and is trying to monetize for
08:46:20 12 plaintiff.
08:46:21 13 And so again, that's -- it makes no logical
08:46:26 14 sense, and we would object under Rule 702 and Daubert for
08:46:31 15 that.
08:46:31 16 THE COURT: Now, when you said before that this
08:46:36 17 might make more sense to raise after the experts testify,
08:46:41 18 wouldn't the expert already have testified about this, so
08:46:44 19 doesn't that suggest that this is not really a reliability
08:46:48 20 issue if you're willing to let the expert talk about this
08:46:50 21 before I rule on it?
08:46:52 22 MR. WARE: Well, Your Honor, we believe that you
08:46:56 23 already considered this issue when you did rule on the
08:46:58 24 initial Daubert motion, and you had suggested that this may
08:47:01 25 be an issue for cross-examination. But you granted Amazon

08:47:06 1 leave to re-raise it at trial. We think and know that if a
08:47:11 2 damages verdict issues by the jury in this Court, that the
08:47:17 3 Federal Circuit will overturn because again, this makes no
08:47:23 4 -- first of all makes no logical sense, and it also fails to
08:47:27 5 follow any recognized cost savings methodology that's ever
08:47:31 6 been accepted by the Federal Circuit.
08:47:35 7 THE COURT: I guess we'll have to find out.
08:47:38 8 Let's hear. And I really want to hear mostly on the first
08:47:44 9 ground which is how can this -- how can this have anything
08:47:49 10 to do with infringing the benefits of the patent when you
08:47:55 11 are comparing infringing versus infringing.
08:48:05 12 MR. MACDONALD: Thank you, Your Honor. Matt
08:48:11 13 Macdonald on behalf of VB Assets. Let me just get straight
08:48:14 14 to Your Honor's question.
08:48:15 15 THE COURT: That would be good without even the
08:48:17 16 present preface of telling me, just get there.
08:48:19 17 MR. MACDONALD: The answer is, we cannot run an
08:48:23 18 experiment inside Alexa that compare what would happen in
08:48:25 19 the but for world, where there is infringement and
08:48:27 20 non-infringement. We can't do that, we don't have that
08:48:29 21 ability because our position is that Amazon continuously
08:48:32 22 infringes, and we don't control their system.
08:48:35 23 So what our expert did was look through Amazon
08:48:38 24 data and Amazon documents to attempt to find evidence that
08:48:41 25 would allow him to proxy for the effect of the infringing

08:48:45 alternative.

08:48:46 Mr. Polish -- starting here, the benefit of the
08:48:50 patent is that it simplifies interaction with Alexa. He
08:48:54 looked for evidence that would help him measure that. What
08:48:57 he found is basically three things that I will highlight.
08:49:00 Number one is an internal Amazon document that is not up on
08:49:03 the screen, even though I was looking there, that Amazon
08:49:06 itself looked at a comparison between Amazon -- or excuse
08:49:11 me, between Alexa and Alexa Shopping and it itself made that
08:49:15 comparison as a way to look at the effect of additional
08:49:19 complexity on user growth. In effect, Amazon in that
08:49:23 document says we see a lot more turns on Alexa Shopping than
08:49:28 on Alexa and then connects that to a ten percent growth
08:49:32 differential between the two.

08:49:33 The second piece of evidence --

08:49:35 THE COURT: Why does that have anything to do
08:49:37 with the patents? Apportioning to the patents, that's what
08:49:40 I'm not understanding. As opposed to some other value use.

08:49:47 MR. MACDONALD: What happens is Amazon in this
08:49:49 document equates a 1, 1-and-a-half I believe turn
08:49:53 differential in Amazon Shopping and attributes that to a ten
08:49:57 percent growth differential.

08:49:58 THE COURT: Yes, but you're not telling me what
08:50:00 that one-and-a-half turns has to do with the patents versus
08:50:06 -- I mean, both things that are supposed to be using the

08:50:11 1 patent, right?

08:50:12 2 MR. MACDONALD: Yes, Your Honor.

08:50:13 3 THE COURT: The one-and-a-half turn, is that a

08:50:15 4 patented feature?

08:50:16 5 MR. MACDONALD: No, the one-and-a-half turns is

08:50:20 6 what Dr. Polish will testify is the consequence of using the

08:50:22 7 patent.

08:50:23 8 THE COURT: But using which patent?

08:50:25 9 MR. MACDONALD: All of the patented technology.

08:50:27 10 THE COURT: Okay. So that's where I'm missing

08:50:29 11 this. You get the benefit of the one-and-a-half turns from

08:50:33 12 that, then how -- don't the blue things also have the

08:50:37 13 benefit of the patents because you're accusing them of

08:50:40 14 infringement?

08:50:41 15 MR. MACDONALD: They absolutely do. So the

08:50:43 16 question that Dr. Reed -- excuse me, that Mr. Reed attempted

08:50:47 17 to answer was what evidence in Amazon's files that I have,

08:50:51 18 that would allow me to quantify the impact of this 1.5

08:50:56 19 turns. And we can't run an experiment --

08:50:58 20 THE COURT: I don't understand how you can say

08:51:00 21 the effect of infringing the patents is you get this benefit

08:51:03 22 when you're accusing things of infringing the patents that

08:51:06 23 don't apparently have that benefit.

08:51:10 24 MR. MACDONALD: I think the answer is that we

08:51:12 25 don't have -- we don't have -- we don't have real world

08:51:16 evidence of what would happen in the but for world --

08:51:18 THE COURT: Are you saying the blue things use

08:51:20 the patents?

08:51:21 MR. YOON: Yes, the blue things use the patents.

08:51:23 THE COURT: How is Dr. Polish saying that that

08:51:25 gives you the one-and-a-half turns, that's what I'm missing.

08:51:28 And stop with this double teaming. Then he's done and

08:51:31 you're making the argument.

08:51:33 MR. YOON: Yes, Your Honor, I apologize.

08:51:34 THE COURT: Fine. One time and that's it.

08:51:36 MR. YOON: I apologize, Your Honor, just to be

08:51:38 efficient.

08:51:39 THE COURT: I hate this lurking in the

08:51:40 background, very distracting, I can't pay attention to

08:51:44 what's being said because I got this lurking.

08:51:46 MR. YOON: My apologies, Your Honor. Your

08:51:48 Honor, in this case, Amazon's own expert, Mr. Johnson,

08:51:53 indicated that a non-infringing alternative to the patent

08:51:58 would add one -- at least one turn to each of the Amazon

08:52:04 devices that would infringe. So what you have here, Your

08:52:07 Honor, is you want to tie it to the patent is Amazon's

08:52:12 expert --

08:52:12 THE COURT: I don't want to tie it to the

08:52:14 patent, I think you have to tie it to the patent if you want

08:52:16 to use it, right?

08:52:17 MR. YOON: I am tying it to the patent, Your
08:52:19 Honor, I understand. Amazon's expert had indicated that you
08:52:22 can have a non-infringing alternative by asking more
08:52:26 questions that would add turns. Also Dr. Polish indicated
08:52:29 that as well in response to Amazon's expert. So you have a
08:52:33 situation where both experts on the technical side agree
08:52:36 that a non-infringing alternative to the patented technology
08:52:39 would add turns.
08:52:43 And then the next step is, okay, what is the
08:52:45 consequence of adding turns? We have an Amazon document
08:52:49 that shows adding 1.5 turns reduced growth by ten percent.
08:52:56 And then this was the issue with the math error, Your Honor,
08:52:59 this was the point, we have technical expert testimony that
08:53:02 shows the patents reduced turns, and that the non-infringing
08:53:06 alternative to the patents adds a turn. Then we have --
08:53:10 when we then have to quantify the economic impact of what
08:53:15 adding a turn would mean, we then go to the Amazon
08:53:19 documents, Amazon's own study which shows that when you
08:53:22 added 1.5 turns, it reduced growth by ten percent.
08:53:26 And so, that is the patents reduce the number of
08:53:31 turns, all the patents reduced the number of turns by at
08:53:34 least one --
08:53:35 THE COURT: Why do you have apportioning value
08:53:37 to the patents at the top, that's why I'm getting confused,
08:53:41 it makes it sound like you're comparing patented to

08:53:44 1 non-patented in this slide, and you're not.

08:53:46 2 MR. YOON: We can remove that title.

08:53:47 3 THE COURT: I understand what you're saying is

08:53:49 4 everybody agrees a non-infringing alternative has an extra

08:53:52 5 turn, and we're just trying to approximate what value not

08:53:57 6 having that extra turn gives.

08:53:59 7 MR. YOON: Exactly, we have Amazon's own

08:54:02 8 internal study as to the consequence of having more turns,

08:54:05 9 and Amazon's own document shows that it reduced growth by

08:54:09 10 ten percent, so the rest of this would go to the weight,

08:54:12 11 Your Honor, and it was addressed earlier.

08:54:14 12 THE COURT: All right. Let me hear the response

08:54:17 13 on that.

08:54:19 14 MR. WARE: Your Honor, regardless of whether the

08:54:22 15 patented benefit is 1.5 turns for Alexa, and 1.5 turns for

08:54:28 16 Alexa Shopping --

08:54:29 17 THE COURT: That's not the question. The

08:54:31 18 question is, respond to his argument. His argument is you

08:54:36 19 all say there is a non-infringing alternative, is that

08:54:40 20 right?

08:54:40 21 MR. WARE: Yes.

08:54:40 22 THE COURT: And the non-infringing alternative,

08:54:43 23 he says, that it would have at least an additional turn.

08:54:50 24 And you guys can disagree as to whether that's true, but

08:54:54 25 that's his understanding. He's saying look -- he's not

08:54:58 1 going to put that title up there, because I don't think this
08:55:00 2 is apportioning value to the patents. He's saying all this
08:55:03 3 is doing is approximating how much value you lose or gain by
08:55:12 4 having an additional turn, or one less turn, so that's what
08:55:15 5 I need you to respond to.
08:55:17 6 MR. WARE: Well, in the lines before
08:55:20 7 September 18th, the blue line has 1.5 less turns because
08:55:27 8 they're infringing the patents according to them, and the
08:55:29 9 red line has 1.5 less turns because they are accused of
08:55:34 10 infringing the patent.
08:55:35 11 THE COURT: That's not the question. You're not
08:55:37 12 responding to his argument. His argument is this is trying
08:55:50 13 -- forget the patents for one minute. He's saying all this
08:55:53 14 is doing is showing one thing has fewer turns, and it has
08:55:57 15 more value, and we're trying to approximate how much value
08:56:02 16 having one fewer turns is.
08:56:04 17 MR. WARE: Your Honor, that's what Mr. Reed is
08:56:06 18 attempting to correct, but the issue with cost savings is
08:56:10 19 then he is trying to compare one thing that enjoys the
08:56:13 20 benefit with another thing that enjoys the benefit of the
08:56:16 21 patented technology, as opposed to cost savings methodology
08:56:21 22 which is where he fails, which is supposed to compare
08:56:25 23 infringing versus non-infringing, and here they're comparing
08:56:28 24 I think --
08:56:29 25 THE COURT: Yes, but he's trying to compare as I

08:56:32 1 understand it, is infringing versus non-infringing. He says
08:56:35 2 the infringing has one-and-a-half fewer turns, and the
08:56:39 3 non-infringing alternative that you all say would be out
08:56:42 4 there has 1.5 more turns. And so he's comparing infringing
08:56:47 5 and non-infringing, and what he's trying to do here is to
08:56:51 6 say well, Amazon doesn't have a product because there is no
08:56:56 7 non-infringing product, so we have to come up with a
08:56:58 8 hypothetical or some way to evaluate what that amount would
08:57:02 9 be.
08:57:03 10 So we try and figure out what the value of that
08:57:05 11 extra turn for the infringing versus non-infringing is.
08:57:10 12 MR. WARE: Well, Your Honor, then they could
08:57:12 13 have performed a survey where they compared the benefit of
08:57:15 14 the patent compared to not -- to implementing the specifics
08:57:21 15 that Dr. Johnson included for non-alternatives in --
08:57:27 16 THE COURT: You're saying that's the only way
08:57:29 17 they could do it is a survey, there is no other way, they
08:57:32 18 can't just compare things that we know what one value is,
08:57:36 19 and we see -- we see what the value of that thing is.
08:57:43 20 MR. WARE: That's not necessarily correct, Your
08:57:45 21 Honor, that is one way to do it.
08:57:47 22 THE COURT: Okay. Well they have a different
08:57:49 23 way of doing it, and you're saying well this is just wrong,
08:57:52 24 but you're not really responding -- you're more saying well
08:57:56 25 they're comparing patented to patented. And what I

08:57:59 1 understand him to be saying is no, we're comparing patented
08:58:02 2 to unpatented, and what we're trying to do on this graph is
08:58:09 3 estimate the value that is given between the alleged
08:58:15 4 non-infringing alternative and the value of the patent.
08:58:23 5 MR. WARE: That's what is claimed, you're
08:58:26 6 correct, Your Honor.
08:58:27 7 THE COURT: All you're telling me is that's not
08:58:29 8 right and I don't understand what you're telling me.
08:58:32 9 MR. WARE: Because for cost savings, he is not
08:58:35 10 again -- each one of these lines, the whole time, both of
08:58:38 11 those lines are enjoying the benefit of the 1.5 turn in a
08:58:44 12 different context. But rather than him comparing the
08:58:48 13 benefit of the 1.5 turns in that context versus not --
08:58:55 14 versus doing these non-infringing alternatives, what he's
08:58:58 15 doing is he's comparing one way of receiving 1.5 turn
08:59:03 16 benefit to another way of receiving a 1.5 turn benefit.
08:59:07 17 THE COURT: All right. You guys aren't going to
08:59:09 18 have any time left to try your case at the rate you're
08:59:12 19 going. What I am going to do is deny this motion again. I
08:59:15 20 think that he is comparing based on the representation I
08:59:19 21 have gotten, he's comparing non-infringing to infringing,
08:59:22 22 and trying to estimate based on the information that they
08:59:26 23 have the value that the patent gives. So I am going to
08:59:31 24 allow that. I also think the fact that when we started this
08:59:35 25 that counsel said look, you can just decide this after the

08:59:38 1 jury has already heard this, also lends credence to the fact
08:59:43 2 that this is an issue for cross rather than an issue of
08:59:46 3 reliability.
08:59:46 4 All right. So what's the next issue that we
08:59:49 5 have?
08:59:55 6 MR. SALMASTLIAN: Your Honor, the issue, the
08:59:57 7 next issue is related to the document AMZ_VB_SCPRINT 74 to
09:00:05 8 77, and the issue here is this is a source code document
09:00:12 9 that Mr. Peck raised for the first time in his reply report,
09:00:16 10 but it was not for any issue related to any patent that's
09:00:20 11 currently at issue in this case. He restricted his opinion
09:00:25 12 to be a response to paragraph 459 of Mr. Johnson
09:00:29 13 non-infringement report. And that part of Mr. Johnson's
09:00:33 14 opinion relates to '536 patent, Claim 3. That claim has
09:00:38 15 been long dropped, and that patent has been dropped from
09:00:41 16 this case. And so we have no idea what Mr. Peck is going to
09:00:44 17 say about this source code document with respect to any
09:00:48 18 patent that's currently at issue in this case.
09:00:51 19 MR. SMITH: Your Honor, Mr. Peck is a source
09:00:55 20 code expert and all he did was opine about the operation of
09:00:58 21 the source code. His opinion is agnostic with respect to
09:01:03 22 any particular patents --
09:01:05 23 THE COURT: So where did he say he's going to
09:01:07 24 rely on this document with respect to any issue that is
09:01:10 25 currently in the case?

09:01:11 1 MR. SMITH: In his reply report --
09:01:14 2 THE COURT: Show me where, because he's saying
09:01:17 3 it's not relevant to any patent left in this case.
09:01:20 4 MR. SMITH: So in his reply report on page --
09:01:25 5 towards the bottom of page 10, if we can put that on the
09:01:29 6 screen. It's exactly -- probably 20 percent of his report
09:01:37 7 is just on this particular file. So of the ten page report,
09:01:41 8 two pages are about this file. And all he's going to
09:01:44 9 testify to is exactly what he says in his report. It's not
09:01:48 10 agnostic as to any particular patent, it's just Dr. Johnson
09:01:52 11 is saying that the Amazon Alexa works in a certain way, and
09:01:56 12 he was disputing based on code.
09:01:57 13 THE COURT: But when you said it's agnostic, he
09:02:01 14 said this is response to '459. What was '459?
09:02:04 15 MR. SMITH: So if we go to '459, this was about,
09:02:09 16 then I'll quote, '459, Johnson says "Polish also accuses "a
09:02:16 17 dialog management feature that determines a context by
09:02:19 18 passing." And he goes on from there. And then he talks
09:02:21 19 about the context interpreter. And the context interpreter
09:02:25 20 comes up again and again and again with respect to all of
09:02:28 21 the VoiceBox patents that are being asserted. So
09:02:31 22 Dr. Johnson has a particular paragraph. I'll say that his
09:02:34 23 report was 500 pages, and every single argument incorporates
09:02:39 24 everything else. So it's basically a spiderweb of opinions
09:02:46 25 that cover everything seemingly.

09:02:50 1 THE COURT: So a Section 459 addresses all the
09:02:53 2 patents and incorporates everything else he said?
09:02:57 3 MR. SMITH: He said it addresses a feature which
09:02:59 4 is the context interpreter, which comes up again and again
09:03:02 5 and again through the allegations of VoiceBox in the
09:03:05 6 different patents he talks about this particular argument,
09:03:08 7 the file in this part of his report, but he's really also
09:03:13 8 just responding to something that Dr. Polish said about the
09:03:16 9 operation of Alexa, a dialogue management feature that
09:03:21 10 determines.
09:03:22 11 THE COURT: How is he going to use this now
09:03:24 12 because it wasn't important for his infringement analysis,
09:03:26 13 he had to do it on reply, how is he going to use this?
09:03:30 14 MR. SMITH: He's going to use it in the same way
09:03:32 15 that he had in his report where he's just --
09:03:34 16 THE COURT: You're holding an iPad, it's really
09:03:37 17 not helpful to me to see what's on it.
09:03:40 18 MR. SMITH: He's going to say exactly what he
09:03:43 19 said in his expert report, he's pointed out there a source
09:03:47 20 file for the rules-based context interpreter, and the fact
09:03:51 21 that Amazon is saying they don't use manual rules and they
09:03:54 22 have two pages in effect where they're saying they do, and
09:03:57 23 that is specifically dead on, it's the context interpreter
09:04:00 24 which is the heart of the infringement allegations for
09:04:03 25 several of these patents. He's pointing out that he uses

09:04:07 1 source code analysis is inconsistent with what Amazon is
09:04:10 2 saying.
09:04:17 3 MR. SALMASTLIAN: Your Honor, the opinion that
09:04:18 4 Mr. Peck provided has nothing to do with whether or not
09:04:21 5 Amazon uses rules, it was specifically in response to --
09:04:25 6 THE COURT: Yes, but Amazon is now saying they
09:04:27 7 don't use rules, right? And so why can't he say, well they
09:04:32 8 have rules. They do have files that are relevant to rules.
09:04:36 9 MR. SALMASTLIAN: I think the issue here, Your
09:04:37 10 Honor, is that he didn't provide an opinion on that with
09:04:40 11 respect to this code. And we didn't have a chance --
09:04:43 12 THE COURT: He did -- he did have an opinion but
09:04:46 13 you're saying he didn't with respect to one patent.
09:04:50 14 MR. SALMASTLIAN: This is all with respect to
09:04:52 15 the '563 patent that has been dropped in this case.
09:04:55 16 THE COURT: What about the fact that he said
09:04:56 17 your guy, this context, whatever interpreter comes up in all
09:05:00 18 the different patents, and everything was incorporated by
09:05:05 19 reference to everything else, so --
09:05:10 20 MR. SALMASTLIAN: So there was no incorporation
09:05:11 21 by reference to other things, Your Honor, and the fact that
09:05:14 22 there is a module in Alexa called context interpreter does
09:05:20 23 not tell us what the opinion is about that context
09:05:24 24 interpreter with respect to any other patent where the
09:05:26 25 claims have different requirement. So here, Dr. Polish only

09:05:31 1 addressed it with claim 3 of '536.

09:05:35 2 THE COURT: I am going to let this go, to the

09:05:37 3 extent that it's just saying look, there is this rules-based

09:05:40 4 context interpreter, if he starts getting into much more

09:05:44 5 specifics that weren't disclosed, then you can object again

09:05:48 6 and I'll reconsider. But I think it's fair to say that with

09:05:52 7 respect to the source code, it was disclosed that he was

09:05:55 8 going to note that there is a rules-based context

09:06:00 9 interpreter in the device.

09:06:02 10 All right. Anything else we need to talk about?

09:06:06 11 MR. SALMASTLIAN: There is just one other issue

09:06:07 12 Your Honor, and it might be resolved at this point, but

09:06:10 13 there was a late add source code exhibit this morning which

09:06:14 14 is Amazon -- go ahead.

09:06:16 15 MR. SMITH: We'll withdraw that in light of your

09:06:19 16 ruling.

09:06:19 17 THE COURT: All right.

09:06:20 18 MR. SALMASTLIAN: Thank you.

09:06:22 19 THE COURT: All right. Mr. Buckson, do you want

09:06:24 20 to get the jury.

09:06:26 21 MR. YOON: Your Honor, just one procedural point

09:06:27 22 just to let you know Your Honor, we're going to be starting

09:06:30 23 with Mr. Freeman. The parties agreed to add some additional

09:06:34 24 testimony to Mr. DiCristo, so we'll start with Mr. Freeman

09:06:39 25 right away. And we'll play DiCristo's testimony a little

09:06:43 bit later today.

09:06:44 THE COURT: All right. So I'm sorry, is Freeman

09:06:48 live?

09:06:48 MR. YOON: Yes, he's coming live.

09:06:50 THE COURT: And what was the agreement -- oh, to

09:06:52 Mr. DiCristo, so then you have to update the times.

09:06:56 MR. YOON: They had asked us, there is no

09:06:57 dispute, we just have to get it done Your Honor, so to not

09:07:01 delay anything, we'll start with Mr. Freeman.

09:09:01 (Jury entering the courtroom at 9:09 a.m.)

09:09:15 THE COURT: All right, everyone. Welcome back.

09:09:24 Thank you for being here. Everyone else can please be

09:09:26 seated.

09:09:27 What's next?

09:09:32 MR. SMITH: Your Honor, plaintiff calls Tom

09:09:36 Freeman to the stand.

09:09:49 COURTROOM DEPUTY: Please raise your right hand.

09:09:53 Please state and spell your full name for the record.

09:09:59 THE WITNESS: My name is Thomas Edward Freeman,

09:10:06 F-R-E-E-M-A-N.

09:10:07 THOMAS EDWARD FREEMAN, having been duly sworn,

09:10:13 was examined and testified as follows:

09:10:16 DIRECT EXAMINATION

09:10:17 BY MR. SMITH:

09:10:28 Q. Good morning, Mr. Freeman.

09:10:31 A. Yes, I am.
09:10:33 Q. Sorry, I said, can you hear me, Mr. Freeman?
09:10:36 A. Oh, can I hear you, I can indeed, thank you.
09:10:39 Q. Good morning sir, how are you?
09:10:41 A. I'm quite well, thank you.
09:10:42 Q. Mr. Freeman, could you briefly introduce yourself to
09:10:45 the jury?
09:10:46 A. Sure. My name is Tom Freeman, I was born in Seattle
09:10:50 Washington, that's the other Washington, and that I lived
09:10:54 there my entire life with a short stint in Anchorage,
09:11:00 Alaska. I was a co-founder of VoiceBox technologies. And I
09:11:05 worked there for nine years. And the reason I am here
09:11:09 today, I guess, is that I was responsible for the invention
09:11:14 of some of the important patents in this particular case.
09:11:19 Q. Mr. Freeman, if we could just go over your background
09:11:23 briefly. Did you attend college?
09:11:25 A. Both graduate and undergraduate.
09:11:27 Q. Where did you go?
09:11:28 A. I went to the University of Washington.
09:11:30 Q. Where is that?
09:11:30 A. That's in Seattle, Washington.
09:11:33 Q. What did you do professionally after college?
09:11:36 A. I was in a technology world in which interpreting the
09:11:41 value of the technology to the customer was extremely
09:11:44 important, so I think you would call this technical sales

09:11:49 and technical product management.
09:11:51 Q. And just generally, what sort of products were you
09:11:53 selling back then?
09:11:54 A. Everything from database publishing to relational
09:12:01 databases, to holographic beam forming metamaterial
09:12:07 techniques, it was quite a broad world of technologies.
09:12:11 Q. Do you have any formal engineering training?
09:12:14 A. Oh, no, I do not, sir.
09:12:15 Q. Did you learn technology on the job?
09:12:17 A. I did indeed, sir, to be able to interpret it is
09:12:21 extremely important.
09:12:22 Q. And Mr. Freeman, do you know the Kennewick brothers?
09:12:26 A. I do know the Kennewick brothers, I know them well.
09:12:30 Q. How did you first come to meet them?
09:12:32 A. Rich and I worked together in a company called
09:12:35 Interleaf, which produced technical publishing software for
09:12:41 the '757, '767, and eventually the '787, and in that he
09:12:46 worked for me, and eventually I worked for him at VoiceBox.
09:12:52 Q. And you mentioned VoiceBox. How did it come about
09:12:56 that you guys founded that company?
09:12:58 A. So the idea of being able to talk to your computer,
09:13:04 talk to the internet, be able to do information retrieval I
09:13:08 found extraordinarily compelling. And so much so that when
09:13:13 we eventually began to form the company, it became clear
09:13:18 that, you know, early money is like yeast, and you have to

09:13:23 1 be able to, be able to address the marketplace, and in such
09:13:31 2 I invested the first $400,000, which I might add caused a
09:13:37 3 little distress in the marriage.
09:13:39 4 Q. For you was that a lot of money?
09:13:41 5 A. That was a significant amount of money.
09:13:43 6 Q. And when you were going to decide to found VoiceBox
09:13:48 7 with the Kennewick brothers, and put all that money in of
09:13:53 8 your own money, were you aware of any other voice-based
09:13:56 9 computer solutions that were out there?
09:13:59 10 A. There were a lot of people attempting there, there
09:14:02 11 was command and control systems, they would follow the menu
09:14:05 12 and say exactly the right words or nothing good would
09:14:08 13 happen.
09:14:09 14 Q. What do you mean by command and control?
09:14:11 15 A. So if I wanted to talk to a pharmacist, you had to be
09:14:17 16 able to say the exact words on the menu tree in order to get
09:14:24 17 to that particular destination. As opposed to just saying
09:14:28 18 one shot and be able to say I would like to talk to a
09:14:32 19 pharmacist about my -- about the drugs that are being
09:14:36 20 assigned to me.
09:14:37 21 Q. When you're saying talk to a pharmacist, is this
09:14:40 22 something on a computer or something like a phone?
09:14:44 23 A. I'm sorry, this would be a telephone tree, the
09:14:46 24 primary form that this took was a telephone, telephone based
09:14:55 25 systems.

09:14:56 Q. Okay. After you founded VoiceBox with the Kennewick
09:15:00 brothers, did you make any sort of a product or service
09:15:05 initially?
09:15:06 A. So, the whole notion became to develop this way that
09:15:11 one could talk to the internet, and talk to your computer.
09:15:14 And that's what the creation of Cybermind was all about.
09:15:18 Q. And what was your responsibility with Cybermind back
09:15:24 in the early days of VoiceBox?
09:15:26 A. I wore a lot of hats, but primarily what we're
09:15:30 talking about is, is the technology useful. There is lots
09:15:34 of people creating lots of technology in which you wouldn't
09:15:37 cross the street in order to use. And so the notion is that
09:15:40 okay, you're talking to the internet. Why are you talking
09:15:44 to the Internet? How is that useful? How does that bring
09:15:48 value and make it compelling to the customer to use.
09:15:54 Q. And were you responsible for any aspect of the
09:16:02 physical design of the Cybermind product?
09:16:04 A. Yeah. So I also wore the hat of creating the
09:16:07 physical head itself, the array microphones, the voice
09:16:11 cancelling system, the beam forming system that was inside
09:16:15 there because the whole system works if you have a robust
09:16:20 utterance that has a lot of detail in it. And that is what
09:16:23 the Bluetooth implementation brought for us.
09:16:27 Q. And could we -- I would like to show a video file,
09:16:35 543, to the witness on his screen, if we could do that?

09:16:39 A. Yes, I'm seeing it.
09:16:40 Q. Okay. Well before we play this. Mr. Freeman, are
09:16:46 you familiar with the video of PTX-543?
09:16:50 A. I am indeed, sir.
09:16:52 Q. Did you participate in the filming of that video?
09:16:56 A. I was what they call in the business a talking head,
09:16:59 yes, sir.
09:17:01 MR. SMITH: Your Honor, at this point plaintiff
09:17:02 would like to introduce into evidence PTX-543.
09:17:06 MR. HADDEN: No objection.
09:17:08 THE COURT: Thank you. It is admitted.
09:17:08 (PTX Exhibit No. 543 was admitted into
09:17:08 evidence.)
09:17:11 MR. SMITH: Mr. Smith, could you start playing
09:17:15 543, please?
09:17:18 (PTX-543 video played.)
09:17:39 MR. SMITH: Can you pause that?
09:17:40 So Mr. Freeman, who is that individual talking
09:17:44 there in the video?
09:17:45 A. That individual would be me, sir.
09:17:49 Q. And here in this -- in the video, what were you
09:17:53 talking to?
09:17:54 A. So that was the Cybermind, what we call the head, the
09:18:02 whole system together was called Cybermind because it
09:18:04 included the software, but that was the Bluetooth speaker

09:18:07 1 and mic.
09:18:09 2 Q. Do you recall approximately when this video of you
09:18:13 3 and the Cybermind head was taken?
09:18:14 4 A. Yeah, it was about 2006 that the NBC affiliate
09:18:20 5 approached us and said this would be useful to do.
09:18:23 6 MR. SMITH: Mr. Smith, if you could continue
09:18:26 7 playing this.
09:18:27 8 (Video played.)
09:19:21 9 BY MR. SMITH:
09:20:06 10 Q. So Mr. Freeman, I think at the end of that video, the
09:20:10 11 newscaster mentioned that there was going to be the hand
09:20:14 12 held thing out by Christmas. And what was -- was VoiceBox
09:20:18 13 able to do that?
09:20:24 14 A. We were not at the time.
09:20:26 15 Q. Why not?
09:20:26 16 A. There were compliments of issues, one was what's
09:20:30 17 called the bottom, the buildable materials to build that
09:20:36 18 thing, was expensive. Number two, there wasn't like a post
09:20:40 19 sales revenue model to be able to bring it down. And three,
09:20:45 20 Toyota, wonderful amazing company, approached us with a $5
09:20:49 21 million deal and said we want that technology in our cars,
09:20:52 22 so would you please defer your resources in order to do it.
09:20:56 23 Very difficult to turn down that kind of offer.
09:20:59 24 Q. What did VoiceBox do then?
09:21:00 25 A. What we did was, we said all right, the NLU and the

09:21:04 1 core technology we will continue to develop over time, and
09:21:07 2 we will make better but we will figure out how to shrink
09:21:11 3 this, in order to fit it into what's called the head unit of
09:21:14 4 an automobile which is the navigation unit, the infotainment
09:21:20 5 unit, the radio, heater, all that kind of stuff.
09:21:22 6 Q. At this point, when you had decided to work with
09:21:25 7 Toyota and put the Cybermind technology into the Toyota
09:21:29 8 cars, at that point, did you give up on Cybermind?
09:21:31 9 A. No. Cybermind was the, the image that we were
09:21:41 10 chasing, the notion of why this would be compelling, again,
09:21:44 11 why would you cross the street. We can make it this good in
09:21:48 12 an automobile, but we can make it this good in a living
09:21:51 13 room.
09:21:52 14 Q. Did you work on any technical improvements for
09:21:56 15 Cybermind in that time period?
09:21:59 16 A. The way software works, as you well know, is you just
09:22:03 17 don't put a formula into -- into software, instead you
09:22:08 18 iterate, and you iterate, and you get better, and you get
09:22:11 19 better, they're called drops, they're called -- they're
09:22:16 20 called release candidates, but yes, to answer your question,
09:22:20 21 completely.
09:22:20 22 Q. Did you work on any potential patents at that point?
09:22:25 23 A. We were working on ways, we believe very much in the
09:22:28 24 patent process, so we were working on ways of improving our
09:22:33 25 technology both from a recognition point of view, from an

09:22:38 1 audio path point of view, from a usefulness point of view so
09:22:43 2 that, we would patent that and be able to incorporate that
09:22:46 3 in our technology.
09:22:47 4 Q. And Mr. Smith, I believe we can -- this has been
09:22:51 5 admitted. Could you show JTX-001, which is the '681 patent
09:22:55 6 on the screen.
09:22:55 7 And Mr. Freeman, if you could take a look at
09:23:02 8 JTX-1, which is on the screen, and particularly in the upper
09:23:07 9 right-hand corner, do you see the last three digits of this
09:23:10 10 patent?
09:23:10 11 A. I do, indeed.
09:23:11 12 Q. Is this what we have been referring to as the '681
09:23:14 13 patent?
09:23:14 14 A. This is indeed.
09:23:15 15 Q. Are you one of the co-inventors of the '681 patent?
09:23:18 16 A. Along with looks like 5, 4 others, yes, correct.
09:23:23 17 Q. And what was the motivation for you and the other
09:23:28 18 co-inventors of the '681 patent to develop this?
09:23:31 19 A. Cooperative conversation. Voice is interesting,
09:23:35 20 right, it's like you talk, I talk, you talk, I talk. But if
09:23:41 21 there is no purpose, if there is no goal, if we're not
09:23:44 22 trying to actually get somewhere, it's kind of jaded idol in
09:23:50 23 the garden of knowledge. Why not be able to sharpen the
09:23:56 24 conversation, so that you're focused on actually achieving a
09:24:00 25 goal between the machine and the human. And that's a

09:24:03 1 primary idea here.
09:24:06 2 Q. Do you see here on the Item 22, there it says
09:24:11 3 October 16, 2006, do you see that?
09:24:13 4 A. I do.
09:24:14 5 Q. And is that consistent with your recollection of when
09:24:17 6 you filed for this patent?
09:24:18 7 A. It's consistent with the video that we were listening
09:24:25 8 to.
09:24:27 9 Q. And at this time, had you thought about the idea of
09:24:29 10 using long-term shared knowledge?
09:24:33 11 A. Absolutely. It turns out that when you and I have a
09:24:38 12 conversation, two things happen. One is we understand
09:24:41 13 what's happening in the conversation right now. But
09:24:45 14 secondly, and importantly, you have a basis of past
09:24:49 15 experience with me that makes a lot more sense. You say I
09:24:54 16 need to get to the freeway, what freeway are you talking
09:24:58 17 about, are you talking about I95, or Interstate 5, there is
09:25:02 18 a big difference between those two. And that by having a
09:25:06 19 long-term memory, we understand what freeway means in this
09:25:10 20 particular context.
09:25:11 21 Q. At the time that you filed the patent, were you aware
09:25:14 22 of any other voice system that had that kind of long-term
09:25:17 23 memory?
09:25:18 24 A. Not to my knowledge, sir.
09:25:20 25 Q. And what was your contribution for the '681 patent,

09:25:24 1 what role did you play?
09:25:26 2 A. So my role was to sit there and give it the litmus
09:25:31 3 test, would this be used, why would it be used. Are we just
09:25:34 4 having a technology chasing an idea, or is this actually
09:25:38 5 useful in the different use cases that we were talking
09:25:42 6 about, in the car, in the living room and different places.
09:25:45 7 Q. And Mr. Smith, if could go to column 11, start around
09:25:51 8 line 10, please.
09:25:55 9 And Mr. Freeman, I moved to a section of the
09:25:58 10 patent, the '681 patent. Do you see that on your screen
09:26:01 11 there?
09:26:02 12 A. I do.
09:26:03 13 Q. And in this section of the patent, it says, for
09:26:07 14 example, a user utterance of quote, what about traffic?
09:26:12 15 Close quote, may have a distinct meaning in various
09:26:18 16 contexts. Do you see that?
09:26:18 17 A. I do.
09:26:18 18 Q. Is this part of the patent that you worked on?
09:26:20 19 A. Yes, it is.
09:26:21 20 Q. And what was the idea behind this example about
09:26:25 21 traffic?
09:26:26 22 A. Well, if you were to look at me and say completely
09:26:29 23 cold, what about traffic, I would look at you and say I
09:26:33 24 think you're crazy. But to be able to establish a context
09:26:38 25 about what traffic means, are we talking about music, are we

09:26:41 1 talking about a movie, did this request come from a movie
09:26:45 2 agent, did this request come from a smart television, did
09:26:50 3 this request come from a navigation system, that helps
09:26:53 4 narrow down what traffic means. It doesn't mean you
09:26:57 5 wouldn't have a follow-up question because follow-up
09:26:59 6 questions are useful -- humans do this all the time. But
09:27:03 7 you can make certain assumptions based on where it's coming
09:27:07 8 from, and what the user is doing, what "traffic" means in
09:27:12 9 this situation.
09:27:12 10 Q. And would the long-term shared knowledge be helpful
09:27:16 11 for the system to figure out what was meant by traffic?
09:27:20 12 A. Oh, yes, very much so. How have you used this word
09:27:24 13 before, and how were you using this word before. And so you
09:27:29 14 have not only the history of those -- of the usage of the
09:27:33 15 word traffic, but also of the different elements within
09:27:35 16 traffic itself.
09:27:36 17 Q. And at that time when you invented the '681 patent,
09:27:42 18 what was it going to do for Cybermind in your view?
09:27:46 19 A. It was -- so the power of the Cybermind was to make
09:27:51 20 -- to take the friction out of a conversation with a human.
09:27:56 21 So it's -- a lot of this user interface stuff is, I don't
09:28:02 22 want to repeat everything I just said, I want you to be
09:28:06 23 smart enough to understand what traffic means. What is
09:28:10 24 traffic like on the way to the hospital? Well if it
09:28:14 25 understands me, it understands the hospital I go to, every,

09:28:18 1 couple times a month, and it understands the route I take
09:28:23 2 every time. Instead of saying gosh, I'm looking for traffic
09:28:26 3 information on the 520 going to Interstate 5 at the
09:28:29 4 interchange of Interstate 90, blah, blah, blah, sorry.
09:28:33 5 Q. No, thank you. I appreciate that.
09:28:36 6 Why don't we move to, Mr. Smith, if you could
09:28:39 7 pull up JTX-003.
09:28:44 8 And Mr. Freeman, do you see on the screen a
09:28:48 9 patent which ends in the numbers '176?
09:28:51 10 A. I do indeed, sir.
09:28:53 11 Q. And do you see in the left hand portion Item 75
09:28:59 12 there, you're listed as one of the inventors with Mike
09:29:02 13 Kennewick, do you see that?
09:29:03 14 A. I do.
09:29:04 15 Q. Did you work with Mike Kennewick on an invention that
09:29:09 16 turned into the '176 patent?
09:29:10 17 A. I did.
09:29:11 18 Q. And at a general level, what was the '176 about?
09:29:17 19 A. The '176 was to start thinking of those agents, the
09:29:22 20 agents, the music agent and a movie agent and the other
09:29:27 21 kinds of things that Cybermind could do, adding to that an
09:29:32 22 advertising aspect. So it would be like gosh, that is a
09:29:37 23 great tune, take some of you back to ring tone land, that I
09:29:41 24 would like a ring tone for that -- for Brown Eyed Girl and
09:29:46 25 be able to download that. The world has moved on of course

09:29:49 since that particular technology. But the aspect is
09:29:53 completely the same.
09:29:55 Q. And so and what was -- so you're saying as an
09:30:02 example, you could tell Cybermind to play a song, and then
09:30:06 Cybermind might suggest that you might want to buy the ring
09:30:10 tone?
09:30:11 A. Or I could ask it, right? So one of the key notions
09:30:15 here, and how advertising expanded in the course of the time
09:30:19 that we were dealing with this, was that instead of it being
09:30:23 a one-way flow from the source -- from the radio, if you
09:30:28 will, into the consumer, if the world switches around where
09:30:33 I could say gosh, could I get a ring tone, that would be
09:30:37 pretty cool, and that advertising becomes interactive.
09:30:40 Q. Why don't we now pull up JTX-005, please, Mr. Smith.
09:30:53 And Mr. Freeman, do you see that JTX-005, the upper
09:30:59 right-hand corner has the last three digits '097?
09:31:02 A. I do.
09:31:03 Q. And this one also lists you as a co-inventor with
09:31:07 Mike Kennewick, do you see that?
09:31:09 A. I do.
09:31:10 Q. What was the '097 patent about?
09:31:14 A. The '097 was an -- there is a term of art you folks
09:31:18 use, an expansion, no, it was an expansion of our patents
09:31:25 that we had in the earlier ones, so it's a continuation,
09:31:29 that was the word I was looking for, it was a continuation

09:31:33 1 of the previous patents that we had in this advertising
09:31:38 2 space.
09:31:39 3 Q. And when you say the previous patent we had, are you
09:31:41 4 talking about the '176 patent?
09:31:43 5 A. I am talking about the '176.
09:31:46 6 Q. Okay. And one question I had for you is with respect
09:31:51 7 to the '097 patent, the '176 patent, with advertising, did
09:31:57 8 the idea of pronouns ever play a role in what you were
09:32:00 9 thinking about at the time?
09:32:02 10 A. Yeah, if I could take you back to that part of the
09:32:06 11 conversation about removing the friction from the human to
09:32:09 12 machine interaction, part of that is as straightforward, but
09:32:15 13 technically not completely easy, of being able to substitute
09:32:20 14 a pronoun for a very long product name, a very long music
09:32:27 15 name, a very long artist, and which one doesn't want to risk
09:32:32 16 misrecognition, if you will. So buy it, get it for me, ship
09:32:37 17 it, and these are words that add a lot of value in the
09:32:44 18 human/machine interface.
09:32:46 19 Q. Okay. Now, I believe that-strike that.
09:32:52 20 What was the '176 and '097 patents intended to
09:32:59 21 do for the Cybermind product?
09:33:01 22 A. So, as I mentioned before, we had this challenge, and
09:33:04 23 that was that when you are a company, and you are trying to
09:33:09 24 produce something new and you're manufacturing it, you have
09:33:13 25 this bill of materials. And the bill of materials are all

09:33:16 1 of the elements that go into Bluetooth, and all the elements
09:33:19 2 that go into the head, and all of the elements that you have
09:33:22 3 to require in terms of your R & D, and we were looking at
09:33:26 4 what is called a BOM cost, building and materials cost, that
09:33:33 5 once we put it in the channels was going to make it
09:33:36 6 prohibitively expensive for your average consumer, for your
09:33:41 7 above average consumer to be able to purchase. So the
09:33:44 8 notion of being able to subsidize that by being able to have
09:33:48 9 an ad-free, or an ad-based system, being able to have people
09:33:52 10 be able to buy products in which we would get a percentage
09:33:58 11 of the opportunity, allowed us to have a post-sales business
09:34:05 12 model that would allow us to sell for less money than we
09:34:10 13 were currently thinking about doing.
09:34:12 14 Q. So would the version of Cybermind that had ads on it
09:34:16 15 be less expensive, or more expensive than the version of
09:34:20 16 Cybermind without the ads?
09:34:21 17 A. It would be less expensive. And consumers became
09:34:24 18 used to this kind of model for this time period that you can
09:34:29 19 buy yourself down to a non-ad environment.
09:34:32 20 Q. When did you leave VoiceBox?
09:34:34 21 A. 2010.
09:34:36 22 Q. And when you left VoiceBox in 2010, had you been able
09:34:42 23 to implement the '681 patent with the long-term memory that
09:34:51 24 we talked about, and the '176 and '097 patents with the
09:34:56 25 advertising in a physical product at that point when you

09:34:59 1 left?
09:35:00 2 A. Okay. I'll respond carefully, because you said in a
09:35:04 3 physical product to be released. The answer to that is no.
09:35:08 4 But the answer is, did we have point releases in engineering
09:35:12 5 that was getting us there, that software development is not
09:35:19 6 magic, sir, it is a lot of I wonder if that worked, let's
09:35:23 7 try this and you do it again and you do it again. And
09:35:26 8 that's -- that iterative process is the process by which
09:35:32 9 software is realized.
09:35:35 10 MR. SMITH: Thank you, Mr. Freeman. Appreciate
09:35:38 11 your time. I'll pass the witness.
09:35:40 12 THE WITNESS: Thank you.
09:35:41 13 THE COURT: Cross-exam.
09:35:43 14 CROSS-EXAMINATION
09:35:45 15 BY MR. HADDEN:
09:35:49 16 Q. Good morning, Mr. Freeman.
09:35:51 17 A. Good morning.
09:35:51 18 Q. Just to go back a little to your background, I
09:35:55 19 understand you went to the University of Washington. What
09:35:58 20 was your degree in?
09:35:59 21 A. So I had a liberal arts degree as an undergraduate,
09:36:03 22 and a masters of business administration from the Foster
09:36:06 23 School which is a prestige school at the University of
09:36:10 24 Washington.
09:36:10 25 Q. And just to be clear, so you quit VoiceBox in May of

09:36:15 1 2010; is that correct?
09:36:18 2 A. That is correct.
09:36:18 3 Q. And over the ten years you spent at VoiceBox before
09:36:23 4 you quit, it's correct that VoiceBox spent more money than
09:36:28 5 it received; correct?
09:36:30 6 A. I think it -- I'm not a subject matter expert in this
09:36:35 7 field, but I would say -- I'm not a subject matter expert in
09:36:40 8 that field.
09:36:40 9 Q. So you don't know?
09:36:42 10 A. It would -- I would answer yes to your question.
09:36:48 11 Q. Now, when you left VoiceBox in 2010, VoiceBox didn't
09:36:55 12 have -- had never enabled eCommerce transactions through
09:37:01 13 voice, correct?
09:37:04 14 A. Not true.
09:37:06 15 Q. Excuse me?
09:37:07 16 A. Not true.
09:37:09 17 Q. So can --
09:37:10 18 A. It depends what you mean, was it commercialized,
09:37:13 19 absolutely true, it was never commercialized.
09:37:16 20 Q. Do you have a copy of your deposition, it should be
09:37:19 21 that binder there, do you remember when you were deposed in
09:37:22 22 this case, sir? Were you able to find it there?
09:37:29 23 A. Deposition transcript? Sure.
09:37:32 24 Q. I would ask you to look at page 47, lines 9
09:37:37 25 through 18.

09:37:42 A. Page 47, lines 9 through?
09:37:46 Q. 18?
09:37:47 A. 18.
09:37:48 Q. Do you see that?
09:37:53 A. And that statement is accurate in that we never
09:37:58 enabled eCommerce transactions through voice, and I
09:38:03 interpreted your previous question is, were we actually
09:38:07 working on that and developing it.
09:38:08 Q. So just to be clear, my question was:
09:38:11 Did you ever actually do it, were you ever able
09:38:15 to actually do it before you left in 2010, and your answer
09:38:20 was, you could not, right?
09:38:22 A. Do it is a very uncomfortable word. In that were we
09:38:27 able to demonstrate it? Yes. Were we able to commercialize
09:38:32 it? No.
09:38:34 Q. Now, could we see JTX-003, Mr. Patterson?
09:38:47 Now, this is one of the voice ad patents that
09:38:50 you're an inventor on, right, sir?
09:38:53 A. Yes, sir.
09:38:53 Q. And you didn't make any technical contributions to
09:38:57 this patent, did you?
09:39:03 A. So the answer to your question is technical
09:39:10 contributions of -- well, did I develop an algorithm by
09:39:18 which engineers would at some point take and build a typical
09:39:22 solution, the answer is yes. Was I an engineer who built a

09:39:26 technical solution, the answer is no.
09:39:29 Q. So in February of 2007, when VoiceBox filed this
09:39:35 patent, did VoiceBox have a system that provided spoken
09:39:39 advertisements based on a user's utterance?
09:39:42 A. In 2007 when we filed for this patent?
09:39:45 Q. Correct.
09:39:46 A. So in 2007 when we filed for this patent, we were
09:39:50 developing concept software that was certainly not ready for
09:39:57 the customer facing, but was proving out of thought.
09:40:02 Q. Let's look at your deposition again, sir, this time
09:40:05 at page 174, lines 22 through 175, line 4?
09:40:18 A. Page 74?
09:40:19 Q. 174, sir.
09:40:24 A. 174. Okay, 174, what line, sir?
09:40:34 Q. Do you see the question begins on line 22, and your
09:40:38 answer ends on line 4. Do you see that? Can we put up that
09:40:45 testimony, Mr. Patterson?
09:40:58 And you testified in 2007, VoiceBox did not have
09:41:25 such an implementation; correct?
09:41:28 A. Certainly one that was not commercially available.
09:41:32 Q. You didn't say anything about commercially available
09:41:34 in your sworn testimony, did you sir?
09:41:38 A. I did not. Implementation is a vague word.
09:41:42 Q. I'm sorry, I didn't hear you?
09:41:44 A. I was just saying that implementation is a very broad

09:41:48 1 word.
09:41:48 2 Q. Let's look at the patent again, please,
09:41:51 3 Mr. Patterson, JTX-003. And let's look at Claim 40.
09:42:05 4 Did you invent a conversational language
09:42:10 5 processor?
09:42:12 6 A. Did I invent a conversational language processor?
09:42:15 7 Q. Yes, sir.
09:42:15 8 A. I did not.
09:42:16 9 Q. Did you invent an adaptive misrecognition engine,
09:42:21 10 sir?
09:42:29 11 A. So did I outline what the algorithm would be to make
09:42:32 12 it useful. The answer is yes. Did I write the code, the
09:42:35 13 answer is no.
09:42:37 14 Q. Can you describe for me what the algorithm is that
09:42:42 15 you developed for an adaptive misrecognition engine?
09:42:45 16 A. So it is about where the utterance comes from and
09:42:49 17 what the purpose of the utterance is. For instance, if
09:42:51 18 you're in a distributed environment, and you have a
09:42:55 19 navigation system and that you're picking up the words about
09:43:00 20 traffic from a navigation system, that's telling you in an
09:43:06 21 enormous amount of information about what the word traffic
09:43:08 22 means in that particular context, poison IV and Poison Ivy
09:43:16 23 are phenetically extremely close to each other, but one
09:43:18 24 comes from medical, and one comes from a gardening app, then
09:43:21 25 you have a clue as to what that information means. So even

09:43:25 1 though it might be misrecognized originally, that because of
09:43:30 2 the source of the information you're able to say, that's
09:43:35 3 probably about a medical situation, and not about a
09:43:39 4 gardening situation.
09:43:40 5 Q. And that's the algorithm that is in this patent?
09:43:48 6 A. So that -- so the algorithm in this patent is about
09:43:51 7 what it is that the engineers are going to be sent off in
09:43:57 8 order to build, in order to make the rest of the situation
09:44:00 9 compelling.
09:44:01 10 Q. And what you told the engineers to build is what you
09:44:06 11 just described to the jury?
09:44:07 12 A. Say again?
09:44:08 13 Q. What you told the engineers to build is what you just
09:44:12 14 described to the jury?
09:44:13 15 A. Yes, sir.
09:44:13 16 Q. Thank you.
09:44:14 17 Let's look at DTX -- I'm sorry, JTX-005, please,
09:44:21 18 Mr. Patterson. And this is the '097 patent that you were
09:44:27 19 asked about by your counsel, do you recall that, sir?
09:44:30 20 A. Yes.
09:44:30 21 Q. And could we go to Claim 23, please, Mr. Patterson?
09:44:35 22 And this is the claim that talks about pronouns.
09:44:45 23 Do you remember being asked about pronouns by your counsel,
09:44:49 24 sir?
09:44:49 25 A. I do.

09:44:49 Q. Did you invent a way for a computer to understand the
09:44:52 meaning of a pronoun?
09:44:56 A. So what I did was to say we have a requirement here
09:45:01 for pronouns to be able to take -- to be able to substitute
09:45:07 for the actual nouns in this situation. How you do it in
09:45:11 engineering, I don't have a dog in this fight, I don't care,
09:45:14 but you need to be able to do this. Engineering in
09:45:16 response, of course, will come up with 5, 10, 15, 20
09:45:21 different ways of doing this, and we'll evaluate them
09:45:24 accordingly.
09:45:25 Q. Just to be clear, you didn't invent a way for a
09:45:27 computer to understand the meaning of a pronoun, correct?
09:45:31 A. I said, sure, I agree.
09:45:37 Q. Let's look at -- there is a document in your binder,
09:45:41 sir, DTX-102, it's in the binder that we gave you.
09:45:47 A. DTX-102?
09:45:50 Q. Yes, sir.
09:45:51 A. Yes.
09:45:53 Q. Okay. Do you recognize that document, sir?
09:45:55 A. I do.
09:45:58 MR. HADDEN: Move DTX-102 into evidence, Your
09:46:02 Honor.
09:46:02 MR. SMITH: No objection.
09:46:04 THE COURT: Thank you. It is admitted.
09:46:04 (DTX Exhibit No. 102 was admitted into

09:46:08 1 evidence.)
09:46:08 2 BY MR. HADDEN:
09:46:09 3 Q. So this is a paper that you wrote with Mr. Baldwin,
09:46:14 4 another inventor on the '681 patent. Is that correct?
09:46:17 5 A. That is correct.
09:46:18 6 Q. And you wrote this paper with Mr. Baldwin, it's a
09:46:25 7 marketing material for VoiceBox, correct?
09:46:27 8 A. It was a vision piece, correct.
09:46:34 9 Q. If we look at page 6 of this paper. Can you move to
09:46:41 10 that Mr. Patterson?
09:46:42 11 At the top is a heading cooperative
09:46:45 12 conversation. Do you see that, sir?
09:46:46 13 A. I do.
09:46:48 14 Q. And then it goes on and it says what is a cooperative
09:46:53 15 conversation, and you wrote in this paper that the
09:46:56 16 cooperative principle was first described by Paul Grice in
09:47:01 17 1975. Do you see that?
09:47:04 18 A. I do.
09:47:04 19 Q. And the cooperative principle is what you were
09:47:07 20 talking about with your counsel, where people need to work
09:47:12 21 together to get to some goal in their conversations,
09:47:14 22 correct?
09:47:15 23 A. Correct.
09:47:15 24 Q. And that idea that you're talking about with your
09:47:19 25 counsel, you got from this paper from Mr. Grice, right?

09:47:22 A. Correct.

09:47:23 Q. And if we look lower down, if we could go down to --
09:47:30 so if we highlight the paragraph that begins a cooperative
09:47:33 conversation, and the material just below it, it would be
09:47:36 the second paragraph, Mr. Patterson.

09:47:40 There is --

09:47:41 MR. SMITH: I object, I think referencing this
09:47:45 is Mr. Patterson, if we could clarify Mr. Freeman for the
09:47:50 record.

09:47:50 MR. HADDEN: I was talking to Mr. Patterson who
09:47:52 is --

09:47:53 THE COURT: Don't talk to each other, talk to
09:47:55 me.

09:47:55 MR. HADDEN: Sorry, Your Honor. I did not mean
09:47:58 to misstate your name. I was talking to Mr. Patterson, sir.

09:48:03 THE WITNESS: No offense taken.

09:48:05 BY MR. HADDEN:

09:48:05 Q. If we look at this from this paper you wrote with
09:48:08 Mr. Baldwin, you're laying out here the principle of a
09:48:12 cooperative conversation that Mr. Grice described in his
09:48:16 paper in 1975, right?

09:48:17 A. Correct.

09:48:18 Q. And that included this notion of the parties in a
09:48:25 conversation using shared knowledge, right?

09:48:28 A. Correct again.

09:48:29 Q. And that included both long-term shared knowledge and
09:48:33 short-term shared knowledge, correct?
09:48:34 A. Correct.
09:48:35 Q. And so the notion of -- the importance of using
09:48:45 long-term and short-term shared knowledge to understand what
09:48:49 is being said in a conversation is something that you got
09:48:52 from Mr. Grice; right?
09:48:54 A. This was part of the intellectual averment, and this
09:49:03 was a very useful way of sharpening our vocabulary around
09:49:09 this, the answer is yes, sir.
09:49:10 Q. I just want to be clear, this was Mr. Grice's idea,
09:49:14 that wasn't your idea?
09:49:15 A. Correct.
09:49:15 Q. And then it talks about hypothesis building, intent
09:49:23 determination, that also came from Mr. Grice's paper?
09:49:26 A. Correct.
09:49:26 Q. And adaptive response, that also came from
09:49:29 Mr. Grice's paper?
09:49:31 A. Correct.
09:49:32 Q. And this marketing paper where you describe these
09:49:37 principles that Mr. Grice had laid out, this is essentially
09:49:40 the '681 patent in ordinary language, right?
09:49:43 A. So this is '681 but '681 takes this to a machine, as
09:49:47 opposed to human to human.
09:49:49 Q. So what you did in the '681 patent is you said let's

09:49:52 take all these ideas from Mr. Grice about a cooperative
09:49:57 conversation, and let's propose that we should do the same
09:50:00 thing when we talk to computers; right?
09:50:05 A. Yes, sir. And he's attributed in the '681.
09:50:11 Q. Actually when you went -- here in this paper you cite
09:50:14 to Mr. Grice, if you look at the references in the back, can
09:50:18 we look at page 7, please, Mr. Patterson, and blow up that
09:50:22 first reference there.
09:50:24 You credit Mr. Grice's paper here in this
09:50:31 marking material you wrote, correct?
09:50:33 A. Yes.
09:50:34 Q. But when you filed the '681 patent, you didn't say
09:50:40 anything about Mr. Grice, did you?
09:50:41 A. We attributed the view paper, and the view paper
09:50:44 attributes this.
09:50:45 Q. There was nothing in the patent itself that says that
09:50:47 these ideas about using long-term shared knowledge and
09:50:51 short-term shared knowledge and intent determination and
09:50:55 adapting a response, that that all came from Mr. Grice,
09:50:59 that's not anywhere in your patent, is it?
09:51:02 A. Well, it is attributed to the -- the paper is
09:51:06 attributed, and the paper attributes that.
09:51:10 Q. So you're saying -- just to be clear, Mr. Grice's
09:51:17 article did not provide --
09:51:21 A. Say again?

09:51:22 Q. Mr. Grice's article that you got all these ideas
09:51:25 from, you did not provide that to the Patent Office during
09:51:30 the prosecution of the '681 patent, correct?
09:51:35 A. It was included by reference.
09:51:36 Q. Well it wasn't included by reference, what was
09:51:39 included by reference was your view paper?
09:51:41 A. Correct.
09:51:41 Q. There is no reference in the '681 patent to
09:51:45 Mr. Grice, or the paper you copied those ideas from,
09:51:50 correct?
09:51:52 A. There was no attempt to deny Mr. Grice's
09:51:56 contribution.
09:51:56 Q. I understand that, sir, and I'm not saying you were
09:52:00 attempting to deny his contribution, I just want to state
09:52:03 the facts. The fact is that the patent that you filed, the
09:52:06 '681 patent, doesn't say anywhere that these ideas came from
09:52:12 Mr. Grice, correct?
09:52:14 A. I disagree.
09:52:19 Q. Well, do you want to look at the patent and show me
09:52:22 where Mr. Grice's name appears?
09:52:26 A. So it is in the reference to the view paper, and the
09:52:29 view paper attributes him.
09:52:31 Q. So you're saying if someone found the view paper and
09:52:35 read it, they would figure out that actually these ideas
09:52:38 came from Mr. Grice, but you didn't tell the Patent Office

09:52:41 1 that, correct?
09:52:43 2 A. We were talking about implementing a lot of these
09:52:47 3 ideas in a machine, and that as such, that was less
09:52:52 4 important.
09:52:57 5 MR. HADDEN: I have no further questions. Thank
09:52:58 6 you for your time.
09:53:00 7 THE WITNESS: Thank you.
09:53:00 8 THE COURT: All right. Thank you.
09:53:01 9 Redirect.
09:53:04 10 MR. SMITH: No redirect. Thank you.
09:53:06 11 THE COURT: All right. Thank you, sir. You are
09:53:08 12 excused.
09:53:09 13 THE WITNESS: Thank you.
09:53:10 14 THE COURT: What's next?
09:53:13 15 MR. YOON: Your Honor, at this time, we are
09:53:15 16 going to call by deposition, Kelly Vanee, who was the former
09:53:23 17 Vice-President of Learning and Engineering at Amazon.
09:53:31 18 (Videotape deposition of Kelly Vanee:)
09:53:44 19 Q. -- PowerPoint presentation with a Bates number ending
09:53:46 20 in 44736?
09:53:49 21 A. Oh, there we go. Yep, it's open now.
09:53:52 22 Q. Well, do you recognize Exhibit 226 to be a slide deck
09:53:57 23 that you presented at Amazon?
09:53:58 24 A. Yes, this is a slide deck that I helped author and
09:54:02 25 present.

09:54:04 1 Q. And do you recall when you presented this slide deck?
09:54:09 2 A. This was presented numerous times throughout the --
09:54:13 3 the history of Alexa. And it was -- the presentation was
09:54:17 4 recorded and watched by a large number of individuals, as
09:54:23 5 well.
09:54:34 6 (End of videotape.)
09:54:36 7 MR. YOON: Your Honor, at this time, we move
09:54:39 8 into evidence PTX-226.
09:54:42 9 MR. HADDEN: No objection.
09:54:46 10 THE COURT: All right. It's admitted.
09:54:48 11 (PTX Exhibit No. 226 was admitted into
09:54:49 12 evidence.)
09:54:49 13 MR. YOON: And Your Honor, at this time,
09:54:53 14 VoiceBox will call its expert, Nathaniel Polish.
09:55:00 15 THE COURT: Okay.
09:55:08 16 COURTROOM DEPUTY: Please raise your right hand.
09:55:14 17 Please state and spell your full name for the record.
09:55:44 18 THE WITNESS: Nathaniel Polish,
09:55:46 19 N-A-T-H-A-N-I-E-L, P-O-L-I-S-H.
09:55:54 20 NATHANIEL POLISH, having been duly sworn, was
09:55:59 21 examined and testified as follows:
09:56:02 22 DIRECT EXAMINATION
09:56:03 23 BY MR. YOON:
09:56:09 24 Q. Good morning, Dr. Polish.
09:56:10 25 A. Good morning.

09:56:11 Q. Could you please introduce yourself to the jury?
09:56:14 A. Sure. My name is Nathaniel Polish. I'm from New
09:56:20 York City, and I'm here to testify on behalf of VB Assets.
09:56:26 Q. And if we could have PDX-32 on the screen.
09:56:32 Dr. Polish, could you briefly describe your educational
09:56:35 background?
09:56:36 A. Sure. I have a bachelor's in physics from Columbia,
09:56:40 and I have a Ph.D. in computer science from Columbia. My
09:56:44 Ph.D. work was in computer speech.
09:56:47 Q. And Dr. Polish, could you briefly describe your
09:56:50 professional experience?
09:56:51 A. Sure. Outside of -- as soon as I got done with high
09:56:55 school, I started with a consulting firm called Daedalus
09:57:02 Technology Group, it's a research and development outsource
09:57:04 for other companies, companies coming to me and ask me to
09:57:06 help them create new technologies, and we and a team of
09:57:11 engineers produce prototypes and early stage concepts for
09:57:15 the clients.
09:57:16 Q. And Dr. Polish, could you briefly describe your
09:57:18 professional experience?
09:57:20 A. Yeah. So in addition to Daedalus, I from time to
09:57:25 time start companies that commercialize technologies that I
09:57:27 invent. I have probably done that seven or eight times over
09:57:32 the last forty years. Two of the companies relevant to
09:57:36 here, one of them is SimplicITY, which developed a

09:57:40 1 conversational speech access to databases, particularly from
09:57:45 2 music and eCommerce in the late '90s. And also I recall
09:57:50 3 that was doing linking of note taking applications to things
09:57:55 4 like audio and video recordings.
09:57:58 5 Q. And Dr. Polish, are you a named inventor on any
09:58:02 6 patents?
09:58:02 7 A. Yes. My work has resulted in I think, at this time,
09:58:05 8 10 patents, there are a few others pending. And at least
09:58:09 9 one of them involves conversational speech systems.
09:58:12 10 MR. YOON: Your Honor, at this time we would
09:58:14 11 like to offer Dr. Polish as an expert on interactive voice
09:58:17 12 and speech technology.
09:58:18 13 MR. HADDEN: No objection.
09:58:19 14 THE COURT: All right. Thank you. He'll be
09:58:20 15 recognized as such. There is an echo going on. Maybe you
09:58:25 16 can turn down the volume or something.
09:58:30 17 THE WITNESS: How is it now? Is it better?
09:58:32 18 MR. YOON: Is that better, Your Honor?
09:58:38 19 BY MR. YOON:
09:58:39 20 Q. If we could go to PDX slide 3 on the screen.
09:58:44 21 Dr. Polish, in general, what areas of technology do the
09:58:46 22 patents in this case relate to?
09:58:48 23 A. So the patents relate to systems that are conversant
09:58:55 24 systems where a person is talking to a computer, and the
09:59:00 25 computer is talking back, and there is some conversation

09:59:03 1 that goes on for some particular purpose, in some cases,
09:59:07 2 involving eCommerce.
09:59:10 3 Q. And if we could have PDX slide four. Dr. Polish,
09:59:16 4 what have you been asked to do in this case?
09:59:23 5 A. So I was asked to look at the patents, look at the
09:59:25 6 accused products, and determine whether I believe that there
09:59:28 7 was infringement of those -- of those patents by those
09:59:32 8 products.
09:59:32 9 Q. And what did you do to prepare yourself to offer such
09:59:37 10 an opinion?
09:59:37 11 A. I looked at a lot of documents. I reviewed the
09:59:40 12 patents. I looked at the prosecution histories.
09:59:43 13 Prosecution histories are the interaction between the
09:59:46 14 patentee and the Patent Office as part of getting the
09:59:49 15 patent. I looked at the Court's claim construction order.
09:59:52 16 I looked at PTAB rulings, I used Alexa extensively. I
09:59:59 17 reviewed many technical documents provided to me in the case
10:00:04 18 relating to Alexa, and then I confirmed my analysis with
10:00:08 19 source code. And I ultimately compared the claims of the
10:00:11 20 patents to the Alexa system.
10:00:14 21 Q. And if I could have PDX slide five.
10:00:18 22 Dr. Polish, as part of your analysis, did you
10:00:22 23 review the Court's claim construction of certain claim terms
10:00:28 24 with respect to the four patents?
10:00:29 25 A. I did.

10:00:30 Q. And for example, do you see there is a reference
10:00:33 there to speech recognition engine/speech recognition?
10:00:38 A. Yes.
10:00:38 Q. And for illustration, do you see that it's software
10:00:42 or hardware that recognizes the words or phrases in the
10:00:46 natural language utterance?
10:00:47 A. Yes.
10:00:47 Q. And in your analysis, did you apply the Court's claim
10:00:52 construction to your infringement analysis?
10:00:53 A. I did.
10:00:56 Q. Now, Dr. Polish, if a claim term was not expressly
10:01:01 interpreted by the Court, did you apply the ordinary meaning
10:01:06 of claim terms to a person of ordinary skill in the field?
10:01:09 A. Yes, I did.
10:01:10 Q. If we could have slide six on the screen.
10:01:13 And Dr. Polish, in your opinion, what's the
10:01:17 level of ordinary skill in the art with respect to the four
10:01:20 patents?
10:01:20 A. So, a person of ordinary skill in my view, is
10:01:25 somebody who has a bachelor's degree in computer science or
10:01:28 related field, plus two years of practical experience in
10:01:34 various kinds of speech systems. You could trade some of
10:01:39 that off with more education or more experience, but this is
10:01:42 approximately the level of skill that you would need to
10:01:45 understand the patents.

10:01:46 Q. And is it your understanding that the patents were
10:01:49 written for people of such skill?
10:01:51 A. Yes.
10:01:51 Q. And that would be the background of someone reading
10:01:54 the patent?
10:01:55 A. Yes, that's right.
10:01:57 Q. And if we could now have on the screen JTX-1, the
10:02:05 '681 patent.
10:02:09 And Dr. Polish, what conclusions have you
10:02:14 reached regarding Claim 13 of the '681 patent?
10:02:19 A. My conclusion is that the Alexa products infringe
10:02:22 Claim 13 of the '681 patent.
10:02:25 Q. And if we could now have on the screen the '176 and
10:02:31 '097 patents.
10:02:34 And Dr. Polish, with respect to the asserted
10:02:39 claims to each of the patents in this case, what are the
10:02:42 opinions that you have reached with respect to whether or
10:02:45 not Alexa infringes those patents?
10:02:46 A. So the asserted claims of each of those patents are
10:02:49 infringed by the Alexa systems.
10:02:52 Q. In particular Alexa Shopping?
10:02:54 A. Yes.
10:02:59 Q. And let's go specifically, if we could have just the
10:03:01 '176 patent on the screen. And what conclusions have you
10:03:05 reached with respect to Claim 40 of the '176 patent?

10:03:09 A. Claim 40 is infringed by the Alexa systems.
10:03:17 Q. And looking at -- would that be Alexa Shopping?
10:03:23 A. Yes.
10:03:23 Q. Looking at the '097 patent, what conclusions have you
10:03:27 reached with regard to Claim 23 of the '097 patent?
10:03:31 A. That Claim 23 is infringed by the Alexa Shopping
10:03:35 systems.
10:03:36 Q. And if we could have the '703 patent on the screen.
10:03:41 And sir, what conclusions have you reached regarding Claim
10:03:44 25 of the '703 patent?
10:03:46 A. Claim 25 is infringed by the Alexa Shopping system.
10:03:50 Q. And Dr. Polish, we're going to walk through element
10:03:53 by element, your infringement analysis for those patents
10:03:56 shortly, but let's talk a little bit about the background
10:03:59 for each patent.
10:04:00 If we could have on the screen the '681 patent.
10:04:10 If we could turn to column one, lines 7 to 8 of the '681
10:04:15 patent.
10:04:16 And -- I'm sorry, column one lines 5 to 9, I
10:04:22 apologize.
10:04:22 What was the field of invention for the '681
10:04:25 patent?
10:04:25 A. Well the invention relates to a cooperative
10:04:30 conversation model for human to machine voice.
10:04:34 Q. And the '681 patent you understand relates to a

10:04:38 1 machine voice user interface, correct?
10:04:41 2 A. That's right.
10:04:44 3 Q. And what is a cooperative conversation model, sir?
10:04:47 4 A. A cooperative conversation model is some kind of
10:04:51 5 model that represents an expected dialogue between a user
10:04:55 6 and a computer. So it's something that represents what
10:05:00 7 interaction would be expected between the human and the
10:05:03 8 computer for some particular purpose.
10:05:06 9 Q. And if we could now have Figure 1 of the '681 patent.
10:05:11 10 And column six lines 64 and 65. Dr. Polish, can you
10:05:20 11 describe the system disclosed in the '681 patent?
10:05:23 12 A. Sure, so this block diagram, a very high level
10:05:31 13 diagram of how this all works. In the upper left is what's
10:05:35 14 labeled as input, that's the human speech that comes in.
10:05:38 15 The next box is labeled ASR, Automatic Speech Recognition
10:05:43 16 that converts it, converts the speech into some form of
10:05:47 17 text. Then there is a Conversational Language Processor
10:05:51 18 which is, also can be an NLU or Natural Language
10:05:55 19 Understanding processor that seeks to understand what the
10:05:58 20 text is talking about. And that, in combination with
10:06:02 21 applications and databases, tries to understand what the
10:06:08 22 user's intentions or goals are, and ultimately constructs an
10:06:14 23 output which goes out on the left as generated speech.
10:06:17 24 Q. If we could now have Figure 3 of the '681 patent, and
10:06:24 25 column seven, lines 1 to 3.

10:06:29 Dr. Polish, what's shown in Figure 3?
10:06:32 A. So this is a little bit of a drill down on how the
10:06:40 models work for a cooperative conversations. It's showing
10:06:45 the input is coming in from the bottom, that's some kind of
10:06:49 recognized speech, and then that is combined with shared
10:06:55 knowledge, that is to say knowledge from prior utterances or
10:06:59 prior conversations, it's combined with that, it's combined
10:07:04 with some kind of hypothesis builder where it's making
10:07:10 guesses if you will, intelligent guesses as to what the user
10:07:13 means. And then based on that, it comes up with some kind
10:07:16 of a response.
10:07:20 Q. And if we could now have column four, lines 45 to 50
10:07:24 and column five, lines 25 to 30 of the '681 patent on the
10:07:30 screen.
10:07:31 Dr. Polish, how is shared knowledge used by the
10:07:36 system in the '681 patent?
10:07:38 A. So the '681 patent talks about several different
10:07:42 kinds of shared knowledge. They referred to long-term
10:07:45 shared knowledge, and short-term shared knowledge.
10:07:48 Generally, the knowledge is used, is combined with new
10:07:54 utterances that comes in to try to figure out what the
10:07:58 intention of the user is.
10:07:59 Q. And if we could focus on, for example, lines 53 to
10:08:05 55, does the patent describe short-term shared knowledge?
10:08:09 A. Yes, it does.

10:08:10 Q. And what does it say?
10:08:12 A. So short-term shared knowledge is accumulated during
10:08:16 a single conversation. So if you're engaged in a
10:08:20 conversation with a computer that's a multiple interaction
10:08:24 with the computer, from interaction to interaction, that's
10:08:29 short-term knowledge.
10:08:30 Q. And if we look at lines 63 to 65, do you see there is
10:08:35 a reference there to build long-term shared knowledge?
10:08:38 A. Yes.
10:08:39 Q. How does the '681 describe how long-term shared
10:08:43 knowledge can be built?
10:08:44 A. So long-term shared knowledge, is knowledge about the
10:08:47 user that extends over multiple conversations, so it would
10:08:54 be things about the user that presumably don't change, it
10:08:57 might involve where their device is, it might involve the
10:09:01 acoustics of their room, it might involve their profile in
10:09:05 some other way that would help with the conversation.
10:09:08 Q. And if we could now look at Figure 3 and have column
10:09:13 13, lines 31 to 36 on the screen.
10:09:19 And does the '681 patent describe how long-term
10:09:25 and short-term shared knowledge can be used?
10:09:28 A. Yes, so it talks about that for long-term and
10:09:34 short-term shared knowledge, it accumulates over a single
10:09:38 conversation to help build user profiles, other information
10:09:42 about the user.

10:09:43 Q. And what is the benefit of using long-term and
10:09:50 short-term shared knowledge?
10:09:51 A. There is a lot of benefits to it. It generally
10:09:55 streamlines the conversations. It means that the user
10:09:58 doesn't have to keep repeating themselves because you have
10:10:01 already set a context of what you're looking for, so you
10:10:04 don't have to keep saying the context over and over again.
10:10:07 It streamlines the conversation, it makes errors less likely
10:10:11 because you're interpreting things from the point of view of
10:10:15 what's come before. With long-term knowledge, you might
10:10:18 have information about whether someone is an expert or a
10:10:22 novice, or whether they're a native English speaker or
10:10:25 something of that nature.
10:10:26 Q. Would the '681 patent reduce the number of turns or
10:10:30 interactions a user would have to have in order to get
10:10:35 something done?
10:10:35 A. Yes, it would streamline a number of things including
10:10:38 the number of times you would have to go around.
10:10:40 Q. And if we could now discuss the '176 and '097 patent
10:10:46 together. And for convenience sake, Dr. Polish, you
10:10:49 understand that the '176 patent and '097 patent share the
10:10:54 same specification?
10:10:55 A. Yes, I do.
10:10:55 Q. And so I'll just for convenience, direct you to where
10:10:59 it is in the '176 patent, but I'll be asking you about both.

10:11:03 1 A. Okay.
10:11:04 2 Q. If I could go to column one, lines 8 to 10 of the
10:11:08 3 '176 patent. What is the field of invention for the '176
10:11:13 4 and '097 patents?
10:11:15 5 A. So the field of the invention of these patents is
10:11:18 6 selecting and presenting advertisements based on natural
10:11:22 7 language processing of voice-based input.
10:11:26 8 Q. And if we could have Figure 1, column two, lines 50
10:11:30 9 to 53 on the screen. And looking at Figure 1 of the '176
10:11:37 10 and '097 patent, can you describe the system disclosed in
10:11:41 11 those patents?
10:11:43 12 A. Sure, similar to what we saw before, the input, which
10:11:47 13 is human speech, is on the upper left connecting to an ASR,
10:11:55 14 which connects to a Conversational Language Processor or
10:12:02 15 NLU. This then connects to applications including an
10:12:06 16 advertising application and connects to databases. From
10:12:12 17 that, output is generated and sent to the user in the form
10:12:17 18 of vocalized speech.
10:12:23 19 Q. If we could also have Figure 1, column 4, lines 17 to
10:12:29 20 34, if you could move it down a little it. Column 17 is cut
10:13:11 21 in half.
10:13:12 22 Dr. Polish, does the '176 and '097 patent
10:13:16 23 explain how the Conversational Language Processor operates?
10:13:20 24 A. Yes, it describes it here, it talks about long-term
10:13:23 25 and short-term knowledge. And using that to understand the

10:13:29 1 current utterance.
10:13:31 2 Q. And looking at line 25, do you see the statement "for
10:13:36 3 example, Conversational Language Processor 120 may use
10:13:41 4 context determination model 130 to establish context for a
10:13:48 5 current utterance by having domain agents." Do you see
10:13:52 6 that?
10:13:52 7 A. Yes.
10:13:52 8 Q. What is a domain agent?
10:13:55 9 A. A domain agent is some piece of code that is
10:13:59 10 associated with a particular topic area. So the domain
10:14:02 11 might be shopping, it might be weather. It might be travel.
10:14:07 12 But it's some topic area where the interactions with the
10:14:12 13 users make sense to group them together within a single
10:14:16 14 module.
10:14:16 15 Q. Do you see there is a reference there to scoring
10:14:20 16 possible interpretations and selecting a highest scoring
10:14:23 17 interpretation?
10:14:24 18 A. Yes.
10:14:25 19 Q. What do you understand that to be part of?
10:14:29 20 A. So these systems are not -- they're not ever
10:14:34 21 completely sure about themselves, I mean, sometimes they may
10:14:37 22 be, but usually they operate with competing hypotheses. So
10:14:42 23 someone is talking to it, and it thinks they're probably
10:14:45 24 talking about the weather, but maybe they're talking about
10:14:48 25 travel. So it will often maintain a collection of competing

10:14:55 1 hypotheses with different levels of confidence. And this is
10:15:00 2 talking about generating those different hypotheses and
10:15:06 3 scoring them.
10:15:06 4 Q. Briefly talking about Alexa for a moment, we'll get
10:15:09 5 into more detail later, does Alexa have domains?
10:15:12 6 A. Yes, it does.
10:15:13 7 Q. And does Alexa score possible interpretations and
10:15:17 8 select the highest?
10:15:18 9 A. Yes, it does.
10:15:19 10 Q. Now, if we could now go to Figure 1 and column four,
10:15:25 11 lines 55 -- 45 to 55. Do the '176 and '097 patents explain
10:15:36 12 how the Conversational Language Processor can adapt to
10:15:40 13 incorrect interpretation?
10:15:42 14 A. Yes. It describes how it can support, essentially
10:15:48 15 changing its mind about a given hypothesis, deciding that
10:15:54 16 what it thought was going on was actually wrong and it
10:15:57 17 should do something else.
10:15:58 18 Q. If we could go to Figure 2, and column two, lines 53
10:16:04 19 to 54. Figure 2.
10:16:15 20 Dr. Polish, what's shown in Figure 2 of the '176
10:16:18 21 and '097 patents?
10:16:20 22 A. So, this is talking about an advertising server, so
10:16:27 23 this is how an advertiser might create advertisements that
10:16:34 24 would be used within the system.
10:16:38 25 Q. And if we could now have Figure 2, column two, lines

10:16:45 7 to 12 on the screen. Does the '176 patent provide
10:16:50 examples of the types of ads that may be used, or how ads
10:16:55 may be selected?
10:16:58 A. Yes, it talks about ways in which ads could be
10:17:02 selected based upon any number of things, including user
10:17:07 profiles, demographics, so that things would be served to a
10:17:11 user that would be appropriate to their needs or might be
10:17:15 more likely to be purchased than some other offering.
10:17:23 Q. And if we could have column two, lines 21 to 25 on
10:17:30 the screen.
10:17:58 Does the '176 and '097 patents provide examples
10:18:15 of the types of advertisements that can be used by the
10:18:20 system?
10:18:20 A. Yes. They talk, for example, about the
10:18:25 advertisements could include sponsored messages, purchase
10:18:29 opportunities, other marketing communications. It would be
10:18:34 -- it could be a pretty broad range of things, but
10:18:38 importantly, it could be just any kind of message that is
10:18:42 offering a purchase.
10:18:43 Q. If we can now go to Figure 3, column two, lines 55 to
10:18:51 57.
10:18:52 Dr. Polish, what's shown in Figure 3 of the '176
10:18:58 and '097 patents?
10:19:00 A. So this is a flow chart that is showing roughly what
10:19:08 happens when voice input is received where voice input comes

10:19:14 1 in at the top, it's identified and then -- and then an ad is
10:19:21 2 chosen, and it's then presented to the user in the form of
10:19:28 3 speech. And then -- and then the user can interact with
10:19:33 4 that, with that advertisement, presumably to purchase
10:19:37 5 something or ask a follow-up question or something like
10:19:40 6 that. And then action is taken based on that. And it can
10:19:44 7 loop where it continues to present output and get
10:19:50 8 interaction from the user, and ultimately the behavior of
10:19:54 9 the user is being tracked.
10:19:57 10 Q. And given the interface between the user and the
10:20:00 11 system, do the '176 and '097 patents streamline the
10:20:10 12 communications between the user and the system such as the
10:20:13 13 number of turns or interactions?
10:20:16 14 A. Yes, as you go through these loops, things become
10:20:19 15 more focused as you're interacting with the system in
10:20:24 16 response to this ad. And more information is being
10:20:29 17 obtained. So this approach to sort of focusing the
10:20:33 18 interaction reduces the amount of cycles through the system
10:20:39 19 you would have to do.
10:20:40 20 Q. If we could go to JTX-2, the '703 patent. And if we
10:20:45 21 could now have column one, lines 11 to 17 on the screen.
10:20:52 22 And what is the field of invention for the '703
10:20:55 23 patent?
10:20:55 24 A. So the '703 patent is related to voice commerce, and
10:21:01 25 particularly to systems and methods around streamlining

10:21:06 1 checkout. So it makes the checkout process more efficient.
10:21:14 2 Q. And if we could now have Figure 1, column three,
10:21:19 3 lines 20 to 21 of the '703 patent on the screen.
10:21:24 4 And what's shown there, sir?
10:21:28 5 A. All right. So this is a diagram that's showing
10:21:33 6 various components of the system. It's showing the voice
10:21:39 7 commerce application that has various components within it
10:21:42 8 that are managing profiles and managing the user interface,
10:21:46 9 that are then talking -- are working with storage devices
10:21:51 10 and processors, and have databases associated with them.
10:21:56 11 And then this whole system communicates to various
10:22:02 12 third-party external systems which would include search
10:22:05 13 engines and retailers and service providers.
10:22:08 14 Q. And if we could now have Figure 2 and column three,
10:22:12 15 lines 20 to 23 on the screen.
10:22:15 16 And what's shown in Figure 2, Dr. Polish?
10:22:20 17 A. So this is showing block diagrams for different
10:22:24 18 components of the systems and how they communicate, it's
10:22:27 19 actually showing something fairly similar to things we've
10:22:31 20 seen before. On the upper left are input devices such as an
10:22:36 21 Alexa, which would go into a speech recognition engine.
10:22:40 22 That would feed a Natural Language Processing system or NLU
10:22:45 23 and also talk to a database, and also talk to applications.
10:22:50 24 And ultimately the result of all that goes to output devices
10:22:55 25 where the system speaks.

10:22:57 Q. And if we could now have Figure 3, column 3, line 24
10:23:03 to 26.
10:23:06 Can you explain what's shown in Figure 3 of the
10:23:11 '703 patent?
10:23:13 A. This is a flow chart that is showing the operation of
10:23:19 the system from where it receives a natural language
10:23:23 utterance, and then the steps it takes to automatically
10:23:29 without further user input, complete the purchase and
10:23:33 ultimately complete the transaction for the user.
10:23:38 Q. And if you look in Step 304 and 310, there is
10:23:43 reference to without further user input. Do you see that?
10:23:46 A. Yes.
10:23:47 Q. What does that indicate?
10:23:48 A. Well that means that it's not waiting for some sort
10:23:52 of step from the user, that when it gets to that point, it
10:23:58 just effectively falls through and gathers the information
10:24:01 it needs from other sources besides the user. So it's an
10:24:06 ultimate streamlining of a purchase.
10:24:08 Q. Does that reduce the number of turns or interactions
10:24:11 the user has with the system?
10:24:13 A. Yes, if it can operate without further user input,
10:24:17 then there is no further turns at all, and it just completes
10:24:20 the transaction.
10:24:25 Q. If we could now have Figure 3, column eight, lines 19
10:24:29 to 24 on the screen.

10:24:37 And let's start to talk a little bit about
10:24:41 Step 304. Does the '703 patent explain how products could
10:24:46 be selected?
10:24:47 A. Yes, it talks about being able to use search and also
10:24:54 to be able to use groupings of products, and categorizations
10:25:01 of products to make the determination as efficiently as
10:25:07 possible.
10:25:07 Q. Now, if we could have Figure 3, column five, lines 44
10:25:11 to 46 on the screen. And let's talk about Steps 306, which
10:25:23 retrieve payment information that is to be used to pay for
10:25:26 the product or service, and 308, retrieve shipping
10:25:30 information that is to be used to deliver the product or
10:25:34 service.
10:25:34 Does the '703 patent explain how user profile
10:25:39 information could be used to perform these steps and
10:25:43 complete a transaction?
10:25:44 A. Yes, it's talking specifically about storing payment
10:25:48 information, storing shipping information, precisely so that
10:25:52 a purchase can be completed without having to ask for it
10:25:56 again.
10:26:01 Q. And, now, Dr. Polish, let's now turn to the product
10:26:06 at issue in this case, Amazon Alexa, okay?
10:26:09 A. Okay.
10:26:10 Q. And if we could have slide four on the screen. And
10:26:19 just kind of refresh, Dr. Polish, when you were doing your

10:26:23 1 analysis to determine whether or not Alexa infringed the
10:26:28 2 VoiceBox patents, what did you review and base your opinion
10:26:31 3 on?
10:26:31 4 A. I reviewed technical documents provided to me in the
10:26:35 5 case. I looked at the Court's construction order. I used
10:26:41 6 Alexa. And I confirmed what I saw by looking at source
10:26:48 7 code. And then I looked at the patents and compared it to
10:26:51 8 what I was seeing in Alexa.
10:26:54 9 Q. If you would open your binder, sir, and do you
10:26:58 10 recognize exhibit PTX-68?
10:27:05 11 A. Yes.
10:27:06 12 Q. And is that the document entitled ASR overview and
10:27:11 13 workflows?
10:27:11 14 A. That's right.
10:27:12 15 Q. Is that one of the technical documents from Amazon
10:27:15 16 that you based your opinion on?
10:27:16 17 A. Yes, it is.
10:27:17 18 MR. YOON: I would like to move into evidence
10:27:19 19 now, Your Honor, exhibit PTX-68.
10:27:22 20 MR. HADDEN: No objection.
10:27:24 21 THE COURT: Thank you. It is admitted.
10:27:26 22 (PTX Exhibit No. 68 was admitted into evidence.)
10:27:26 23 BY MR. YOON:
10:27:27 24 Q. And so do you see here sir on the bottom right corner
10:27:32 25 you can see there is an AMZ_VB_00046286. Do you see that?

10:27:41 A. **Yes.**

10:27:41 Q. **And that would be what I referred to earlier in this case as the Bates number and that would indicate that this document came from Amazon?**

10:27:44 A. **Yes, my understanding.**

10:27:48 Q. **And do you see that in this document here it's entitled ASR overview and work flow?**

10:27:49 A. **Yes.**

10:27:51 Q. **That's for Alexa, correct?**

10:27:56 A. **That's correct.**

10:28:00 Q. **If we can scan back. And for example, at the bottom, if you could just expand there, there is a logical -- and this document provides information such as a logical architecture overview of Alexa?**

10:28:02 A. **Yes, that's right.**

10:28:07 Q. **If we could now go to page 6. And if we could just expand what's shown on the top part of page 6, the diagram. Does Alexa have an ASR, sir?**

10:28:12 A. **Yes, it does.**

10:28:15 Q. **And where is that indicated?**

10:28:17 A. **So the box on the upper right of this diagram is labeled ASR, that's where the speech recognition takes place.**

10:28:19 Q. **Does the Amazon Alexa include an NLU engine?**

10:28:25 A. **Yes.**

10:28:34
10:28:37
10:28:38
10:28:39
10:28:43
10:28:48
10:28:48
10:28:53

10:28:53 Q. Where is that?
10:28:54 A. That's just in the box below, it's labeled NLU on it.
10:28:58 Q. Now, if we can now expand that document back.
10:29:01 And if you could now expand the bottom there
10:29:04 that says flow description. -- oh, actually, would go back
10:29:13 Mr. Smith, I apologize. If we can just expand out that top
10:29:17 part.
10:29:18 Dr. Polish, what's an ASR?
10:29:20 A. It's a speech recognizer, it's something that takes
10:29:24 speech and does some kind of recognition on it, turning it
10:29:27 into text or some other kind of more compact representation.
10:29:32 Q. And what is an NLU, sir?
10:29:34 A. An NLU is a Natural Language Understanding system,
10:29:39 which is some kind of -- can be referred to as Artificial
10:29:46 Intelligence, it tries to make sense of speech in the
10:29:50 context of a particular purpose. So it puts the speech into
10:29:54 some kind of context, you think of it as understanding or
10:29:59 assigning meaning.
10:30:00 Q. If we could have a split screen with this figure and
10:30:04 the flow description below. If we can just focus on the
10:30:33 flow description through steps 1 to 5 for a moment.
10:30:39 A. Okay.
10:30:40 Q. Mr. Smith, can you expand it more? And I'm only
10:30:47 doing this, Dr. Polish, to try to make it a little easier to
10:30:51 read.

10:30:52 A. Yes.
10:30:53 Q. Now, does this flow description and document describe
10:30:58 how Alexa operates?
10:31:00 A. Yes, in part it does.
10:31:01 Q. Could you walk through that for us?
10:31:04 A. Sure, it starts off at Item 1 where the examiner asks
10:31:13 Echo or Alexa what's the weather today. So you then have,
10:31:18 it notices the wake word, which is Alexa, and it streams the
10:31:23 rest of it, what is the weather today to the ASR. So that
10:31:31 goes there. The ASR takes it and does further processing on
10:31:38 it, and it comes up with what it's referring to as the NBest
10:31:43 utterances, which is what its NBest guesses are as to what
10:31:49 the audio stream contains.
10:31:52 Q. Is that Step 2?
10:31:54 A. Yes.
10:31:57 Q. And what happens next?
10:32:04 A. So the NBest gets sent to a, what is referred to as a
10:32:12 speechlet which is some application code that is oriented to
10:32:19 a particular task such as shopping, or getting the weather.
10:32:29 And ultimately, the NLU detects what is referred to as the
10:32:38 domain which is "Weather", in this case, and the intent is
10:32:42 "get the weather", and it says here no slot is detected.
10:32:47 And that means that the intent "GetWeather" would have slots
10:32:56 associated with it which are pieces of information that
10:33:02 would be needed to fulfill the request such as location and

10:33:05 1 time.
10:33:05 2 Q. If you look at 4(a), it says it goes in parallel
10:33:11 3 through both FST's, rule-based and exact match and
10:33:15 4 statistical models paths and then goes to a range to choose
10:33:20 5 the best output, do you see that?
10:33:22 6 A. Yes.
10:33:22 7 Q. We discussed in this case about rules-based and AI
10:33:25 8 and machine learning, do you recall that?
10:33:26 9 A. Yes.
10:33:27 10 Q. Does the NLU of Amazon Alexa use rules to determine?
10:33:32 11 A. Yes, my understanding is that there are rules-based
10:33:38 12 components to the NLU among other things that it does. So
10:33:44 13 it's definitely using rules and that's indicated here.
10:33:48 14 Q. If we can now go to and re-expand it, and focus on
10:33:54 15 steps 6 through 9. What happens after the NLU detects the
10:34:00 16 domain which is "Weather," the intent, which is
10:34:04 17 "GetWeather," and no slot is detected?
10:34:07 18 A. So what it is talking about here is that it's
10:34:11 19 noticing that it doesn't have all the information that it
10:34:16 20 needs, so it's saying, if you look at 6(b), where it says
10:34:24 21 that the "Weather" speechlet detects a missing slot, this
10:34:28 22 case location, and you can have a pre-configured rule that
10:34:32 23 says that it should now perform an action which is ask for
10:34:36 24 location in that case, you can have a different rule which
10:34:42 25 says use the location where the device is located, or some

10:34:46 1 other rule, but in this case, the rule is set to if you
10:34:49 2 don't have a location, ask for it. So it then goes off and
10:34:53 3 prompts the user for the location.
10:34:55 4 Q. And would the slot eventually be filled?
10:35:00 5 A. Yes, if the user cooperates and provides the
10:35:04 6 location, the slot would eventually be filled and that would
10:35:08 7 be satisfied and they could fulfill the request.
10:35:11 8 Q. If we could now turn in your binder to PTX-78.
10:35:16 9 A. Okay.
10:35:18 10 Q. And do you recognize exhibit PTX-78?
10:35:22 11 A. Yes, I do.
10:35:23 12 Q. And what is the title?
10:35:26 13 A. The title is history of Alexa Shopping and Alexa, and
10:35:30 14 it's dated December 17th, 2020.
10:35:35 15 Q. And this is -- was one Amazon technical documents you
10:35:40 16 reviewed to base your opinion on?
10:35:42 17 A. Yes.
10:35:42 18 MR. YOON: I would now like to move into exhibit
10:35:46 19 PTX Exhibit 78.
10:35:47 20 MR. HADDEN: No objection.
10:35:49 21 THE COURT: Thank you. It is admitted.
10:35:49 22 (PTX Exhibit No. 78 was admitted into evidence.)
10:35:49 23 BY MR. YOON:
10:35:50 24 Q. If we could now turn to page Amazon 8041. Generally,
10:35:58 25 what's shown in the diagram at the top of the page?

10:36:01 A. So this is a diagram similar to what we saw before,
10:36:06 it's showing the connection of the device, in this case the
10:36:10 Echo on the far left, passing information to this operating
10:36:15 system called DeeSOS, which communicates with the ASR and
10:36:23 communicates with the NLU and communicates with speechlets,
10:36:28 and potentially other Amazon internal services.
10:36:31 Q. See the input to the ASR is audio, correct?
10:36:34 A. Yes.
10:36:35 Q. And the output of the ASR then would be the text,
10:36:42 list of possible interpretations?
10:36:43 A. Yes.
10:36:43 Q. And would that be an input into the NLU?
10:36:46 A. Yes, it would.
10:36:47 Q. And what is the output of the Amazon Alexa NLU?
10:36:53 A. The output of the Amazon Alexa NLU is a collection of
10:37:07 intents and slots, each of them having a domain associated
10:37:12 with them. The output of the NLU is a collection of the
10:37:15 hypotheses of what the user's intent is, and how different
10:37:19 slots have been filled.
10:37:20 Q. Is the intent output indicated on this Amazon graph?
10:37:26 A. Yes, the right most are coming out of NLU, it's
10:37:30 labeled intent.
10:37:30 Q. Now, if we expand the page back. Let's go to
10:37:36 Section 2.1, the glossary, how Amazon defines the terms.
10:37:43 How does Amazon define ASR or Automatic Speech Recognition?

10:37:55 It's the sub-bullet point under SLU?
10:38:02 A. They say the ASR, Automatic Speech Recognition, is
10:38:04 the act of taking audio and turning it into recognized text.
10:38:09 Q. How does Amazon define NLU?
10:38:11 A. They say NLU is natural language recognition, I think
10:38:14 that may be a typo, I think it's Natural Language
10:38:18 Understanding, but the act of taking recognized text and
10:38:21 turning it into an intent.
10:38:22 Q. How does Amazon define intent?
10:38:27 A. They say intent is a language understanding concept
10:38:30 referring to the goal of the speech, what the speaker was
10:38:33 trying to communicate.
10:38:35 Q. And do you see the bullet point below it says at
10:38:40 Amazon, though, refer to intent, we are almost always
10:38:45 referring specifically to the blueshift intent format. Do
10:38:48 you see that?
10:38:50 A. Yes.
10:38:50 Q. Is that the acronym BIF?
10:38:54 A. Yes.
10:38:54 Q. Is that the output of the NLU?
10:38:56 A. Yes, the output of the NLU, my understanding of it is
10:38:59 a BIF, and a BIF is a collection of intents and scores for
10:39:11 them.
10:39:11 Q. If we could now go to PTX-84?
10:39:14 A. Okay.

10:39:15 Q. PTX-84, what's the title of that document?
10:39:18 A. Dialogue platform.
10:39:19 Q. Is this another one of the Amazon technical documents
10:39:23 that you based your opinion on?
10:39:24 A. Yes, it is.
10:39:25 MR. YOON: I would like to now move into
10:39:27 evidence, PTX-84.
10:39:29 MR. HADDEN: No objection.
10:39:30 THE COURT: Thank you. It is admitted.
10:39:32 (PTX Exhibit No. 84 was admitted into evidence.)
10:39:32 BY MR. YOON:
10:39:34 Q. Do you see on this document here, it has the
10:39:36 reference to dialogue platform?
10:39:38 A. Yes.
10:39:39 Q. And do you see there is a primary owner of this
10:39:42 Amazon technical document is Amazon Shopping?
10:39:47 A. Yes, I do.
10:39:48 Q. And do you see that it provides -- if we can go back.
10:39:55 If you can go to the section there, it provides an overview,
10:39:59 do you see that, Section 1?
10:40:00 A. Yes.
10:40:00 Q. And that includes an overview on conversation run
10:40:04 time, what's new to Alexa Shopping, conversation management,
10:40:09 do you see that?
10:40:09 A. Yes.

10:40:10 Q. If we can now turn to page 5509. And if we could
10:40:20 expand conversation runtime.
10:40:27 Dr. Polish, what are the three primary
10:40:33 components of the Alexa Shopping dialogue platform?
10:40:36 A. There is the Shopping FrontDoor, the Shopping
10:40:43 Orientation, and Shopping Conversation State Management.
10:40:49 Q. If we could expand shopping conversation state
10:40:53 management, let's start with the word "state." What is
10:41:03 state?
10:41:04 A. State generally refers to the state of a system or
10:41:09 all the different parts of the system at the current moment.
10:41:18 So in a person it might be, you know, weight, height, heart
10:41:26 rate, things like that, things that might change, but at a
10:41:30 given moment, it has a particular state.
10:41:33 Q. Does Alexa use states?
10:41:34 A. Yes, it does, it uses states as part of taking
10:41:38 somebody through a conversation to make something happen.
10:41:42 Q. Now, with regards to the Alexa Shopping platform,
10:41:49 what does shopping conversation state management do?
10:41:54 A. Well, what it does, what it is explaining here is
10:41:59 that it's keeping track of what's been done across time so
10:42:03 that it can help -- it can help a user make progress through
10:42:08 their purchase, or make progress through their transaction
10:42:12 and keep track of what they've done and what they need to
10:42:15 do.

10:42:15 Q. **Do you see the statement "we enable a smarter**
10:42:20 **conversation with customers by paying attention to what they**
10:42:23 **have said and done to make better decisions later." Do you**
10:42:27 **see that?**
10:42:27 A. **Yes.**
10:42:28 Q. **Is that done by Alexa?**
10:42:29 A. **Yes, it is.**
10:42:30 Q. **If we can now go to PTX-226. And this document, Your**
10:42:40 **Honor, was already admitted during the deposition of**
10:42:45 **Mr. Vanee.**
10:42:48 **THE COURT: Okay.**
10:42:49 **By Mr. Yoon:**
10:42:49 Q. **Looking at Exhibit 226, do you see this is entitled**
10:42:52 **Natural Language Understanding Deep Dive?**
10:42:54 A. **Yes.**
10:42:54 Q. **And you were here, you understand that one of the**
10:42:57 **authors of this document is an Amazon Alexa engineer, Kelly**
10:43:02 **Vanee?**
10:43:04 A. **Yes.**
10:43:05 Q. **And was this one of the documents you based your**
10:43:08 **opinion on with regards to Alexa?**
10:43:10 A. **Yes, it is.**
10:43:14 Q. **And if we could now go to page 2. And do you see**
10:43:21 **that this slide is entitled where does NLU fit in?**
10:43:25 A. **Yes.**

10:43:25 Q. And do you see there is a reference there to what's
10:43:28 the weather tomorrow? Do you see that?
10:43:32 A. Yes.
10:43:32 Q. Would that be an utterance by a user?
10:43:35 A. Yes. Where you're showing it there, that's -- the
10:43:40 utterance is spoken by the user into a device, and that is
10:43:45 then sent to the ASR.
10:43:47 Q. And you understand those wavy lines below would
10:43:50 indicate kind of the sound of a person's voice?
10:43:53 A. Yeah, that's sort of a time domain representation of
10:43:56 human speech.
10:43:57 Q. And does this slide demonstrate how the ASR in Alexa
10:44:02 and the NLU interact?
10:44:05 A. Yes, at a very high level.
10:44:07 Q. And could you explain that?
10:44:11 A. Sure. What it's showing here is that the person
10:44:14 speaks to the device, the speech goes to the ASR which then
10:44:23 turns it into text, which turns it into an NLU, and the NLU
10:44:28 outputs an intent which is in the form of get weather
10:44:32 forecast, and it has a slot, and the slot is the date,
10:44:38 March 18, 2016.
10:44:39 Q. And so when the user says "what's the weather
10:44:43 tomorrow," the NLU in Alexa would interpret that as the
10:44:49 intent of wanting the weather forecast?
10:44:51 A. Yes, it's the intent of the weather forecast and it

10:44:55 1 also has the -- it's interpreted tomorrow to mean that date.
10:44:59 2 Q. And if could go to the next page? And does this
10:45:07 3 slide indicate the stages of the Alexa NLU engine?
10:45:12 4 A. Yes, it gives you a little more detail, it's showing
10:45:15 5 there is an intent classification where it comes up with the
10:45:19 6 notion that this is a get weather forecast intent, and then
10:45:22 7 it has a separate step which is slot resolution where it's
10:45:26 8 trying to figure out how to fill in the date slot which it's
10:45:31 9 doing here with the word tomorrow.
10:45:33 10 Q. Do you see at the bottom the reference there to BIF?
10:45:36 11 A. Yes.
10:45:37 12 Q. That's the output of the NLU right?
10:45:39 13 A. Yes, that's the same BIF we saw before.
10:45:41 14 Q. So the NLU engine of Alexa would go through a
10:45:45 15 recognition step, and then a resolution step before it
10:45:50 16 outputted the BIF?
10:45:51 17 A. That's right.
10:45:52 18 Q. What is recognition?
10:45:54 19 A. I'm sorry, say that again.
10:45:55 20 Q. What does recognition refer to?
10:45:59 21 A. Recognition refers to that it's taking the recognized
10:46:04 22 speech, and it's breaking it up into pieces that it can then
10:46:09 23 use to further process and determine the intent.
10:46:13 24 Q. And then what's resolution, sir?
10:46:16 25 A. So resolution is where it's taking any number of

10:46:19 1 possible words that the user could have used to indicate
10:46:24 2 time, and turning it into a recognizable time entry.
10:46:32 3 Q. So that, for example, when you ask about the weather,
10:46:36 4 Alexa, before it gave you the weather would want to know
10:46:39 5 what date you wanted the weather associated with?
10:46:42 6 A. That's correct.
10:46:46 7 Q. If we could now go to page 10?
10:46:51 8 What is illustrated on this page?
10:46:53 9 A. This is another view on this process, this is where
10:46:56 10 somebody says "what is the weather in Seattle," recognition
10:47:01 11 takes place, and it turned it into domain is weather, and
10:47:06 12 the intent is get weather forecast.
10:47:10 13 Q. And this would be done by the NLU?
10:47:11 14 A. Yes.
10:47:17 15 Q. If we go to page 11. Now, we have been talking about
10:47:24 16 the recognition in the NLU. Do you recall that, sir?
10:47:29 17 A. Yes.
10:47:29 18 Q. And what's shown on this page?
10:47:31 19 A. So this is a particular view on what's taking place
10:47:34 20 in the NLU, particularly around recognition. It's showing
10:47:39 21 that a number of things are happening simultaneously. So
10:47:45 22 the input comes in on the left where it says input request,
10:47:48 23 and it's going to, in this case, four different places. The
10:47:53 24 first one it goes to is labeled exact match rules. So those
10:47:57 25 are rules that it -- that the recognition system would have

10:48:03 to look for particular words, to look for particular
10:48:06 patterns, and then the next thing over that diamond is a
10:48:12 decision, and if it matches, then it's done, it bypasses the
10:48:16 whole thing and goes to the slot filling step, it thinks
10:48:24 it's now figured out what the intent is.
10:48:26 If it doesn't match exactly, it tries to match a
10:48:30 little bit more collectively with that pattern of reject
10:48:35 rules, and if that matches, then it says okay, we're done.
10:48:39 If it doesn't work, then it waits to see if this other step
10:48:44 called FST returns anything, and the FST is going to be
10:48:50 slower than the rule matching. And finally, if that doesn't
10:48:54 return things, it will wait to see whether these models or
10:48:57 statistical models will complete and come up with their idea
10:49:02 of what the intent is.
10:49:04 So it will do all three of these things at once,
10:49:10 if the rules complete, then it bypasses the rest and doesn't
10:49:16 waste time, but if they don't, then it falls through to ever
10:49:20 more computationally approaches, as a fourth way, using this
10:49:26 totally separate system called the EVI recognizer, it can
10:49:32 also contribute.
10:49:32 Q. During the opening, we would been hearing about
10:49:35 machine learning and AI. Do you recall that sir?
10:49:39 A. Yes.
10:49:39 Q. Would that be the step relating to models?
10:49:41 A. Yes. I think when Amazon talks about machine

10:49:44 1 learning and AI, they're talking about the model there.
10:49:48 2 Q. What would be the FST step?
10:49:51 3 A. The FST step is something involving something
10:49:55 4 referred to as parsing, where it's trying to, in a more
10:50:05 5 deterministic rules-based approach, try to take apart the
10:50:10 6 words and understand what they mean using rules rather than
10:50:13 7 statistics.
10:50:14 8 Q. Let's talk about one possible operation in the Alexa
10:50:17 9 system. Okay, sir?
10:50:19 10 THE COURT: At the end of this, let's take our
10:50:21 11 break.
10:50:21 12 MR. YOON: Yes, let me just finish this
10:50:24 13 paragraph.
10:50:25 14 BY MR. YOON:
10:50:25 15 Q. In this example here, an utterance is received at the
10:50:28 16 ASR, it interprets it and detects and it provides that to
10:50:31 17 the recognition engine of the NLU core pipeline. Okay, sir?
10:50:37 18 A. Yes.
10:50:37 19 Q. And if that input satisfies the rules in the system,
10:50:45 20 it would reach resolution, correct?
10:50:46 21 A. Yes.
10:50:47 22 Q. And so, in the Alexa system when the NLU is
10:50:51 23 interpreting intent, there are situations where it will only
10:50:55 24 use rules and will not use AI or machine learning?
10:50:58 25 A. Yes, if a rule can be satisfied, it will satisfy the

10:51:03 1 rule preference over the other things because it's faster.
10:51:07 2 Q. And you have been -- I will?
10:51:09 3 MR. YOON: I'll stop there, Your Honor.
10:51:11 4 THE COURT: All right. Let's take our 15-minute
10:51:14 5 morning break.
10:51:15 6 COURTROOM DEPUTY: All rise.
10:51:17 7 (Jury exiting the courtroom at 10:51 a.m.)
10:51:44 8 THE COURT: All right. Anything we need to
10:51:46 9 discuss?
10:51:47 10 MR. YOON: No, Your Honor.
10:51:47 11 MR. HADDEN: No, Your Honor.
10:51:48 12 (A brief recess was taken.)
11:10:52 13 COURTROOM DEPUTY: All rise.
11:10:54 14 (Jury entering the courtroom at 11:10 a.m.)
11:11:14 15 THE COURT: All right, everyone welcome back.
11:11:20 16 Everyone else, you can be seated.
11:11:23 17 Please continue.
11:11:24 18 MR. YOON: Yes, Your Honor.
11:11:25 19 BY MR. YOON:
11:11:26 20 Q. Dr. Polish, do you recall earlier this morning we
11:11:29 21 were talking about how the NLU would score, come up with
11:11:35 22 hypotheses and provide scores?
11:11:37 23 A. Yes.
11:11:37 24 Q. And if we could now go to page 18 of Exhibit 226.
11:11:46 25 What is illustrated in this Amazon document regarding Alexa?

11:11:51 A. So this is showing if you look on the top, it says
11:11:56 utterance play Hunger Games, so the user has presumably said
11:12:00 play Hunger Games, and the system is trying to figure out
11:12:04 what that means, so it's come up with, in this case, four
11:12:09 different intent hypotheses. Two of them are music related,
11:12:18 one of them is books, one of them is video. So those are
11:12:22 potentially different domains. And it scored them with
11:12:27 these different ranked values on the right, and the blue
11:12:33 arrow means that it's picked the first one as the best
11:12:37 hypothesis, so it's assuming that when the user said play
11:12:42 Hunger Games, it's talking about a song named Hunger Games.
11:12:49 Q. If we go to the next page. With Alexa, if you make
11:12:56 the same utterance play Hunger Games, is there information
11:12:59 that the system could use that would result in the NLU
11:13:03 selecting something different than the scoring of just
11:13:07 Hunger Games?
11:13:08 A. Yes, so what this is showing at the top is in
11:13:11 addition to saying play Hunger Games, the device that that
11:13:15 request came from was a Fire TV, which is a video play back
11:13:22 device, so in that case, it has decided that the best one
11:13:27 was actually the last one, which was play a video called
11:13:32 Hunger Games.
11:13:33 Q. So, in the first example, just the words play Hunger
11:13:39 Games, it ranked play the song Hunger Games number one, with
11:13:45 just that information, right?

11:13:47 A. That's correct.
11:13:47 Q. But when it knew that the request came from a Fire
11:13:52 TV, it altered the selection because it came from a TV, to
11:13:56 video?
11:13:56 A. Right, that was additional information that was
11:13:59 available that changed the score, that changed the results
11:14:02 of its decision making.
11:14:03 Q. If we go to the next page. Does information stored
11:14:10 by the user, could that affect what's selected?
11:14:13 A. Right. So in this case, it's saying that intent
11:14:20 resolution, which is a component of the system which is
11:14:23 described in the document found that the user had an album
11:14:27 called Hunger Games in their cloud player. So in that case,
11:14:30 the user had previously purchased, or somehow obtained an
11:14:35 album called Hunger Games and stored it. So now when it
11:14:39 goes through this recognition process, the hypothesis that
11:14:44 you mean the album Hunger Games, that becomes the best
11:14:50 hypothesis.
11:14:51 Q. Could that purchase of the album Hunger Games come
11:14:56 from buying it over Alexa?
11:14:57 A. Certainly.
11:14:58 Q. Could that be based on a prior conversation?
11:15:01 A. Yes.
11:15:01 Q. And so, how does the information that Alexa has
11:15:06 determine what the intent interpretation is?

11:15:09 A. So, the information that Alexa has is both -- is
11:15:18 long-term information about the user, it is information
11:15:21 provided by the user from previous interactions, and all
11:15:24 that goes into interpreting what the user said and going
11:15:30 into ranking the hypotheses and picking out which one is the
11:15:34 most likely to be correct.
11:15:36 Q. If we go to the next page. There is a page 22,
11:15:44 what's this show?
11:15:47 A. So this is showing that this is talking about this
11:15:50 sort of re-ranking process of reconsidering what the user
11:15:57 means, and it's showing that the re-ranker can use system
11:16:03 features such as the ASR, the recognizer, what they call
11:16:09 entity resolution, and context.
11:16:11 Q. And the example we discussed earlier, it used entity
11:16:16 resolution, correct?
11:16:17 A. Yes.
11:16:17 Q. And that would result in a change from range 1 or 2?
11:16:22 A. Yes, that's what happened in that case.
11:16:25 Q. So originally, it was the song Hunger Games had a
11:16:30 score of .89, and the album Hunger Games had a 4.7?
11:16:37 A. Yes.
11:16:37 Q. After the resolution entity occurred, what was the
11:16:41 ranking of the album?
11:16:43 A. The album was .82.
11:16:45 Q. What was the song?

11:16:46 A. .6.
11:16:49 Q. Could context also be used by the Alexa system to
11:16:52 cause a re-ranking or readjustment?
11:16:57 A. Yes.
11:16:58 Q. And if we go to the previous page to this one. And
11:17:04 what's shown here on this Alexa technical document?
11:17:08 A. This is just some text on this presentation that says
11:17:11 based on additional signals or features from the rest of the
11:17:13 system, the recognition results can be rescored and
11:17:18 re-ranked. So the signals are features is a very broad
11:17:23 statement of pretty much any kind of input could be taken to
11:17:28 influence what the ranking is on these different hypotheses.
11:17:33 Q. One of those inputs could be context?
11:17:35 A. That's correct.
11:17:36 Q. If we could now go to PTX-472. And Dr. Polish, do
11:17:48 you recognize exhibit PTX-472?
11:17:50 A. Yes, I do.
11:17:51 Q. Is that one of the documents you relied on in forming
11:17:53 your opinion and understanding regarding the design and
11:17:56 operation of Alexa?
11:17:57 A. Yes.
11:17:57 MR. YOON: I would like to now move into
11:17:59 evidence exhibit PTX-472.
11:18:01 MR. HADDEN: No objection.
11:18:03 THE COURT: Thank you. It is admitted.

11:18:03 1 (PTX Exhibit No. 472 was admitted into
11:18:04 2 evidence.)
11:18:04 3 BY MR. YOON:
11:18:07 4 Q. And do you see there is a reference there to speaker
11:18:16 5 ID cloud. Do you see that?
11:18:18 6 A. Yes.
11:18:18 7 Q. And what is that, sir?
11:18:21 8 A. So speaker ID is where the system identifies
11:18:27 9 different speakers from within a household, so there is a
11:18:31 10 speaker registration process where you can say a few
11:18:36 11 sentences, and it learns your voice so that particularly if
11:18:39 12 you have multiple people in the same household, it can
11:18:43 13 distinguish those different people within the household and
11:18:46 14 do different things depending on who they are.
11:18:48 15 Q. If could go to page 2 of this document, and if you
11:18:53 16 would expand speaker ID cloud, speaker identification
11:18:57 17 service. And what's that, sir?
11:19:00 18 A. So this is just -- this is talking about what I just
11:19:05 19 said, which is that it's using this service to distinguish
11:19:12 20 individual users of the device in a single household. They
11:19:15 21 would use the speaker ID service.
11:19:19 22 Q. And is this -- what type of information is this in
11:19:22 23 the system?
11:19:23 24 A. So this would be long-term information, so your
11:19:28 25 registration with the system when you enroll yourself in

11:19:33 1 this, would retain some piece of information that would be
11:19:36 2 retained from conversation to conversation, presumably over
11:19:40 3 the course of months or years, that would be used by the
11:19:44 4 system to customize your experience depending upon who you
11:19:49 5 are.
11:19:49 6 Q. Looking at the last sentence there, it says "the
11:19:51 7 incoming audio stream is forked to the speaker ID service,
11:19:56 8 and this service helps in determining who in the household
11:19:59 9 is currently using the device." Do you see that?
11:20:02 10 A. Yes.
11:20:03 11 Q. What does that indicate?
11:20:04 12 A. When they say forked, what they mean is that the
11:20:09 13 stream, audio stream is split, so it's going to both the
11:20:14 14 speaker ID service, as well as, the speech recognizers, this
11:20:20 15 is happening without causing some sort of disruption, so
11:20:24 16 it's happening in parallel if you will to the regular speech
11:20:31 17 recognition.
11:20:31 18 Q. Does Alexa, ASR and NLU use this speaker ID
11:20:36 19 information in interpreting what is said?
11:20:39 20 A. Yes.
11:20:39 21 Q. How so?
11:20:40 22 A. Well, it knows, it learns from this -- who the user
11:20:45 23 is, and loads various profiles and information about the
11:20:49 24 user that is then used to direct the conversation.
11:20:54 25 Q. And if we could now expand back and do you see there

11:20:57 1 is another reference there at the top to NLU cloud-Natural
11:21:04 2 Language Understanding service?
11:21:06 3 A. Yes.
11:21:06 4 Q. And what does that indicate?
11:21:08 5 A. So that's sort of an overall wrapper for the NLU,
11:21:13 6 it's saying that the input is a SIRF, S-I-R-F, which is just
11:21:21 7 the speech, and then some other things including user info,
11:21:25 8 information about the user, the user's history, so it's
11:21:30 9 taking all these inputs and its output is a BIF, which we've
11:21:35 10 seen before is a collection of scored intents.
11:21:44 11 Q. And so do you see that with the input to the NLU and
11:21:48 12 the ASR be the SIRF, the hint, user info and history?
11:21:56 13 A. Yes.
11:21:56 14 Q. And the output is the intent determination of the
11:21:59 15 NLU?
11:22:00 16 A. Yes.
11:22:00 17 Q. If we could take that down. Thank you.
11:22:02 18 Now, Dr. Polish, do you know of Mr. John Peck?
11:22:07 19 A. Yes.
11:22:07 20 Q. And how did you work with Mr. Peck for your analysis
11:22:10 21 in this case?
11:22:11 22 A. So John Peck is a -- is a very experienced programmer
11:22:19 23 and software analyst, and he reads lots of source code and I
11:22:28 24 read source code too, but I didn't have enough time to put
11:22:31 25 into looking at everything that was required. So he was

11:22:36 1 retained by the lawyers in the case to look at Amazon's
11:22:40 2 source code, and when I had questions about things I wanted
11:22:44 3 to understand about the source code, I told lawyers about
11:22:48 4 it, they instructed Mr. Peck, and he went off and did the
11:22:53 5 research. And I got reports back from him, including a
11:22:58 6 report that he filed in this case of what he found.
11:23:02 7 Q. Now, Dr. Polish, have you helped prepare a video
11:23:06 8 illustrating a conversation with Alexa?
11:23:09 9 A. Yes, I have.
11:23:10 10 Q. And if we could have PDX-3-7 on the screen.
11:23:22 11 And when you were having your actual
11:23:25 12 conversation with Alexa, what features did you look for?
11:23:27 13 A. So I went through this conversation with Alexa
11:23:33 14 particularly looking for Voice ID, the presence of long-term
11:23:41 15 information, the presence of short-term information, I
11:23:45 16 showed voice purchasing, I showed pronoun use which is
11:23:50 17 important in one of the patents, and I showed a voice
11:23:54 18 advertisement. And when you watch the video, you can see
11:23:58 19 all these -- all these elements come out in the video.
11:24:03 20 Q. Why don't we play the first video now.
11:24:12 21 (PDX-3-7 video played.)
11:27:17 22 MR. YOON: If we can play the next video.
11:27:22 23 (Video played.)
11:28:00 24 MR. YOON: Your Honor, at this time, we would
11:28:07 25 like to move into physical evidence PTX-674.

11:28:13 1 MR. HADDEN: No objection, Your Honor.
11:28:15 2 THE COURT: Thank you. It is admitted.
11:28:15 3 (PTX Exhibit No. 674 was admitted into
11:28:21 4 evidence.)
11:28:21 5 BY MR. YOON:
11:28:27 6 Q. Now, Dr. Polish, let's now go into the specifics of
11:28:35 7 your infringement analysis. If we could have PDX3-9 on the
11:28:41 8 screen.
11:28:45 9 And focusing on the '681 patent, Dr. Polish,
11:28:50 10 what's your opinion with respect to Claim 13 of the '681
11:28:54 11 patent?
11:28:55 12 A. It's my opinion that Claim 13 is infringed by the
11:28:59 13 Alexa system.
11:29:00 14 Q. And now let's discuss Claim 13. If I could have
11:29:05 15 PDX3-10 on the screen. Do you see on that screen is the
11:29:10 16 full language of Claim 13 of the '681 patent?
11:29:14 17 A. Yes.
11:29:15 18 Q. Dr. Polish, what's a nontransitory computer readable
11:29:20 19 medium containing computer-executable instructions?
11:29:24 20 A. That's patent language for some piece of storage
11:29:30 21 computer code, so the computer executable instructions are
11:29:35 22 just computer code, and nontransitory means that it's stored
11:29:41 23 somewhere.
11:29:41 24 Q. So focusing on the first part of Claim 13, what we
11:29:44 25 call the preamble, if we can just expand that Claim 13

11:29:56 1 **begins a nontransitory computer readable medium containing**
11:30:00 2 **computer-executable instructions for providing a cooperative**
11:30:05 3 **conversational voice user interface, the computer-executable**
11:30:10 4 **instructions operable when executed to." Is that present in**
11:30:16 5 **Alexa?**
11:30:16 6 A. **Yes, it is.**
11:30:17 7 Q. **And what's the basis of that opinion?**
11:30:19 8 A. **Well, it's Alexa runs on servers, the servers are**
11:30:27 9 **computers, they're running code, I have seen code. And**
11:30:32 10 **those servers are loaded from non-transitory sources, so the**
11:30:40 11 **code is written, and it's loaded into those server computers**
11:30:44 12 **that is then operable when executed.**
11:30:47 13 Q. **When the code that's stored on the Amazon servers is**
11:30:52 14 **executed, does it practice the invention of Claim 13?**
11:30:55 15 A. **Yes, it does.**
11:30:56 16 Q. **If we go to the next element of Claim 13, if we could**
11:30:59 17 **expand that. Do you see that the next element is "receive**
11:31:02 18 **an utterance at a voice input device, during a current**
11:31:07 19 **conversation with a user, wherein the utterance includes one**
11:31:10 20 **or more words that have different meanings in different**
11:31:14 21 **contexts." Do you see that?**
11:31:17 22 A. **Yes.**
11:31:17 23 Q. **Is that present in the Alexa system?**
11:31:19 24 A. **Yes, it is.**
11:31:20 25 Q. **And what's the basis of that opinion?**

11:31:22 A. Well, technical documents that I've seen, as well as,
11:31:29 from my own use of the system.
11:31:33 Q. And if you would look at now PDX3-11.
11:31:45 Does this slide illustrate how the Alexa system
11:31:50 meets the receive an utterance requirement?
11:31:53 A. Yes. This is showing an Alexa device receiving
11:32:00 speech and that the speech, the utterance is being sent to
11:32:06 the Alexa voice service.
11:32:07 Q. And when someone speaks to Alexa, can we speak in one
11:32:11 or more words that have different meanings and different
11:32:15 contexts?
11:32:15 A. Yes, they can.
11:32:16 Q. If we now go back to the next element of the claim,
11:32:20 PDX3-10. If we could have the next one on the screen,
11:32:25 accumulate short-term. This element recites "accumulate
11:32:30 short-term shared knowledge about the current conversation,
11:32:33 wherein the short-term shared knowledge includes knowledge
11:32:36 about the utterance received at the voice during the current
11:32:40 conversation."
11:32:41 Is that element present in that by Alexa?
11:32:48 A. Yes, it is.
11:32:48 Q. What's the basis of that opinion?
11:32:50 A. Well, again, technical documents that I've seen, my
11:32:54 own use -- and my own use of the system.
11:32:57 Q. If we go to the next slide, please.

11:33:05 1 Was the accumulation of short-term knowledge
11:33:09 2 illustrated in the demonstration that you provided in the
11:33:13 3 video?
11:33:13 4 A. Yes.
11:33:15 5 Q. How so?
11:33:16 6 A. Well, so when I was setting the alarm and I said
11:33:21 7 6:30, it knew 6:30, it wasn't -- it had multiple meanings,
11:33:28 8 6:30, it could be in the morning or in the evening, so it
11:33:31 9 retained that short-term knowledge for the further
11:33:36 10 conversation to disambiguate between 6:30 a.m. and p.m.
11:33:44 11 Q. Is that when Alexa asked is that 6:30 in the morning
11:33:47 12 or the evening?
11:33:48 13 A. Yes, that was when Alexa realized that that -- that
11:33:54 14 it didn't have enough information to set the alarm, its slot
11:33:59 15 wasn't full, wasn't fully filled out, so it had to get more
11:34:03 16 information which it did by soliciting me for more
11:34:07 17 information.
11:34:08 18 Q. If we can now go to PDX3-14. Does this also provide
11:34:13 19 support for your opinion that Alexa stores short-term
11:34:20 20 knowledge and meets that element?
11:34:30 21 A. So this is from a technical document, actually we
11:34:34 22 looked at that document before, this is what it does when it
11:34:37 23 has ambiguity that it needs to figure out, and so here it
11:34:41 24 takes what they call a two-step approach where it creates
11:34:46 25 another, effectively another intent looking for some kind of

11:34:51 1 response from the user, and it goes out and makes a request
11:34:57 2 of the user and then tries to interpret what the user says
11:35:01 3 as a response to that disambiguous request.
11:35:04 4 Q. So in the example here, when it hears the number six,
11:35:10 5 it could be either an intent to set the volume, it could be
11:35:16 6 a song name, or it could mean this is a put multi-turn?
11:35:26 7 A. That's correct.
11:35:26 8 Q. In the example you gave in your video, that was
11:35:30 9 multi-turn?
11:35:31 10 A. That was 6:30, followed by a request for morning or
11:35:35 11 evening, followed by my saying morning.
11:35:37 12 Q. And PTX-226, that's the deep dive document that
11:35:42 13 Mr. Vanee was the author of what we discussed, correct?
11:35:46 14 A. That's right.
11:35:46 15 Q. Go to the next element.
11:35:48 16 Can we go to PDX3-15, please. And what is --
11:36:01 17 does that document that you have been discussing provide
11:36:04 18 further support for your opinion about whether the
11:36:07 19 accumulate short-term knowledge is met?
11:36:10 20 A. Yes, it does. This is talking -- this is again from
11:36:13 21 that same document. This is talking about a component
11:36:17 22 called the context interpreter that merges the results of
11:36:22 23 the disambiguation request of what came before so as to come
11:36:29 24 up with the intent.
11:36:36 25 Q. If could go to PDX3-16. And this is also an

11:36:41 1 Exhibit 226, how does this support your opinion regarding
11:36:46 2 accumulating short-term knowledge about a current
11:36:50 3 conversation?
11:36:50 4 A. Right, so this is a work example from that slide
11:36:56 5 deck. The utterance is set for 6 o'clock. It thinks it's
11:37:03 6 setting an alarm, and it sees there are two possibilities,
11:37:10 7 0600 or 1800 or 6:00 p.m., and so it needs to have a
11:37:18 8 dialogue which is slot solicitation, that is asking for more
11:37:27 9 information.
11:37:27 10 Q. Here it doesn't know whether it's 6:00 a.m. or
11:37:31 11 6:00 p.m.?
11:37:32 12 A. That's correct.
11:37:32 13 Q. And 6:00 p.m. is indicated by 1800 hours?
11:37:37 14 A. Right.
11:37:37 15 Q. If we can go to the next element in the claim, go
11:37:42 16 back to PDX3-10, if we can expand the next element. This
11:37:48 17 also requires, "accumulate long-term shared knowledge about
11:37:51 18 the user, wherein the long-term shared knowledge includes
11:37:55 19 knowledge about one or more past conversations with the
11:37:58 20 user." Do you see that?
11:37:58 21 A. Yes.
11:37:58 22 Q. Is that element practiced in that by Alexa?
11:38:01 23 A. It is.
11:38:02 24 Q. What's the basis of your opinion?
11:38:04 25 A. Technical documents that I have seen, and my own use

11:38:07 1 of it including from that demo you saw.
11:38:09 2 Q. And if we could now go to PDX3-17. And this is
11:38:16 3 another document that we've -- that you've relied on.
11:38:23 4 What's shown here in PDX3-17?
11:38:25 5 A. Right, so this is a document, it's a technical
11:38:28 6 document that's talking about among other things, the
11:38:31 7 speaker ID system that we were talking about before where it
11:38:37 8 talks about that if people in a household can register
11:38:41 9 themselves or enroll themselves in this program which allows
11:38:45 10 the system to differentiate people in the household. And
11:38:48 11 this would be a form of long-term shared knowledge.
11:38:51 12 Q. And in order to do the speaker ID, you had a
11:38:56 13 conversation with Alexa, correct?
11:38:58 14 A. Yes, it was done within Alexa, as you saw in the
11:39:01 15 demo, it's part of Alexa.
11:39:03 16 Q. If we can now go to the next slide. And how did your
11:39:07 17 video here illustrate that -- how did your video illustrate
11:39:19 18 how Alexa accumulates long-term shared knowledge about the
11:39:23 19 user?
11:39:23 20 A. Well, in this case, you can see that I'm changing an
11:39:28 21 alarm. So the long-term information was the old setting of
11:39:35 22 the alarm, and I am changing it to a new time. So it has --
11:39:40 23 it's accumulated long-term information that informs the
11:39:45 24 conversation and makes the conversation more streamlined.
11:39:47 25 Q. At this time you had already done the speaker ID?

11:39:51 A. Yes.
11:39:51 Q. And Alexa recognized you as Nat?
11:39:53 A. That's right.
11:39:54 Q. Was that also a form of long-term shared knowledge
11:39:57 from the communication?
11:39:58 A. Yes, it was.
11:40:00 Q. If we could now go to slide PDX3-19. And how do the
11:40:07 documents here that you based your opinion on illustrate the
11:40:11 use of long-term shared knowledge?
11:40:13 A. So this is a document that talks about what they call
11:40:18 an enrollment profile, mutable model, which as it's
11:40:28 described here, this is something that would listen to your
11:40:31 voice, understand some of the acoustic properties of your
11:40:34 voice and your environment and make changes to the, to this
11:40:39 model that it stores that would cause the ASR to do a better
11:40:44 job over time.
11:40:48 Q. And now let's go to the next element. And we could
11:40:52 have that next element on the screen identified. The next
11:41:07 element of Claim 13 refers to identify a context associated
11:41:12 with the utterance, wherein a conversational speech engine
11:41:16 identifies the context associated with the utterance from
11:41:19 the short-term shared knowledge and the long-term shared
11:41:24 knowledge." Do you see that?
11:41:25 A. Yes.
11:41:25 Q. And is that element met by Alexa?

11:41:29 A. It is.
11:41:30 Q. And what's the basis for your opinion, sir?
11:41:32 A. The technical documents and my own use of the system.
11:41:36 Q. And would one of the documents be the NLU deep dive
11:41:41 document which is PTX-226?
11:41:45 A. Yes.
11:41:47 Q. Can we go -- if we can now go to the next element.
11:41:56 And this element says "establish an intended meaning for an
11:42:01 utterance within the identified context, wherein the
11:42:04 conversational speech engine establishes the intended
11:42:08 meaning within the identified context to disambiguate an
11:42:12 intent that the user had in speaking the one or more words
11:42:16 that have the different meanings in the different contexts."
11:42:21 Is that element met by Alexa?
11:42:23 A. Yes, it is.
11:42:24 Q. And what do you base that opinion on?
11:42:26 A. The technical documents that we've seen, as well as,
11:42:29 my own use of it.
11:42:30 Q. If we can now go to PDX3-21. How did your opinions,
11:42:44 how were they supported by PTX-226?
11:42:53 A. We talked about this before, what's going on here is
11:42:56 where you have some kind of ambiguity that the system needs
11:43:00 to resolve and it has, and it gets more information and then
11:43:06 this context interpreter needs to run to combine the result
11:43:11 of this ambiguation request with what came before. So this

11:43:18 1 is combining information from different con -- from
11:43:24 2 different utterances into the context.
11:43:26 3 Q. If we can go now to the next slide?
11:43:31 4 MR. HADDEN: Objection, Your Honor. This is not
11:43:33 5 a slide that was in the demonstratives we were given in the
11:43:38 6 exchange.
11:43:39 7 THE COURT: Can you talk to me.
11:43:43 8 MR. HADDEN: Your Honor, our understanding is
11:43:45 9 this slide was withdrawn per our agreement this morning.
11:43:49 10 MR. YOON: Your Honor, there appears to be an
11:43:52 11 error, counsel is correct. We had made a change, but the
11:43:55 12 change did not pop up. We can put the correct slide on now,
11:43:58 13 and I'll remove it from the demonstrative, I'll redo it now
11:44:02 14 Your Honor with the proper demonstrative.
11:44:04 15 MR. HADDEN: I'll move to strike the testimony
11:44:07 16 with the demonstrative that we agreed wouldn't be shown.
11:44:10 17 THE COURT: We'll strike that.
11:44:12 18 This was some testimony given, it was just a
11:44:15 19 mixup with the slide. They're going to put the correct
11:44:18 20 slide up and redo that testimony, and the testimony that you
11:44:20 21 saw on the prior slide, we'll strike.
11:44:26 22 BY MR. YOON:
11:44:26 23 Q. And so let me just go back really quick. If we could
11:44:30 24 have PDX3-10 on the screen. PDX3-10.
11:44:43 25 Looking at -- could I have the element -- we can

11:44:49 do it this way. If I could just have the element on the
11:44:53 screen, identify a context associated, we can go over that
11:44:57 again. Just expand the identify. Thank you.
11:45:05 Dr. Polish, with respect to this element, do you
11:45:09 see that it says, "identify a context associated with the
11:45:12 utterance, wherein a conversational speech engine identifies
11:45:16 the context associated with the utterance from the
11:45:20 short-term shared knowledge and the long-term shared
11:45:23 knowledge."
11:45:24 A. I do.
11:45:25 Q. And is that element met by Alexa?
11:45:27 A. It is.
11:45:28 Q. And what's your basis for your opinion?
11:45:31 A. Technical documents and my own use of the system.
11:45:34 Q. Can you explain how the technical documents supported
11:45:36 your opinion?
11:45:37 A. Well, what this is talking about is identifying the
11:45:40 context which is, in part anyway, an intent, so the NLU
11:45:49 generates an intent, or a collection of intents based upon
11:45:56 the speech of the user, and the long-term information that
11:46:02 informs it, as well as, the short-term information from
11:46:08 prior utterances within the same conversation.
11:46:12 Q. Let's go to the next element. We'll just expand it
11:46:15 to the next element now. If we could go to establish. The
11:46:20 next element in the claim says "establish an intended

11:46:24 1 meaning for the utterance within the identified context,
11:46:27 2 wherein the conversational speech engine establishes the
11:46:30 3 intended meaning within the identified context to
11:46:34 4 disambiguate an intent that the user had in speaking the one
11:46:40 5 or more words that have different meanings in the different
11:46:43 6 contexts." Do you see that?
11:46:45 7 A. Yes.
11:46:45 8 Q. Is that element met by Alexa?
11:46:47 9 A. It is.
11:46:48 10 Q. What's the basis for your opinion?
11:46:50 11 A. Technical documents and my own use of the system.
11:46:52 12 Q. And if we could now go, I want to go to PDX3-22,
11:46:58 13 we're skipping 21, go to 22. 22. PDX3-22.
11:47:37 14 Could you just advance it to 22, please. If we
11:48:11 15 could just go back up with regards to, put PDX3-10 up again.
11:48:22 16 And sir, can you explain how Alexa does this, please?
11:48:27 17 A. So this is talking about the disambiguation of words
11:48:34 18 that have one or more meanings, that have more than one
11:48:40 19 meaning. So we've seen this already with when I said 6:30
11:48:47 20 which could mean two different things, and it went off and
11:48:52 21 established which it meant. There was also, we saw in the
11:49:00 22 technical document, we saw Hunger Games being interpreted,
11:49:05 23 there were different possible meanings for that, a song, an
11:49:09 24 album, a video, and Alexa went off and figured out,
11:49:13 25 disambiguated which it meant, and it used short-term

11:49:18 information to do that to establish what the meaning was in
11:49:26 the different topics.
11:49:29 Q. Going back to Exhibit 226, how is this demonstrated
11:49:36 in terms of the establishing the intended meaning?
11:49:39 A. Right. So this is the example of Hunger Games, and
11:49:43 so Hunger Games can mean in this case, I guess four
11:49:47 different things, a song, an album, a book and a video. And
11:49:51 it was able to use, in this case long-term information which
11:49:57 was whether somebody had purchased the album before to
11:50:01 disambiguate it and decide that it meant the album.
11:50:07 Q. We can keep that language up and then turn to slide
11:50:23 22 of 226. Advance two pages from the right. Thank you.
11:50:39 And we looked at this during your testimony.
11:50:42 How does this support the basis for your opinion that the
11:50:47 identified context is to disambiguate an intent that the
11:50:52 user had in speaking one or more words?
11:50:55 A. So, this is showing that you can re-rank, or in this
11:51:03 case, that would mean disambiguate the intent of the user
11:51:10 and the meaning of the words based upon the ASR, the ER, or
11:51:17 the content, and that can be long-term information,
11:51:22 short-term information, it could be further information from
11:51:25 the user.
11:51:28 Q. And did that support your opinion, sir?
11:51:30 A. It does.
11:51:31 Q. If we could now go back to PTX3-10 and go to the last

11:51:37 1 element of Claim 13 of the '681 patent. You see this is
11:51:42 2 "generate a response to the utterance where in the
11:51:44 3 conversational speech engine, grammatically or syntactically
11:51:49 4 adapts the response based on the intended meaning
11:51:52 5 established within the identified context." Do you see
11:51:54 6 that?
11:51:54 7 A. Yes.
11:51:55 8 Q. Is that element met by Alexa?
11:51:58 9 A. It is.
11:51:58 10 Q. What is the basis of that opinion?
11:52:00 11 A. Documents that I have seen, and how I have used the
11:52:03 12 system.
11:52:04 13 Q. We now have PDX3-24. And is that element met, the
11:52:39 14 last element of Claim 13 met by Alexa and supported by the
11:52:44 15 technical documents?
11:52:45 16 A. Yes. So this technical document is showing after the
11:52:50 17 utterance has gone through the system, ultimately what comes
11:52:54 18 out is a voice response that is played back through the
11:52:57 19 speaker.
11:52:58 20 Q. And if we could go to Exhibit 226, slide 32. This is
11:53:11 21 a multi-turn example for Alexa, correct?
11:53:14 22 A. Yes.
11:53:14 23 Q. And this example is very similar to what you used in
11:53:19 24 your video, correct?
11:53:20 25 A. Yes, it is.

11:53:20 Q. And this is the example set forth in the Amazon
11:53:24 technical documents that was the deep dive?
11:53:27 A. Yes, it is.
11:53:28 Q. And would Claim 13 of the '681 patent be practiced
11:53:34 during the multi-turn example?
11:53:36 A. Yes, absolutely.
11:53:39 Q. Now, there is some discussion in this case about like
11:53:43 a single-turn conversation. Are you familiar with that?
11:53:46 A. Yes, generally.
11:53:47 Q. And would single-turn conversations include
11:53:53 short-term shared knowledge?
11:53:54 A. It certainly could. It wouldn't -- not all would
11:53:58 necessarily, but you certainly could have a single-turn that
11:54:02 would have short-term information in it.
11:54:05 Q. And where the single-turn have short-term
11:54:09 information, would Alexa also use long-term information?
11:54:13 A. The long-term faster information is used with speaker
11:54:16 ID, there is lots of long-term information that would be
11:54:20 used all the time, so in a situation where a single-turn is
11:54:25 using short-term, it would be using both.
11:54:27 Q. And in the case where a single-turn utterance
11:54:30 included short-term shared knowledge, would Claim 13 be
11:54:34 practiced?
11:54:34 A. Yes.
11:54:35 Q. Now, let's turn to your opinion as to the '176

11:54:41 1 patent, the first of the voice ad patents. And if we could
11:54:48 2 have PDX3-26 on the screen, that's Claim 40.
11:54:58 3 A. Okay.
11:54:59 4 Q. Dr. Polish, what is your opinion with respect to
11:55:02 5 Claim 40 of the '176 patent?
11:55:05 6 A. That it is infringed by the Alexa system.
11:55:09 7 Q. Now, this is Alexa Shopping, correct?
11:55:12 8 A. Yes.
11:55:12 9 Q. Would Alexa Shopping involve multi-turns?
11:55:17 10 A. Yes, it certainly can.
11:55:19 11 Q. And would that -- now, let's talk about the preamble
11:55:24 12 of Claim 40. Do you see that this preamble recites a system
11:55:32 13 for selecting and presenting advertisements in response to
11:55:36 14 processing natural language utterances, do you see that?
11:55:40 15 A. Yes.
11:55:41 16 Q. Is that present in Alexa?
11:55:42 17 A. Yes, it is.
11:55:43 18 Q. And what's the basis for your opinion?
11:55:45 19 A. Technical documents and the use of it that I showed.
11:55:51 20 Q. So let's now go to the next element. Does Alexa
11:56:00 21 include an input device that receives a natural language
11:56:03 22 utterance containing at least one request at an input
11:56:06 23 device?
11:56:07 24 A. Yes.
11:56:07 25 Q. And what's the basis for your opinion?

11:56:09 A. Technical documents and my own use of it.
11:56:14 Q. And if we could look at now PDX3-27. And does this
11:56:22 support your opinion that Alexa includes an input device
11:56:28 that receives a natural language utterance?
11:56:30 A. Yes, this is showing the input device and it's
11:56:34 showing the user's utterance being sent to the Alexa voice
11:56:37 service.
11:56:37 Q. If we could now go back to PDX3-10 -- PDX3-26. Thank
11:56:46 you. Looking at the next element, "a speech recognition
11:56:51 engine coupled to the input device, wherein the speech
11:56:55 recognition engine recognizes one or more words or phrases
11:56:58 in the natural language utterance." Is that element met by
11:57:04 Alexa?
11:57:05 A. Yes, it is.
11:57:05 Q. What's the basis for your opinion?
11:57:07 A. The technical documents and my own use.
11:57:09 Q. And if we can now go to the next slide. And how does
11:57:19 the technical documents show that Alexa practices this
11:57:26 element?
11:57:26 A. Right. So these documents are showing that the
11:57:29 speech comes in on the input device and is sent to the ASR
11:57:34 which is in the upper right, the Automatic Speech
11:57:38 Recognition system which is also shown in Figure 1 of the
11:57:41 '176 patent as the ASR in the Figure 1.
11:57:46 Q. Going to the next element. This element of the claim

11:57:53 1 recites "a conversational language processor coupled to the
11:57:58 2 speech recognition engine, wherein the conversational
11:58:02 3 language processor is configured to." Do you see that?
11:58:04 4 A. Yes.
11:58:05 5 Q. Is that present in Alexa?
11:58:07 6 A. Yes, it is.
11:58:08 7 Q. And what's the basis for your opinion?
11:58:11 8 A. From the technical documents that I have seen.
11:58:14 9 Q. And how does Alexa meet this?
11:58:16 10 A. The conversational language processor is the NLU and
11:58:24 11 some other components, it's also including -- it's also
11:58:29 12 including speechlets as well, as part of that, that we
11:58:33 13 talked about, but in this diagram that's contained within
11:58:36 14 the NLU.
11:58:41 15 Q. And if we can go to the next slide. And how does
11:58:50 16 this support your opinion regarding whether or not the Alexa
11:58:55 17 Shopping system has a conversational language processor
11:59:00 18 coupled to the speech recognition engine?
11:59:02 19 A. So this is from the technical document we have seen
11:59:07 20 before, the coupling to the speech recognition system is via
11:59:13 21 this thing on the left labeled SIRF, S-I-R-F, that's what's
11:59:20 22 coming from the speech recognition, and goes into this
11:59:24 23 recognition engine within the NLU.
11:59:26 24 Q. And so there is a recognition engine within the NLU
11:59:30 25 engine of Alexa?

11:59:33 A. That's right.
11:59:34 Q. If could go to now, the next element. Can you expand
11:59:42 that, interpreting, interpret the recognized. The next
11:59:50 element recites "interpret the recognized words or phrases,
11:59:54 wherein the interpreting the recognized words or phrases
11:59:57 includes establishing a context for the natural language
12:00:00 utterance." Is that met by Alexa Shopping?
12:00:03 A. It is.
12:00:04 Q. And what's the basis for your opinion?
12:00:05 A. The technical documents.
12:00:07 Q. And how do those documents support this opinion?
12:00:11 A. So, the interpretation of the words and phrases is
12:00:16 what we've seen before, the coming up with an intent and
12:00:25 slots based on the words that are presented.
12:00:29 Q. And is that supported by the technical documents such
12:00:32 as the one shown on the screen?
12:00:34 A. Yes.
12:00:35 Q. And how so?
12:00:36 A. So the output of the NLU which we've seen other
12:00:40 places is, I believe it's called a BIF, which is this
12:00:44 collection of intents that are -- that establish the context
12:00:53 for those words.
12:00:54 Q. So the output of the NLU is intent; correct?
12:00:58 A. Yes.
12:00:58 Q. And that establishes context?

12:01:00 A. Yes.
12:01:05 Q. Can you go to the next element?
12:01:10 Do you see this element recites select an
12:01:16 advertisement in the context established for the natural
12:01:19 language utterance. Do you see that?
12:01:22 A. Yes.
12:01:22 Q. Is that found in Alexa Shopping?
12:01:26 A. It is.
12:01:26 Q. And what's the basis for your opinion?
12:01:28 A. The technical documents.
12:01:30 Q. And how do they support your opinion?
12:01:33 A. So what happens is the user says that they want to
12:01:40 shop for something, that is -- that's recognized, and that
12:01:46 information is sent to a shopping speechlet which is a piece
12:01:49 of application code that is connected to the NLU, and it
12:01:55 selects a device or selects a product which is then
12:02:00 presented. So the example that I gave in my demo was that I
12:02:05 said "Alexa, I want to buy an iPhone case," and the
12:02:09 speechlet went off, shopping speechlet went off and figured
12:02:14 out that it wanted to sell me this case.
12:02:17 Q. And stay on this slide, stay on this slide, but I
12:02:21 notice here that this also has the next element recited in
12:02:24 the claim, press a selected advertisement via an output
12:02:28 device, do you see that?
12:02:29 A. Yes.

12:02:30 Q. **Is that found in Alexa?**
12:02:31 A. **Yes, it is.**
12:02:32 Q. **And do you base that opinion on the technical**
12:02:36 **documents you reviewed?**
12:02:36 A. **Yes, it's based on the technical documents and my own**
12:02:40 **use of it.**
12:02:40 Q. **And in the example here, did the Echo device present**
12:02:44 **the advertisements?**
12:02:45 A. **Yes, it did.**
12:02:48 Q. **If we could now go back to the language of Claim 40,**
12:02:53 **the last element. Go to slide 26. The last element of**
12:03:09 **Claim 40 recites "an adaptive misrecognition engine**
12:03:14 **configured to determine that the conversational language**
12:03:18 **incorrectly interpreted the words or phrases in response to**
12:03:22 **detecting a predetermined event, wherein the conversational**
12:03:28 **language processor reinterprets the words or phrases in**
12:03:30 **response to the predetermined event." Do you see that?**
12:03:33 A. **Yes.**
12:03:34 Q. **Is that element practiced by Alexa Shopping?**
12:03:37 A. **Yes, it is.**
12:03:38 Q. **And what's the basis of that opinion?**
12:03:40 A. **The technical documents.**
12:03:42 Q. **And how do the technical document that you reviewed**
12:03:46 **support that opinion?**
12:03:47 A. **So, we've seen that Alexa can get information that**

12:03:54 1 causes it to rescore, or to reevaluate what -- which intent
12:04:02 2 is the right one. So that can be done based upon any number
12:04:07 3 of signals that are passed into it to change the scoring of
12:04:14 4 the intent.
12:04:15 5 Q. If we could have Exhibit 226, page 22 on the screen.
12:04:22 6 This is the NLU deep dive document that we were discussing.
12:04:33 7 And does re-ranking support your opinion with regard to the
12:04:39 8 last element?
12:04:39 9 A. Yes, so this re-ranking could happen in response to
12:04:45 10 getting some information that indicated that you had gotten
12:04:49 11 it wrong, that the top scoring intent was wrong, and then it
12:04:54 12 would re-rank it and try a different intent.
12:04:57 13 Q. And one of the events that could cause that, an input
12:05:01 14 would be context, correct?
12:05:03 15 A. Yes, that's one of the inputs listed here.
12:05:07 16 Q. All right. Let's now turn to the '097 patent.
12:05:20 17 Looking at Claim 23 of the '097 patent, starting with the
12:05:24 18 preamble, do you see that this recites a system providing
12:05:31 19 natural language processing based on advertisements, the
12:05:34 20 system comprising. Do you see that?
12:05:36 21 A. Yes.
12:05:37 22 Q. Is that present in Alexa Shopping?
12:05:39 23 A. Yes, it is.
12:05:40 24 Q. And what's the basis for that?
12:05:42 25 A. The documents and my own use.

12:05:44 Q. And if we could go to the next element. You see that
12:05:59 it says one or more physical processors programmed with the
12:06:02 computer program instructions, which when executed, cause
12:06:05 one or more physical processors to, do you see that?
12:06:08 A. Yes.
12:06:09 Q. Is that met by Alexa Shopping?
12:06:10 A. Yes, it is.
12:06:11 Q. Why is it met by Alexa Shopping?
12:06:14 A. Well, the Alexa voice system is running on servers,
12:06:17 the servers have processors, those processors execute
12:06:23 instructions, so it has to be running on a server somewhere,
12:06:27 it's a physical server.
12:06:29 Q. And looking at the exhibit here from DTX-586, do you
12:06:34 see there is a reference there to Alexa voice services?
12:06:38 A. Yes.
12:06:39 Q. Do you see at the bottom it says AWS?
12:06:41 A. Yes.
12:06:42 Q. What do you understand AWS to be?
12:06:44 A. AWS is Amazon's, is the cloud hosting service for
12:06:50 hosting servers, I use it as part of my business all the
12:06:54 time. It's an enormous infrastructure that supports many,
12:06:59 many servers.
12:07:00 Q. Is it fair to say that those many servers have many
12:07:04 physical processors?
12:07:04 A. Yes.

12:07:05 Q. And those processors are programmed with the computer
12:07:09 instruction; correct?
12:07:11 A. Yes.
12:07:11 Q. Can we go to the next element? This one recites
12:07:17 "provide an advertisement associated with a product or
12:07:20 service for presentation to a user." Is that element
12:07:25 present in Alexa Shopping?
12:07:27 A. Yes, it is.
12:07:27 Q. And why is it?
12:07:31 A. From the technical documents and from my own use of
12:07:33 it.
12:07:35 Q. And how did your own use illustrate providing an
12:07:41 advertisement associated with a product or server for
12:07:44 presentation to the user?
12:07:45 A. When I asked Alexa for an iPhone case, it came up
12:07:50 with one to offer me. It offered it to me, it presented it,
12:07:54 it gave me its price. That offering is an advertisement,
12:08:01 certainly as it was discussed in the patent.
12:08:03 Q. And what the Alexa advertisement offered is Amazon's
12:08:13 choice is a Lago compatible with iPhone 13 Pro case liquid
12:08:18 silicone case, right?
12:08:20 A. That's right.
12:08:21 Q. And that was selected by Alexa Shopping, correct?
12:08:26 A. Yes, that was selected by Alexa Shopping, I had
12:08:30 nothing to do with choosing it.

12:08:35 Q. Can we go to the next element? Does Alexa Shopping
12:08:38 receive the natural language utterance of a user?
12:08:40 A. Yes, it does.
12:08:42 Q. And that was illustrated when you spoke to Alexa to
12:08:45 order the iPhone case, right?
12:08:47 A. That's right.
12:08:48 Q. Can we go to the slide for that?
12:08:51 And so in the example here, was colored in
12:08:57 white, Alexa what color or what color is it, or the previous
12:09:02 one about I want to buy an iPhone case, those were
12:09:06 utterances, right?
12:09:07 A. Correct, those were all natural language utterances
12:09:11 of the user.
12:09:12 Q. If we go to the next element. The last element says
12:09:25 "interpret the natural language utterance based on the
12:09:28 advertisement and responsive to the existence of a pronoun
12:09:31 in the natural language utterance, determine whether the
12:09:35 pronoun refers to one or more of the product or service or a
12:09:40 provider of the product or service." Do you see that?
12:09:42 A. Yes.
12:09:42 Q. Is that element present in Alexa Shopping?
12:09:47 A. Yes, it is.
12:09:48 Q. And why is it present?
12:09:49 A. Based upon technical documents and my own use of it.
12:09:53 Q. If we go to the next slide.

12:09:55 And how did your communication with Alexa
12:10:05 Shopping indicate that this element was present?
12:10:08 A. So, when I asked it -- so it had just offered me this
12:10:12 iPhone case, and then I asked it what color, and it came
12:10:17 back with the definition of the word peach. So it clearly
12:10:25 didn't know what I was talking about because the iPhone case
12:10:28 isn't peach. But when I later said, Alexa what color is it,
12:10:34 it being the pronoun, it recognized that I was referring to
12:10:39 the color of the iPhone case, and it correctly said the
12:10:43 color is stone. So there it's clear that Alexa is using the
12:10:51 presence of the pronoun to understand that I was referring
12:10:56 to the product in the advertisement and not just some random
12:11:01 color.
12:11:01 Q. And turning to now the last of our patents, the '703
12:11:05 patent. On the screen is the language of Claim 25 of the
12:11:09 '703 patent, okay?
12:11:12 A. Okay.
12:11:12 Q. And is it your opinion that Alexa Shopping infringes
12:11:18 Claim 25 of the '703 patent?
12:11:21 A. Yes.
12:11:22 Q. And let's focus on, as we have been, element by
12:11:25 element. Can we have the first one? "A system for
12:11:31 providing voice commerce, the system comprising." Is Alexa
12:11:36 Shopping a system for providing voice commerce?
12:11:38 A. Yes, it is.

12:11:39 Q. Can we go to the next element?
12:11:42 This element recites, "one or more physical
12:11:49 processors programmed with computer program instructions
12:11:52 which, when executed, cause the one or more physical
12:11:56 processors to." Do you see that?
12:11:58 A. Yes.
12:11:58 Q. Is that element present in Alexa Shopping?
12:12:02 A. Yes, it is.
12:12:03 Q. And why is it?
12:12:04 A. Same as we had before, these are -- this is running
12:12:08 on servers, the servers are running in AWS which has
12:12:14 physical processors, so there are physical processors
12:12:17 running the computer programs embodying Alexa Shopping.
12:12:21 Q. Going to the next element. Receive a user input
12:12:30 comprising a natural language utterance. Does Alexa receive
12:12:36 a user input comprising a natural language utterance?
12:12:39 A. Yes.
12:12:42 Q. Can we go to the next slide. Did your demonstration
12:12:45 illustrate how Alexa receives a user input comprising a
12:12:51 natural language utterance?
12:12:52 A. Yeah, certainly when I said, "Alexa I want to buy an
12:12:55 iPhone case," that was the user input was natural language.
12:13:00 Q. Can you go to the next element. This element
12:13:06 requires "provide, without further user input after the
12:13:10 receipt of the user input, the natural language utterance as

12:13:13 an input to a speech recognition engine."

12:13:18 Is that element present in Alexa Shopping?

12:13:21 A. Yes, it is.

12:13:22 Q. And why is it?

12:13:23 A. Based upon the technical documents and my own use.

12:13:26 Q. And how do the technical documents illustrate that it

12:13:32 would provide without further user input after the receipt

12:13:36 of user input the natural language utterance that's an input

12:13:41 to the speech recognition engine?

12:13:47 A. What goes on is I speak to the Echo, the Echo

12:13:51 receives it, and sends it straight to the ASR, there is no

12:13:56 further action I have to take, I just speak and it goes.

12:13:59 Q. If we go to the next element. Obtain without further

12:14:05 user input after the receipt of the user input, one or more

12:14:09 words or phrases recognized from the natural language

12:14:13 utterance as an output of the speech recognition engine. Is

12:14:18 that element present in Alexa Shopping?

12:14:21 A. Yes.

12:14:21 Q. And why is it present?

12:14:23 A. Technical documents and the fact that the output of

12:14:33 the ASR goes straight to the NLU without my having to do

12:14:40 anything further.

12:14:41 Q. So the natural utterance that was said is received by

12:14:45 the ASR?

12:14:46 A. Yes.

12:14:47 Q. And the ASR will provide it to the NLU?
12:14:50 A. Yes.
12:14:51 Q. And will it require any additional input to do that?
12:14:55 A. No.
12:14:56 Q. If we go to the next element. This element is
12:15:03 "determine, without further user input after the receipt of
12:15:06 the user input, a context based at least on the one or more
12:15:11 words or phrases." Do you see that?
12:15:13 A. Yes.
12:15:14 Q. Is that element met by Alexa Shopping?
12:15:17 A. Yes, it is.
12:15:18 Q. And why is it met?
12:15:19 A. It's met because the NLU comes up with an intent, and
12:15:25 that intent is part of the context. And that's done without
12:15:31 any further input from the user.
12:15:33 Q. And do the Amazon technical documents support the
12:15:37 fact that it's a context based on one or more words or
12:15:42 phrases?
12:15:42 A. Yes, that's -- we have seen it in technical documents
12:15:46 so far including PTX-226.
12:15:48 Q. And one of the things that the Alexa NLU has is a
12:15:52 context interpreter?
12:15:54 A. That's right.
12:15:56 Q. Is that illustrated in the red box here from PTX-226?
12:16:03 A. Yes the context interpreter is there, it's a 2 step

12:16:08 1 approach that combines context information from different
12:16:10 2 sources.
12:16:20 3 Q. Let's go to the next element, identify without
12:16:24 4 further user input after the receipt of the user input, the
12:16:27 5 product or service to be purchased on behalf of the user
12:16:30 6 based at least on the determined context."
12:16:33 7 A. Yes.
12:16:34 8 Q. Is that element practiced by Alexa Shopping?
12:16:37 9 A. Yes.
12:16:37 10 Q. And why is it practiced?
12:16:39 11 A. Well, we've seen it in documents and you could see it
12:16:43 12 in my use of it that it found an iPhone case to offer me
12:16:49 13 without my having to do anything further.
12:16:52 14 Q. If we go to the next slide. And how is that
12:16:55 15 illustrated in your video demonstration of use with Alexa?
12:16:59 16 A. Yeah, so I asked for an iPhone case, and it without
12:17:03 17 any further queries from me, just came up with that
12:17:07 18 particular offer.
12:17:07 19 Q. Which is the Lago compatible with iPhone 13 proliquid
12:17:14 20 silicone case?
12:17:16 21 A. Yes.
12:17:16 22 Q. If we go to the next element here. "Obtain payment
12:17:22 23 information with which to pay for the product or service."
12:17:26 24 Do you see that?
12:17:27 25 A. Yes.

12:17:27 Q. **Is that element met by Alexa Shopping?**
12:17:31 A. **Yes.**
12:17:32 Q. **And was that illustrated in your demonstration?**
12:17:36 A. **Yes, it was.**
12:17:37 Q. **Could we go to the next slide. Also here the next**
12:17:50 **element is "obtain, without further user input after the**
12:17:54 **receipt of the user input, shipping information with which**
12:17:57 **to deliver the product or service, wherein the shipping**
12:18:00 **information specifies a name or address of a recipient to**
12:18:05 **which the product or service is to be delivered after the**
12:18:08 **product or service is purchased." Is that element present**
12:18:12 **in Alexa Shopping?**
12:18:15 A. **Yes.**
12:18:15 Q. **And why is it present?**
12:18:16 A. **Well, it's present from my own use of it, it's also**
12:18:21 **in documents that are discussed.**
12:18:23 Q. **Next slide. Let me just ask you. Sir, in your**
12:18:38 **example of the video, when you said buy it, did Alexa**
12:18:43 **Shopping ask for your credit card or purchasing information?**
12:18:47 A. **No, it didn't.**
12:18:48 Q. **Did Alexa Shopping ask for your address?**
12:18:52 A. **No.**
12:18:52 Q. **Did it, when you said buy it, ship the case to you**
12:18:57 **without requiring any input with respect to either financial**
12:19:02 **information or address information?**

12:19:06 A. That's right, it did it without any of that
12:19:08 information being given by me at the time.
12:19:12 Q. Going to the next element, "provide, without further
12:19:15 user input after the receipt of the user input, a request
12:19:19 for user confirmation to use the payment information and the
12:19:23 shipping information for a purchase transaction for the
12:19:26 product or service."
12:19:29 Is that element met by Alexa Shopping?
12:19:34 A. Yes.
12:19:35 Q. And what's the basis of that opinion?
12:19:37 A. My own use of it and their documents.
12:19:42 Q. Now looking at the last three elements that we have
12:19:45 been talking about, but in particular, provide without
12:19:49 further user input after the receipt of the user input a
12:19:52 request for user confirmation to use the payment information
12:19:56 and the shipping information for a purchase transaction for
12:20:00 the product or service, was that illustrated in your video
12:20:04 demonstration?
12:20:04 A. Yes. It provided -- it provided -- it provided the
12:20:13 information about what was going on.
12:20:15 Q. So when you said, "Alexa, buy it now," what happened?
12:20:20 A. It made a beep tone, and said that it was -- that it
12:20:28 was placing the order and it will be shipped to me and what
12:20:31 the costs were.
12:20:32 Q. And did it indicate that it would be delivered to New

12:20:35 1 York, 10025?
12:20:37 2 A. Yes.
12:20:37 3 Q. That's your address?
12:20:38 4 A. That's my Zip code, yes.
12:20:40 5 Q. And that, did it indicate that the total including
12:20:45 6 tax after $0.65 in discounts and credits, was $13.44?
12:20:52 7 A. Yes, that's what it said.
12:20:53 8 Q. Was that charged to your credit card without any
12:20:56 9 additional input?
12:20:57 10 A. Yes.
12:20:57 11 Q. Were all the elements of Claim 25 met by Alexa
12:21:01 12 Shopping?
12:21:01 13 A. Yes.
12:21:16 14 Q. Just a few more questions, Dr. Polish. In this case,
12:21:22 15 did you review the expert report of Amazon expert
12:21:27 16 Dr. Johnson?
12:21:27 17 A. Yes.
12:21:28 18 Q. And did Dr. Johnson offer an opinion regarding
12:21:32 19 possible non-infringing alternatives to the VoiceBox patent?
12:21:36 20 A. Yes, he did.
12:21:37 21 Q. And what impact did these non-infringing alternatives
12:21:41 22 have on the number of turns that a user would have to use to
12:21:47 23 perform the operations without practicing the patents?
12:21:50 24 A. In Dr. Johnson's view, the number of turns would
12:21:56 25 increase.

12:21:57 Q. And would that increase -- let me ask this first.
12:22:08 Would it increase by approximately how much, at least?
12:22:13 A. I don't remember the exact number, but it was on the
12:22:16 order of, you know, of 2 or 3 number of turns, it may have
12:22:22 been one-and-a-half turns, something like that.
12:22:25 Q. And do the VoiceBox patents reduce the number of
12:22:28 turns that a user needs to make?
12:22:30 A. Yes, they will.
12:22:32 Q. And does that streamline the user's interface with
12:22:35 respect to how they interact with Alexa?
12:22:40 A. Yes, the whole point is to reduce the amount of
12:22:44 effort the user has to go through.
12:22:45 Q. Would the patents in this case reduce the friction
12:22:49 between the user and the system performing it?
12:22:54 A. Yes.
12:22:55 MR. YOON: Your Honor, at this time, we would
12:22:56 also like to admit Exhibit 586.
12:23:00 MR. HADDEN: I don't recall what that is.
12:23:04 MR. YOON: That's the DTX-586 document. I'm
12:23:08 sorry. It's the one in the slide.
12:23:14 MR. BELGAM: Your Honor, can we confer?
12:23:21 MR. HADDEN: No objection, Your Honor.
12:23:23 THE COURT: Thank you. It is admitted.
12:23:23 (DTX Exhibit No. 586 was admitted into
12:23:24 evidence.)

12:23:24 MR. YOON: Thank you, Dr. Polish. No questions.
12:23:27 THE COURT: Did he talk about that document?
12:23:30 MR. YOON: Those were the documents in all the
12:23:32 slides, it was the slide, each slide that said PDX, I forgot
12:23:35 to mention that, he did talk about, that's the main, if we
12:23:40 can call up one of the slides, he repeatedly talked about
12:23:43 those documents.
12:23:44 THE COURT: I just want to make sure --
12:23:46 MR. YOON: He did.
12:23:47 THE COURT: All right. Thank you. How about
12:23:49 before cross we take our lunch break.
12:23:51 MR. HADDEN: Sure.
12:23:51 THE COURT: Let's take our lunch break. We'll
12:23:53 come back after, a little bit after 1:00.
12:23:58 COURTROOM DEPUTY: All rise.
12:23:59 (Jury leaving the courtroom at 12:23 p.m.)
12:24:21 THE COURT: All right. You guys can be seated.
12:24:33 And sir, you can step down.
12:24:36 Claim construction, is there going to be an
12:24:38 issue on that? I didn't hear Dr. Polish talk about it, so
12:24:43 I'm just curious.
12:24:45 MR. HADDEN: Yeah, I think we can avoid that
12:24:47 issue at this point. He kind of pointed to something but he
12:24:53 didn't redefine it.
12:24:54 THE COURT: And there was no objection, so I

12:24:56 1 didn't hear that, but if you think there is, then you're
12:24:59 2 going to have to tell me so I can deal with it. All right.
12:25:03 3 COURTROOM DEPUTY: All rise.
12:25:04 4 (A luncheon recess was taken.)
13:08:31 5 COURTROOM DEPUTY: All rise.
13:08:32 6 MR. YOON: I just have one brief procedural
13:08:43 7 thing. Your Honor, the next witness after this one will be
13:08:47 8 Mr. Peck, and Mr. Peck is the source code witness. Your
13:08:50 9 Honor, the protective order, the parties agreed we wouldn't
13:08:52 10 be making copies so for Exhibit 541(a) and 541(b), Mr. Smith
13:08:59 11 will be using the Elmo but the witness won't be given a
13:09:03 12 copy.
13:09:03 13 THE COURT: How are you going to give it to the
13:09:05 14 jury when they deliberate?
13:09:07 15 MR. YOON: We'll submit it to the jury when they
13:09:10 16 deliberate, there is no objection to being submitted, but
13:09:13 17 the order limited the number of copies, but it will be
13:09:18 18 submitted to jury, there is no dispute there, it's just two
13:09:21 19 copies.
13:09:21 20 THE COURT: You guys agree?
13:09:24 21 MR. HADDEN: Fine, Your Honor.
13:09:31 22 (Jury entering the courtroom at 1:09 p.m.)
13:09:39 23 THE COURT: All right. Everyone, welcome back.
13:09:48 24 Please proceed.
13:09:52 25 CROSS-EXAMINATION

13:09:52 1 BY MR. HADDEN:
13:09:53 2 Q. Good afternoon, Dr. Polish.
13:09:55 3 A. Good afternoon.
13:09:56 4 Q. Now the Alexa NLU or Natural Language Understanding,
13:10:01 5 uses machine learned neural networks, correct?
13:10:08 6 A. There are some neural networks components of the NLU.
13:10:12 7 Q. And those neural networks are domain models, those
13:10:18 8 are machine learned DNN's and they're used to classify the
13:10:22 9 user's intent, correct?
13:10:24 10 A. That's part of what the NLU does, it also does other
13:10:28 11 things as I talked about.
13:10:29 12 Q. But you agree that the domain models are machine
13:10:34 13 learned DNN's used to classify the user's intent, correct?
13:10:40 14 A. As I said, there are machine learned models within
13:10:45 15 the NLU that are also rules and other things used to
13:10:50 16 classify intent.
13:10:51 17 Q. Well, you prepared an expert report in this case that
13:10:55 18 included your opinions about how Alexa worked and your
13:10:59 19 opinions on infringement, correct?
13:11:01 20 A. Yes.
13:11:02 21 Q. Okay. And you have that in your binder in front of
13:11:05 22 you. Let's look at your expert report. Mr. Patterson, can
13:11:12 23 we look at Dr. Polish's opening report at paragraph 90. On
13:11:23 24 page 31?
13:11:24 25 A. Tell me which tab it is in this.

13:11:28 Q. You should have a tab for your report, pretty thick,
13:11:32 probably the thickest thing in your cross binder. Do you
13:11:36 have that? There should be two binders there.
13:11:58 A. Yes, there is another binder labeled -- oh, and
13:12:03 expert reports. Okay.
13:12:08 Q. It's on page 31 of your opening report. If you look
13:12:31 at paragraph 90, you start by saying the domain models are
13:12:36 machine-learned DNN's used to classify the user's intent,
13:12:40 such as does the customer want to play music, get the
13:12:44 weather, or shop, do you see that?
13:12:46 A. Yes.
13:12:47 Q. And that's accurate, right?
13:12:55 A. Yes, for the -- yes.
13:12:58 Q. And you don't say anything in here about rules, you
13:13:02 talk about machine-learned DNN's, right, in fact, if we go
13:13:17 up to paragraph 88, Mr. Patterson. Highlight that sentence.
13:13:26 You said in your expert report when you're describing
13:13:30 Alexa's NLU, the first step in the 1P pipeline, "X-domain
13:13:40 recognition and re-ranking" refers to running all the 1P
13:13:44 domain models in parallel and running them through a
13:13:47 re-ranking model that merges all results. Do you see that?
13:13:50 A. I see that.
13:13:51 Q. When you talk about 1P models, those are those deep
13:13:57 neural networks, those machine learning neural networks,
13:14:00 right?

13:14:01 A. Just give me one second to review the rest of this
13:14:06 section in the report here.
13:14:36 Q. You wrote this report, right?
13:14:38 A. Yes.
13:14:44 Q. Okay.
13:14:45 A. So the section we're talking about refers to one
13:14:50 particular pipeline through the NLU.
13:14:52 Q. Right, and that's the 1P pipeline, that's the
13:14:54 pipeline that is used to determine whether the user's
13:15:01 request relates to one of the functionalities of Alexa that
13:15:05 Amazon itself provides rather than a third-party skill,
13:15:10 right, that's why it's called 1P, correct?
13:15:14 A. It's one of the pipelines that the NLU runs, it's
13:15:20 part of what the NLU does.
13:15:21 Q. Well, let's be clear here. 1P refers to it being a
13:15:26 pipeline for the Amazon speechlet, like shopping or weather,
13:15:35 or alarms, and the other examples that you used that you
13:15:37 just talked about with VB Assets's counsel, right? Let me
13:15:47 ask a clearer question, maybe. Every example that you
13:15:51 talked about earlier today, those are all examples of these
13:15:56 domain models running in the 1P pipeline, those generated
13:16:04 intents, right?
13:16:04 A. What I talked about with counsel this morning was
13:16:09 about generating intents within the NLU.
13:16:12 Q. Right. And every one of these examples you used,

13:16:15 1 those intents were generated by these deep neural networks
13:16:19 2 that are run in the 1P pipeline; right? Do you not know how
13:17:05 3 the intents were generated that you talked about all
13:17:08 4 morning?
13:17:08 5 A. I know that the intents were generated within the
13:17:11 6 NLU. I don't know -- what I'm not clear about is whether
13:17:15 7 they were part of the 1P pipeline in particular.
13:17:19 8 Q. Okay. So you don't know other than that they came
13:17:21 9 out of the NLU box, how they were actually generated?
13:17:24 10 A. Well, we talked about they were generated based upon
13:17:28 11 the document we had from Mr. Vanee, that they were generated
13:17:33 12 by executing a variety of different pipelines, including
13:17:41 13 ones that contained rules including ones that contained
13:17:47 14 finite data machines including one that contained models.
13:17:48 15 Q. You don't know whether these are generated by the 1P
13:17:52 16 models that you talk about in your report?
13:17:54 17 A. I don't recall whether the -- whether what was in
13:17:59 18 Mr. Vanee's diagram was entirely part of 1P or not entirely
13:18:05 19 part of 1P.
13:18:06 20 Q. So in your report here, you claim that the
13:18:13 21 information that's in paragraph 90, the information in a
13:18:19 22 SIRF, is passed to all the domain models, roughly 37 to
13:18:22 23 produce intents and token labels, do you see that?
13:18:29 24 A. Yes.
13:18:29 25 Q. And those 37 domain models, those are 37 of those

13:18:34 1 machine learned deep neural networks, right?
13:18:36 2 A. Yes, that's my understanding.
13:18:37 3 Q. And each of those domain models produces multiple
13:18:42 4 results, including intent classification and name entity
13:18:48 5 recognition; right?
13:18:50 6 A. Yes.
13:18:51 7 Q. And then the Alexa NLU -- before I get there, yeah,
13:19:02 8 and the Alexa NLU sorts those results based on confidence
13:19:10 9 scores and retains the NBest to be passed along to
13:19:18 10 downstream processing.
13:19:20 11 A. Yes.
13:19:22 12 Q. Yes, and the output of the NLU retains multiple
13:19:26 13 interpretations of a customer's request right?
13:19:29 14 A. Yes, the different intents are different
13:19:31 15 interpretations.
13:19:37 16 Q. At least what you described in your report is what I
13:19:41 17 showed in the opening, right, so we have this NLU and it
13:19:46 18 receives the words from ASR, that's what's in that SIRF,
13:19:53 19 right?
13:19:54 20 A. The SIRF contains the recognized words.
13:19:57 21 Q. Right, that's the output of ASR, right?
13:20:00 22 A. The SIRF is the output of the ASR.
13:20:03 23 Q. And those words, that's what is provided to these
13:20:07 24 deep neural network domain models, and here I only show six,
13:20:13 25 but as you explained in your report, there is actually 37,

13:20:17 1 right?
13:20:17 2 A. Yes.
13:20:18 3 Q. And those guys receive whatever the words are, right,
13:20:23 4 and they process those neural networks, and each one comes
13:20:29 5 up with its own meaning, right?
13:20:34 6 A. So you used -- so first of all, the implication here
13:20:38 7 is that the entirety of the NLU is of these -- are these
13:20:43 8 DNN's, which it's not, and also you're calling them
13:20:48 9 meanings. They're different intents and associated slots.
13:20:52 10 Q. So let's just follow along with what you said in your
13:20:55 11 expert report. What you describe in your expert report is
13:20:58 12 exactly what I showed on the screen, which is the output of
13:21:01 13 ASR goes through each of the deep neural networks and they
13:21:06 14 all provide at least one, maybe multiple interpretations,
13:21:10 15 correct?
13:21:12 16 A. They each can provide one or more intents and slots.
13:21:19 17 Q. Well you called them interpretations in the expert
13:21:22 18 report, are they interpretations?
13:21:24 19 A. Yes, that's in the form of those intents.
13:21:26 20 Q. Right, so these are interpretations of what the words
13:21:31 21 meant that the user spoke; right?
13:21:38 22 A. I'm cagey about using the word meaning here because I
13:21:42 23 think it implies --
13:21:43 24 Q. Meaning today in the claim, and you talked about
13:21:48 25 determining the meaning of the words that are ambiguous in

13:21:51 an utterance, right?
13:21:51 A. Yes.
13:21:53 Q. And in that use of meaning, you were referring to
13:21:56 exactly this, which are the intents and slots that are
13:21:59 determined by the NLU, right?
13:22:02 A. Broadly speaking, an interpretation of what comes out
13:22:06 of the ASR is an intent and slot.
13:22:12 Q. Right.
13:22:12 A. Yes.
13:22:13 Q. That's what I'm showing exactly here on the diagram
13:22:16 right, you have different meanings for the words that come
13:22:19 out of each of these neural networks, right, and they're
13:22:23 collected in this list; right, that's the NBest words,
13:22:28 correct?
13:22:28 A. This is your slide, this is -- are you saying that
13:22:35 change is the intent and 6:45 is the slot.
13:22:43 Q. I'm using meaning to incorporate both the intent and
13:22:45 the slot, do you understand that?
13:22:47 A. Okay.
13:22:48 Q. Okay, so given that interpretation, this diagram is
13:22:53 correct, right?
13:22:56 A. This diagram is extremely high level and it's -- I
13:23:05 was trying to make reference to actual Amazon documents that
13:23:08 had more technical terminology there. I'm not sure I see
13:23:13 anything wrong with this diagram, but it's not -- I'm not

13:23:16 adopting it as the truth.
13:23:18 Q. One important thing about this diagram and also when
13:23:21 you say in your expert report is that the Alexa NLU runs all
13:23:28 of these deep neural networks, 37 in your report. And they
13:23:34 all do their job and come up with their interpretation
13:23:38 before any context is selected, right?
13:23:44 A. My understanding is that as I testified to this
13:23:51 morning, there are rules -- there are rules approaches taken
13:23:57 that would complete before these NLU -- these neural
13:24:03 networks would finish processing.
13:24:05 Q. I understand that. But the question is, if we take
13:24:10 music as an example of a context, Alexa doesn't determine
13:24:15 that the user's request had to do with music before it runs
13:24:21 all 37 neural networks and comes up with all the different
13:24:24 meanings, right?
13:24:26 A. Well, there is more to the context than just -- than
13:24:31 just this one output. In a multi-turn environment, you're
13:24:35 going to have, as I showed, you're going to have prior --
13:24:38 you're going to have prior -- prior intents as the current
13:24:42 context.
13:24:43 Q. Let's just focus on what we're talking about now.
13:24:47 Speech comes into through the ASR, the words are recognized,
13:24:52 every utterance goes through this process, and we create a
13:24:56 list of meanings before Alexa has picked which of these
13:25:00 contexts is the right one; correct?

13:25:03 1 A. Well, also going into these are the current contexts
13:25:08 2 and the information about the prior -- the prior dialogue
13:25:16 3 acts.
13:25:16 4 Q. That's not true, you know that's not true. The input
13:25:19 5 to the deep neural network, that's only the words that come
13:25:23 6 out of the ASR, right?
13:25:25 7 A. I don't know if that's solely the inputs.
13:25:27 8 Q. You don't know how these work, right? Do you know
13:25:32 9 how these work or not?
13:25:34 10 A. My understanding is that the deep neural networks
13:25:37 11 take into account contexts, as well as the speech that comes
13:25:42 12 out.
13:25:42 13 Q. So you believe that when the NLU processes the words
13:25:48 14 from SIRF, it somehow has other inputs?
13:25:52 15 A. My understanding is that the NLU takes into account
13:25:57 16 context.
13:25:58 17 Q. Let's be more precise than that. We're not talking
13:26:04 18 about everything that's done in the NLU, I'm talking about
13:26:06 19 the processing of the words by these deep neural networks
13:26:10 20 that you talk about in your report. What they receive is
13:26:13 21 only the words from the utterance, correct?
13:26:16 22 A. I'm looking at the NLU, the NLU takes into account
13:26:20 23 the other things. It may be that there is a portion of the
13:26:23 24 code that doesn't take into account the contexts. But the
13:26:28 25 NLU takes into account the contexts.

13:26:30 Q. Answer my question. My question is, I'm not talking
13:26:34 about the NLU, I'm talking about the deep neural networks
13:26:37 that you described in your report, and it's true fact, isn't
13:26:40 it, that the only thing that they receive as input are the
13:26:44 words from NLU, isn't that a true fact?
13:26:50 A. I don't know whether the particular deep neural
13:26:52 network that you're pointing to just takes in speech or
13:26:56 takes in other things. The NLU takes into other things
13:26:59 beside just the speech.
13:27:00 Q. As far as you know, what I have got on the screen is
13:27:02 absolutely accurate, right?
13:27:04 A. I'm not adopting your slide, as you already said the
13:27:07 NLU does more than just run the deep neural networks.
13:27:11 Q. I'm not asking about that, I'm asking about the
13:27:13 process that you describe in your report where the output of
13:27:17 the ASR is processed by deep neural networks, this is
13:27:21 accurate, right?
13:27:25 A. The deep neural -- I don't know whether the deep
13:27:28 neural networks use more than just speech.
13:27:30 Q. You don't know what they do because you actually
13:27:35 didn't look at any of Alexa source code in this case, did
13:27:40 you?
13:27:40 A. Well, yes, I certainly did look at source code.
13:27:44 Q. You did? Do you recall being deposed in this case?
13:27:47 A. I do.

13:27:48 1 Q. Let's look in your deposition. Let's look at page 7,
13:28:02 2 lines 1 through 4. Mr. Patterson, if you could pull up a
13:28:15 3 clip, if you have it. Did you find that in your deposition,
13:28:23 4 sir?
13:28:23 5 A. Yes.
13:28:25 6 Q. Can we play the clip?
13:28:26 7 (Videotape deposition:)
13:28:31 8 Q. Is it correct that you did not review any of the
13:28:34 9 Alexa source code that Amazon made available in this case?
13:28:38 10 A. That is correct.
13:28:39 11 MR. YOON: Objection, Your Honor.
13:28:41 12 THE COURT: Go ahead.
13:28:42 13 MR. YOON: I understood, Your Honor, under the
13:28:45 14 procedure they're not supposed to be able to look at it.
13:28:48 15 THE COURT: We don't play it up there, that's
13:28:51 16 not even impeachment. Why don't you ask him if he read it
13:28:55 17 and looked at it, tell him what it says, but don't play it.
13:28:59 18 MR. HADDEN: Okay, Your Honor.
13:29:00 19 THE COURT: You did not ask him any questions I
13:29:03 20 just looked at it --
13:29:05 21 BY MR. HADDEN:
13:29:05 22 Q. Do you see your testimony from your deposition on
13:29:08 23 page 7, line 1?
13:29:09 24 A. Yes.
13:29:09 25 Q. Were you asked the question:

13:29:10 "Question: Is it correct you did not review any
13:29:14 Alexa source code that Amazon made available in this case?"
13:29:16 Is it correct that you answered:
13:29:18 "Answer: This is correct."
13:29:21 A. I said that is correct, yes.
13:29:23 Q. Is that a true statement?
13:29:24 A. Yes, I did not review code. I didn't sit at a source
13:29:28 code review computer and review it.
13:29:31 Q. And is it also true that you relied entirely on
13:29:36 Mr. Peck's analysis of the Alexa source code?
13:29:40 A. I relied -- yes, I relied on Mr. Peck's analysis of
13:29:45 the source code and I looked at the code that was in his
13:29:47 report and that he provided to me.
13:29:51 Q. Is it correct that -- did you direct any of
13:30:03 Mr. Peck's source code review or analysis?
13:30:06 A. Indirectly through the lawyers.
13:30:12 Q. Let's turn to your deposition again, page 8. Do you
13:30:15 have it? Line 5 through 7. Were you asked:
13:30:22 Question: Did you direct any of Mr. Peck's
13:30:25 source code review or analysis?
13:30:27 Did you answer:
13:30:29 Answer: No, I did not.
13:30:34 A. Yes, I said that.
13:30:35 Q. Is that how you testified?
13:30:37 A. That is.

13:30:37 Q. And you didn't tell or ask Mr. Peck to analyze
13:30:42 certain features of the Alexa source code, is that right?
13:30:45 A. Yes, I did not tell Mr. Peck directly to do anything.
13:30:51 I saw the questions that the lawyers were asking him, I
13:30:56 contributed to that, and I saw what was going on through
13:30:59 that interaction.
13:31:00 Q. And you could have reviewed the Alexa source code, it
13:31:05 was all made available to you, correct?
13:31:08 A. Yes, I think so.
13:31:10 Q. But you didn't bother to look at any of that code
13:31:14 that was made available to you?
13:31:16 A. I resist the term bother there, sir. I had somebody
13:31:21 who was looking at the code very diligently and I interacted
13:31:26 through the lawyers with him and had his report. I
13:31:30 certainly took seriously my responsibility to confirm what I
13:31:35 found with the code.
13:31:35 Q. But you had the opportunity to see exactly how what
13:31:39 we saw on the screen work in the source code, and you chose
13:31:43 not to look at it, correct?
13:31:46 A. I did not look at the source code. I did not review
13:31:50 the source code on a source code computer.
13:31:54 Q. Can we put the demonstrative back up.
13:31:56 Now, you agree that the output of NLU is a list
13:32:03 of possible interpretations, right?
13:32:06 A. The output-- my understanding is the output of these

13:32:12 1 models is an NBest list of intents and slots.
13:32:17 2 Q. Right, that's what I have here is this list and
13:32:21 3 meaning, right?
13:32:21 4 A. Okay.
13:32:23 5 Q. Now, if Alexa chose a context for all these DNN's
13:32:31 6 processed the user's words, then my example from the opening
13:32:40 7 of "do I need an umbrella," versus "I need an umbrella,"
13:32:47 8 wouldn't work, right?
13:32:49 9 A. I'm not sure I'm following.
13:32:51 10 Q. Sure. So if Alexa decided before figuring out the
13:32:57 11 meaning of the user's words in all possible contexts, they
13:33:01 12 decided beforehand that it was going to be focused on the
13:33:05 13 shopping context, and we only ran the neural network for
13:33:11 14 shopping, then if I said "Alexa, do I need an umbrella," I
13:33:17 15 would get intent about buying an umbrella, and not a weather
13:33:22 16 forecast, right?
13:33:24 17 A. No, I don't think that's -- I don't think that's
13:33:27 18 right.
13:33:27 19 Q. Well, the weather forecast intent would come from the
13:33:33 20 operation of the weather model, when it operates on the
13:33:39 21 words, right, isn't that how it works?
13:33:42 22 A. So maybe we can re-examine this hypothetical. Give
13:33:47 23 me the example that you're talking about.
13:33:49 24 Q. Sure. I say "Alexa, do I need an umbrella?"
13:33:55 25 A. Okay.

13:33:55 Q. And we agree the words the way it works now, the
13:33:58 words "Do I need an umbrella?", would be provided as input
13:34:03 to all thirty-seven of those neural networks in the 1P
13:34:07 pipeline; correct?
13:34:08 A. Among other things, yes.
13:34:09 Q. And now if instead of doing that, Alexa decided this
13:34:14 is a shopping context, and only provided those words to the
13:34:20 model for shopping, then I would never get a weather intent,
13:34:25 right?
13:34:25 A. No, as I was saying, the NLU takes into account
13:34:30 context, so you could absolutely have all those words
13:34:34 supplied to all these DNN's, they would all put up what they
13:34:38 thought things were and then they were interpreted in the
13:34:42 context of the context. So if you said "do I need an
13:34:45 umbrella," and you were in the middle of discussing weather
13:34:49 with Alexa, there would be a propensity to decide this was
13:34:54 weather, not shopping.
13:34:55 Q. That wasn't my hypothetical. The hypothetical was,
13:34:59 you decide a context before the words are processed by the
13:35:03 DNN, right?
13:35:04 A. Okay.
13:35:05 Q. You do that, you're never going to get a weather
13:35:09 intent, you picked shopping as the context?
13:35:11 A. So you're assuming that having decided on the
13:35:14 context, you would shut down the other DNN's.

13:35:16 Q. Sure.
13:35:18 A. I don't believe that that's what occurs.
13:35:20 Q. I agree that doesn't occur, we'll get into the
13:35:24 claims, but the claims require identifying a context before
13:35:27 you determine anything, right?
13:35:30 A. The claim talks about finding a context. And as I've
13:35:37 talked about, that context can come from a variety of
13:35:41 sources.
13:35:42 Q. Let's look at the claim to make it clear.
13:35:45 Mr. Patterson, could we get PDX3-10? This is Claim 13 of
13:36:06 the '681 patent from the counsel's presentation this
13:36:11 morning. And before we get to the context, if we just blow
13:36:16 up the preamble again. I think as you testified this
13:36:23 morning, this is a claim to computer executable
13:36:28 instructions, right?
13:36:29 A. Yes.
13:36:30 Q. And that's computer code; right?
13:36:34 A. That's part of the preamble, yes.
13:36:37 Q. And so if we want to find out whether this specific
13:36:42 computer code that's required by this claim was in Alexa,
13:36:49 the best place to look for that would be in the Alexa source
13:36:52 code, right?
13:36:55 A. I'm not sure I understand. My understanding is that
13:36:59 there is servers that are running Alexa, and that those
13:37:05 servers are embodied in executable code that's sitting on a

13:37:10 1 server.
13:37:11 2 Q. Sure, but if we're going to understand whether or not
13:37:14 3 Alexa has the specific computer instructions that are
13:37:18 4 required by this claim, the place we should look is in the
13:37:22 5 computer instructions in Alexa's source code; right?
13:37:27 6 A. The computer source code and the computer executable
13:37:31 7 instructions are not the same thing.
13:37:34 8 Q. The computer source code is where it gets compiled to
13:37:38 9 make the executable instructions the computer to run, right?
13:37:42 10 A. There are many steps between the source code and
13:37:45 11 computer executable instructions. The point is, that what
13:37:50 12 Alexa is doing, the Alexa server is running codes sitting on
13:37:57 13 computers and that's what is -- what's being talked about
13:38:01 14 here.
13:38:01 15 Q. Right, and if you're going to look at the code in a
13:38:05 16 form that a human can read it, it would be in the source
13:38:08 17 code, right?
13:38:09 18 A. Well, code in human readable form is source code.
13:38:15 19 But this preamble is not talking about source code.
13:38:19 20 Q. You're telling me that the computer executable
13:38:22 21 instructions in Alexa are not created from source code?
13:38:28 22 A. Well, of course they're created, they're created from
13:38:32 23 source code. What this preamble is talking about is
13:38:35 24 computer readable medium, not human, computer readable and
13:38:40 25 computer executable instructions that are -- that are

13:38:44 1 providing this user interface. So I'm talking about not
13:38:47 2 source code here, but the code that runs the server.
13:38:50 3 Q. Right, and you didn't look at the code that runs the
13:38:54 4 servers, right?
13:38:55 5 A. I did not look at the machine code that was sitting
13:38:58 6 on the servers, no.
13:39:00 7 Q. Let's look at the -- one of the requirements of this
13:39:09 8 claim is receiving an utterance; right?
13:39:17 9 A. Yes.
13:39:18 10 Q. And that utterance then has to be processed by those
13:39:23 11 computer executable instructions in a specific way; right?
13:39:28 12 A. Okay.
13:39:29 13 Q. And if we look at the accumulate short-term shared
13:39:35 14 knowledge element, that has to be knowledge about the
13:39:40 15 utterance. Can you see that? I have it highlighted there,
13:39:50 16 sir, right?
13:39:51 17 A. Well it includes knowledge about the utterance that
13:39:53 18 was received, but the short-term shared knowledge is about
13:39:56 19 the current conversation. So those are -- the conversation
13:40:00 20 can contain multiple utterances and the knowledge, the
13:40:04 21 short-term knowledge is knowledge that goes from utterance
13:40:06 22 to utterance within the same conversation.
13:40:09 23 Q. Right, but it has to be included knowledge about the
13:40:12 24 utterance and that has to be the utterance that was received
13:40:16 25 in the receive an utterance step, right? It has to be the

13:40:21 1 same utterance, right?
13:40:25 2 A. Well, the short-term shared knowledge includes
13:40:28 3 knowledge about the -- it includes other things as well.
13:40:31 4 Q. Sure, but the utterance in this element has to be the
13:40:34 5 same utterance that we receive when we receive an utterance,
13:40:38 6 right?
13:40:40 7 A. Well, it is accumulating short-term knowledge from
13:40:44 8 the current utterance. The short-term knowledge includes
13:40:49 9 knowledge accumulated from prior, that's what it means to be
13:40:52 10 accumulating.
13:40:54 11 Q. Right, this is not that complicated. All I'm trying
13:40:59 12 to confirm is when it says knowledge about the utterance,
13:41:02 13 that has to be the same utterance that's referred to above
13:41:05 14 where it says receive an utterance, right?
13:41:10 15 A. Again, it's not that complicated, short-term
13:41:15 16 knowledge about the current conversation and it's
13:41:17 17 accumulating some knowledge about the current utterance.
13:41:22 18 Q. Right, and then we go down to the identify context
13:41:28 19 step. And we have to identify a context associated with the
13:41:33 20 utterance again, right?
13:41:37 21 A. Yes.
13:41:37 22 Q. And that has to be the utterance that we received
13:41:40 23 above; right?
13:41:42 24 A. Yes, it's the utterance.
13:41:45 25 Q. And if we go to the next element, we have to

13:41:51 1 establish an intended meaning for the utterance, it has to
13:41:56 2 be the same utterance we started with, right? Correct?
13:42:03 3 A. Yes, this is talking about establishing an intended
13:42:07 4 meaning for this particular utterance.
13:42:08 5 Q. Right, and then we have to go down to the last one
13:42:13 6 where we generate a response to the utterance, right?
13:42:17 7 A. Yes.
13:42:18 8 Q. So nowhere this morning did you identify what-strike
13:42:26 9 that. Nowhere this morning did you tell the jury this is
13:42:30 10 the utterance and here is how every one of these steps is
13:42:36 11 performed on that utterance, did you?
13:42:39 12 A. So I reviewed the example of a demonstration and I
13:42:51 13 talked about how utterances were received and processed by
13:42:55 14 the NLU.
13:42:57 15 Q. No, my question is very specific. This claim
13:43:01 16 requires taking a specific utterance and doing these
13:43:06 17 highlighted things to it. You didn't identify an utterance
13:43:09 18 and show how each of these steps is performed on that
13:43:13 19 utterance, did you?
13:43:14 20 A. That is why I walked through my infringement
13:43:18 21 analysis, I walked through my infringement analysis by
13:43:21 22 looking at the documents and about my demonstrated use.
13:43:25 23 Q. Understood. But going back to my football/soccer
13:43:30 24 ball example in the opening, to prove infringement, you have
13:43:33 25 to prove that Alexa receives an utterance, the specific

13:43:38 1 utterance, and that each of these things required, right?
13:43:43 2 A. I have to prove that Alexa does each of the elements
13:43:46 3 of the claims.
13:43:47 4 Q. Right, and each of the elements have to be performed
13:43:51 5 consistently on a single utterance as recited here; right?
13:43:59 6 A. Well, it traces through what happens with an
13:44:03 7 utterance and how knowledge is gained through the utterance
13:44:07 8 and knowledge gained previously is used with the utterance.
13:44:11 9 Q. So what utterance did you show this that each of
13:44:17 10 these steps maps to?
13:44:20 11 A. Well, I showed -- I showed the conversation that I
13:44:26 12 had with Alexa, and I showed the technical documents to show
13:44:29 13 what Alexa is doing.
13:44:31 14 Q. That's not what's required. What's required is you
13:44:34 15 have to identify an utterance and show to the jury that each
13:44:37 16 of these things is gone through. So what is the utterance
13:44:40 17 that you think you showed that for?
13:44:43 18 A. I did not show it that way, I showed what Alexa is
13:44:47 19 doing and -- showed that Alexa is doing each and every thing
13:44:52 20 that is in the claim.
13:44:53 21 Q. It's a fact, isn't it sir, that you can only prove
13:44:57 22 infringement if you prove that every element of this claim
13:44:59 23 was performed; right?
13:45:01 24 A. Yes.
13:45:02 25 Q. Okay. And to perform every element of this claim,

13:45:05 1 you have to have a specific utterance and show that each of
13:45:09 2 these highlighted steps is performed with respect to that
13:45:13 3 utterance; isn't that what is required, isn't that the law?
13:45:17 4 A. I'm not a lawyer. I know that I showed that Alexa
13:45:22 5 was doing each of these things, and I showed documents
13:45:25 6 around it and I showed an example of Alexa in use.
13:45:30 7 Q. Let's be very clear. True fact or not, did you show
13:45:36 8 the specific utterance that is processed in the way required
13:45:39 9 by Claim 13?
13:45:52 10 A. It's possible that my demo example would fit this and
13:45:58 11 could be analyzed that way. That isn't how I presented it.
13:46:03 12 It's possible that my demo example shows receiving an
13:46:08 13 utterance and using -- and accumulating short-term
13:46:11 14 knowledge, identifying the context and establishing intended
13:46:14 15 meaning and generating a response, I didn't analyze it in
13:46:20 16 that way.
13:46:21 17 Q. But you didn't show any utterance that Alexa has ever
13:46:27 18 processed, that's processed as required by the steps of this
13:46:31 19 claim, right?
13:46:32 20 A. I showed a demonstration and I analyzed documents.
13:46:37 21 Q. You're not answering my question. Showing a
13:46:41 22 demonstration doesn't prove anything to the jury. What the
13:46:44 23 jury has to see is that you proved that each step of this
13:46:48 24 claim is performed on a specific utterance so that we can
13:46:51 25 show that, you can show that the claim is met. You didn't

13:46:55 do that, right?
13:46:56 A. I did not prove it by showing one specific utterance.
13:47:00 Q. So let's look at your presentation. And let's go to
13:47:17 the PDX3-12. Now, PDX3-12, you cite to that deep dive
13:47:32 Amazon document on the right. Is that correct?
13:47:36 A. That's correct.
13:47:37 Q. And what you show above, is the output of the NLU
13:47:44 when user says six, right?
13:47:49 A. I don't know if that would be considered the output
13:47:52 of the NLU. It certainly -- it's the NLU has reached a
13:47:56 number of different hypotheses about what six means.
13:48:00 Q. But we all agree that the output of the NLU is this
13:48:06 NBest list which is, it looks like this, right, it's got
13:48:12 intent and slots and multiple contexts, right?
13:48:15 A. Yes, that is true.
13:48:17 Q. And so this has already been done before any context
13:48:24 has been identified in your mapping of the claim, right?
13:48:34 A. So no, there is a context associated with what the
13:48:42 state of the system is already, and this is being
13:48:47 interpreted within the context of that context. And there
13:48:51 is further example of how it processes this two-step
13:48:58 approach that takes into account the context that it had
13:49:01 before, and looks at new utterances and interprets them.
13:49:07 Q. I'm confused now. So let's look at the language that
13:49:11 you highlighted and red in the claim. One or more words

13:49:20 that have different meanings in different contexts. Do you
13:49:22 see that?
13:49:23 A. Yes.
13:49:23 Q. You're saying that this slide proves that, right?
13:49:26 A. I'm showing that this slide -- I'm saying that this
13:49:30 slide shows an example of Amazon considering the case where
13:49:34 the word six has different meanings and different contexts.
13:49:37 Q. And the different contexts you identify are volume
13:49:41 level and song name, or are those a different meaning?
13:49:48 A. So those are different slots within different
13:49:54 intents.
13:49:55 Q. Just to be clear, when you say volume level equals
13:49:59 six, that's one meaning for the word six; correct?
13:50:02 A. Yes.
13:50:03 Q. And time equals six is another meaning for the word
13:50:07 six, right?
13:50:08 A. That's right.
13:50:08 Q. And song name equals six is another meaning for the
13:50:13 word six, right?
13:50:14 A. Yes, because they each have different intents
13:50:18 associated with them.
13:50:19 Q. And those different meanings are in different
13:50:23 contexts, so the different contexts you show here are global
13:50:28 set volume intent, is that a context?
13:50:31 A. That's an intent, these are intents, and it's trying

13:50:35 -- it's going to try to disambiguate, figure out which is
13:50:39 the correct intent.
13:50:42 Q. That wasn't the question, though. What the claim
13:50:45 requires is different meanings and different contexts.
13:50:48 A. Yes.
13:50:48 Q. We've gone over what the different meanings are. So
13:50:51 where are you showing the different contexts on these
13:50:54 slides?
13:50:54 A. Well, those different intents, set volume, multi-turn
13:51:03 notification, and multi-turn music are different -- in the
13:51:08 context of -- in the way of thinking about it in the claim,
13:51:12 those are different contexts.
13:51:13 Q. That was what I said originally. So you agree that
13:51:17 the different -- that set volume is a context, notifications
13:51:22 multi-turn intent is a context, music multi-turn intent is a
13:51:28 context, right?
13:51:30 A. In the sense of the patent, those are different
13:51:32 contexts, yes.
13:51:33 Q. We have three contexts here with a separate meaning
13:51:36 for each of them, right?
13:51:38 A. Yes.
13:51:39 Q. Okay. And we haven't picked a context, right?
13:51:50 A. Well, what do you mean we haven't picked a context?
13:51:55 Q. We'll get to that.
13:51:57 Now, this, what we see on this slide, this is an

13:52:01 1 example of what I showed on my demonstrative of that table
13:52:07 2 with the different meanings; right?
13:52:10 3 A. I'm not sure I know what you're referring to.
13:52:12 4 Q. Well, this is a portion -- right. Here we go. This,
13:52:16 5 what I showed here, what I have changed 6:45, meaning two,
13:52:21 6 meaning three, that is equivalent, if could go back, to what
13:52:25 7 you showed at the top of your slide where it says get set
13:52:33 8 volume intent, that would be one of the outputs from one of
13:52:36 9 the deep neural networks and it would include six, the
13:52:42 10 meaning, the meaning of six is volume level, right?
13:52:46 11 A. I don't know that this is the output of an DNN. This
13:52:51 12 is an output -- this is output from the NLU which as I said
13:52:55 13 does things besides the DNN, this could have been the result
13:53:01 14 of an DNN, but I don't know how it produced this.
13:53:05 15 Q. And the same is true for the music intent song name
13:53:11 16 equals six, so six went into the music DNN and it figured
13:53:16 17 out that that could be the name of a song, right?
13:53:18 18 A. Again, I don't know that it went through the DNN to
13:53:21 19 come up with this, it could have gone in through a rule or
13:53:25 20 gone in through an FST.
13:53:28 21 Q. And the result, though, is that the NLU has already
13:53:37 22 determined the meaning of the word six in each of these
13:53:41 23 possible contexts, right?
13:53:43 24 A. Those different -- those different meanings are
13:53:45 25 hypothetically in an NBest list.

13:53:50 Q. They are meanings for the word six in each of the
13:53:54 possible contexts, right?
13:53:55 A. I think that's what I just said.
13:54:02 Q. Now -- and at this point, we haven't used any
13:54:12 short-term shared knowledge, right?
13:54:17 A. I don't know that that's -- that that's true. The
13:54:22 NLU uses information from prior speech acts and from other
13:54:28 pieces of information.
13:54:30 Q. Well let's look at that. Let's look at PTX-226,
13:54:42 Mr. Patterson. Let's look at page 32. This is that
13:54:57 multi-turn example you talked about, right?
13:55:00 A. Yes.
13:55:01 Q. And in fact, the user is saying six, that's what we
13:55:07 saw on the top of your slide, PDX3-12, right, the result --
13:55:16 that six corresponds to the user saying six in this example,
13:55:20 right?
13:55:20 A. They're both saying six.
13:55:23 Q. Now, if we look in the same document and we go to the
13:55:34 next slide, 33, it explains that when the NLU processes that
13:55:45 word six to come up with those different meanings, it says
13:55:50 explicitly here, NLU models are currently not aware of
13:55:55 dialogue context, right?
13:55:56 A. It says that, yes.
13:55:59 Q. This is saying in fact, when Alexa interprets the
13:56:03 word six, it's not using any of that short-term shared

13:56:08 1 knowledge about what the user had said before; right?
13:56:11 2 A. That's not how I interpret this, I interpret this as
13:56:14 3 saying the NLU model, the statistical models, the DNN's, may
13:56:19 4 not be aware of dialogue contexts but the rules and FST's
13:56:24 5 are.
13:56:25 6 Q. You haven't seen any Alexa source code or
13:56:30 7 documentation at least that you have shown the jury that
13:56:33 8 says that rules are FST's are where dialogue context, have
13:56:40 9 you?
13:56:40 10 A. I haven't shown the jury any source code.
13:56:43 11 Q. Or technical documents that said rules and FST's use
13:56:47 12 dialogue contexts, right?
13:56:49 13 A. That's not true. This document has examples of rules
13:56:55 14 and FST's running within the NLU along with models, and it's
13:57:01 15 showing use of dialogue contexts in this very example. So
13:57:06 16 from that I think it's quite reasonable --
13:57:09 17 Q. Let us walk through this example here to be a little
13:57:12 18 clear.
13:57:13 19 A. Okay.
13:57:13 20 Q. Because what happens in this example, is we run six
13:57:18 21 through the model with no dialogue context, as the Amazon
13:57:29 22 document itself says. We get to the next slide, 34, please,
13:57:34 23 Mr. Patterson. And this is what happens, right? The
13:57:39 24 recognizer generates these intents and slots, these contexts
13:57:46 25 and meaning of words that we just saw in your chart, right?

13:57:51 A. This is showing for each of those utterances six,
13:57:57 yeah, and nope, the three different groups of three
13:58:02 different interpretations of those utterances.
13:58:04 Q. Right, and the one that you used in your chart is the
13:58:07 one for six, the top line, right?
13:58:10 A. That's right.
13:58:11 Q. We just saw in this document that those meanings were
13:58:16 generated without using any dialogue context, right, that's
13:58:21 what the document said?
13:58:22 A. Well, at this point in the process, it's showing you
13:58:26 what we get without dialogue contexts.
13:58:30 Q. Right, so the meanings of the word six have all been
13:58:33 determined without any dialogue contexts, right?
13:58:37 A. I don't know about determined, it's this is what it
13:58:42 is coming up with for each of these utterances just by
13:58:46 themselves.
13:58:46 Q. Right, but the intents and slots, set volume intent
13:58:52 and volume equals six, for example, that doesn't get changed
13:58:56 anywhere in NLU after it gets generated?
13:59:00 A. I mean, this example as I went through disambiguates
13:59:08 six to mean -- to mean what it's supposed to mean and then
13:59:13 uses the context interpreter to join back up with the point
13:59:21 in the dialogue where things have become ambiguous.
13:59:27 Q. Let's just be clear. Each of these examples here,
13:59:31 global set volume intent, notifications multi-turn intent,

13:59:39 music multi-turn intent, they don't get changed anywhere in
13:59:42 the NLU process after they get generated in this step?
13:59:45 A. I don't know that they get changed or not, that isn't
13:59:48 what this example is using it for.
13:59:50 Q. Right, so as far as you know, all the meanings of the
13:59:53 words have already been determined before we have used any
13:59:59 dialogue context, right?
14:00:02 A. The words are -- this is talking about how the words
14:00:07 are received and how the words can then -- and then later on
14:00:12 it talks about how the words can be used to disambiguate an
14:00:17 existing dialogue.
14:00:18 Q. Let's be clear. You just identified volume level
14:00:23 equals six as a meaning of the word six, right?
14:00:27 A. That is -- that's one of the possible meanings, yes.
14:00:30 Q. Right. And time equals six is another meaning of the
14:00:33 word six, right?
14:00:34 A. Yes.
14:00:35 Q. And song name equals six is another meaning of the
14:00:38 word six, right?
14:00:39 A. Yes.
14:00:39 Q. And those meanings have all already been established
14:00:42 once we get to this step in the NLU process; right?
14:00:46 A. I don't know what you mean by established.
14:00:48 Q. They're right there.
14:00:50 A. So those are three hypotheses.

14:00:52 Q. Right, and the output of NLU is a set of hypotheses, right?
14:00:57
14:00:58 A. It's a NBest list of hypotheses.
14:01:03 Q. Right, so these are exactly the hypotheses that are going to be output from the NLU, right?
14:01:08
14:01:12 A. This is showing what the NLU comes up with. I don't -- I think calling that establishing the meaning is misleading.
14:01:17
14:01:19
14:01:20 Q. Okay. Let's go forward in your example. After we accumulate the short-term shared knowledge, we accumulate the long-term shared knowledge, we have to identify a context. Right?
14:01:49
14:01:55
14:02:00
14:02:01 A. I'm sorry, we're talking about the '681?
14:02:06 Q. Claim 13.
14:02:07 A. Okay.
14:02:08 Q. And look at slide PDX3-21, please, Mr. Patterson. Sorry, that's the wrong slide. Go to slide 20. Yeah, that was wrong. You actually didn't show a slide where you explained how Alexa identifies a context associated with the utterance, did you?
14:02:23
14:02:46
14:02:51
14:03:01
14:03:02 A. I don't remember -- I know I talked about it, I don't remember whether it was in the slide or not.
14:03:05
14:03:07 Q. And you didn't actually identify what the context was, did you? You have to have a word, it can have different meanings in multiple contexts and then you're
14:03:12
14:03:20

14:03:24 1 supposed to somehow identify one of those contexts from this
14:03:28 2 long-term and short-term knowledge; right? Isn't that
14:03:40 3 correct?
14:03:42 4 A. I'm sorry, which element are we talking about here?
14:03:45 5 Q. Well, let's go to--well, I don't have it highlighted
14:03:49 6 because you don't have a slide that shows it, but it's the
14:03:53 7 identify context element.
14:03:57 8 A. Okay.
14:03:57 9 Q. Right, we can go here. This is the element where we
14:04:04 10 have this utterance, you have some long-term shared
14:04:09 11 knowledge and some short-term shared knowledge, you're
14:04:12 12 supposed to use that to identify context; right? And you
14:04:17 13 didn't show that step to the jury at all this morning, did
14:04:21 14 you?
14:04:23 15 A. So we talked about examples involving the Hunger
14:04:35 16 Games example. We had an example involving disambiguating
14:04:44 17 the word -- you know, 6:30 based upon short-term knowledge.
14:04:53 18 Q. But that's not what's required here. Here we have to
14:04:57 19 take the utterance we started with and we have to identify a
14:05:00 20 context associated with it, and we have to do that in a
14:05:04 21 specific way. It uses both the short-term shared knowledge
14:05:09 22 and also long-term shared knowledge, right? That's what's
14:05:14 23 required here. And you didn't show a single example of
14:05:17 24 that, did you?
14:05:23 25 A. I don't remember what the examples would be. The

14:05:32 1 Harry Potter example, where it was deciding whether or not
14:05:37 2 it was a song or an album was using knowledge about whether
14:05:47 3 the person had purchased the album.
14:05:50 4 Q. Well, that's not both long-term shared knowledge and
14:05:55 5 short-term shared knowledge, is it? You got to use both,
14:06:00 6 right?
14:06:03 7 A. Let me think about that example.
14:06:09 8 Q. You certainly didn't explain this to the jury, did
14:06:12 9 you?
14:06:12 10 A. Well, I explained to the jury that we were
14:06:15 11 identifying a context associated with the utterance where it
14:06:22 12 was using short-term knowledge that came from additional
14:06:34 13 elements in the conversation.
14:06:36 14 Q. You didn't show that. Tell me what the utterance
14:06:40 15 was, what the additional conversation information was, and
14:06:44 16 what the long-term shared knowledge was. Because that is
14:06:48 17 all required to show that you did this.
14:06:52 18 A. So, I showed that 6:30 was a.m. -- was a.m. based
14:07:03 19 upon subsequent information that came in on the next
14:07:08 20 utterance.
14:07:11 21 Q. That doesn't have anything to do with what this claim
14:07:14 22 requires. You're just talking about a Harry Potter example,
14:07:18 23 now you're switching to 6:30 a.m. Do you have any example
14:07:22 24 where you actually did this?
14:07:25 25 A. I don't know that I have a specific example in mind

14:07:28 1 other than the ones that I presented.
14:07:29 2 Q. So you acknowledge that you did not show the jury
14:07:33 3 that this required element of the claim is actually done by
14:07:37 4 Alexa, right?
14:07:38 5 A. Alexa is in the NLU as we saw from the documents, is
14:07:46 6 -- is disambiguating words that have multiple meanings based
14:07:53 7 upon different pieces of information.
14:07:57 8 Q. Okay. But that is not what this claim requires. The
14:07:59 9 claim does not say disambiguate words using different pieces
14:08:05 10 he have information, that is not what this claim requires,
14:08:07 11 the claim requires these specific steps and you haven't
14:08:10 12 shown that Alexa does any of them, and specifically not this
14:08:19 13 one; right?
14:08:20 14 Let's go forward to the last element -- well,
14:08:38 15 let's go forward to the last element which is to generate
14:08:44 16 response to utterance, this would be PDX3-9, I think, can
14:08:52 17 you see that, Mr. Patterson. Sorry, that's not the right
14:09:09 18 number. We don't have a slide for this either. Let's go
14:09:15 19 back in to the claim itself, Mr. Patterson. Here we go.
14:09:24 20 So this claim element requires not just
14:09:28 21 generating a response, but it requires wherein the
14:09:33 22 conversational speech engine grammatically or syntactically
14:09:39 23 adapts the response based on the intended meaning, do you
14:09:43 24 see that?
14:09:43 25 A. Yes.

14:09:44 Q. You do not cite a single Amazon document that talks
14:09:49 about grammatically or syntactically adapting this response,
14:09:55 correct?
14:09:56 A. I showed examples of responses that Alexa was
14:10:04 generating that were responsive to the needs of the user.
14:10:10 And those responses were grammatically or syntactically
14:10:17 adapted.
14:10:17 Q. You didn't show that. You showed a response, but
14:10:20 this requires more than generating a response, it
14:10:23 specifically requires grammatically or syntactically
14:10:28 adapting it, and what you showed was this slide here which
14:10:31 doesn't show anything about grammatically or syntactically
14:10:35 adapting anything, does it?
14:10:37 A. No, I think one of ordinary skill would understand
14:10:40 that grammatically or syntactically adapted simply means
14:10:44 that it's not just a canned response, it's not just okay, or
14:10:49 some canned response, it's something which is customized to
14:10:53 the situation.
14:10:54 Q. You haven't seen any Amazon source code that
14:10:57 grammatically or syntactically adapts a response, correct?
14:11:02 A. I'm sure I have, Amazon respond -- and I showed in my
14:11:07 demo that Amazon responds to things in ways that are
14:11:10 specific to the user's request.
14:11:12 Q. No, that wasn't my question. My question was it's a
14:11:15 true fact you haven't seen any Amazon source code that

14:11:19 1 grammatically or syntactically adapts a response, correct?
14:11:22 2 A. I don't recall specifically.
14:11:24 3 Q. You haven't seen any testimony from any Amazon
14:11:28 4 witness or engineer that says that Amazon grammatically or
14:11:33 5 syntactically adapts a response, correct?
14:11:35 6 A. They haven't so-they don't use those words, they say
14:11:38 7 that the response is in response to what the user has
14:11:42 8 requested.
14:11:43 9 Q. And the only evidence you provided to this jury today
14:11:47 10 is this diagram that doesn't say anything about
14:11:52 11 grammatically or syntactically adapting a response, correct?
14:11:57 12 A. I provided this and the fact that Amazon is
14:12:00 13 responding to a request from a user.
14:12:02 14 Q. Okay. Let's move to the '176 patent. If we look at
14:12:22 15 slide PDX3-32. You put up this slide in this element about
14:12:39 16 interpreting the recognized words or phrases wherein
14:12:45 17 interpreting the recognized words or phrases includes
14:12:49 18 establishing a context, do you see that?
14:12:51 19 A. Yes.
14:12:51 20 Q. And your example utterance was Alexa, I want to buy
14:12:57 21 an iPhone case; right?
14:13:08 22 A. That was one example I gave, yes.
14:13:10 23 Q. Right. When you say Alexa, I want to buy an iPhone
14:13:15 24 case, this context interpreter is not used at all?
14:13:20 25 A. This particular example here that was shown was

14:13:25 1 associated with the two-step approach for resolving
14:13:29 2 ambiguities and things of that nature.
14:13:32 3 Q. You're not answering my question. In the example
14:13:35 4 that you used to try to show that Alexa does this, your
14:13:38 5 utterance was Alexa, I want to buy an iPhone case, and it's
14:13:43 6 a true fact that when you say Alexa, I want to buy an iPhone
14:13:47 7 case, what you have shown on the right-hand side of this
14:13:50 8 slide is not used at all?
14:13:54 9 A. I don't believe it's using the two-step approach
14:13:58 10 that's talked about in this part of the document.
14:14:00 11 Q. So you agree with me that this is inapplicable to
14:14:04 12 your example, right?
14:14:08 13 A. The context interpreter may well be being used but
14:14:12 14 it's definitely not part of a two-step approach.
14:14:21 15 Q. Let's go to slide PDX3-35, please, Mr. Patterson.
14:14:28 16 Now, this is the element that requires reinterpreting the
14:14:32 17 words or phrases in response to your predetermined event.
14:14:37 18 Right?
14:14:38 19 A. Yes.
14:14:38 20 Q. And what you show on the right is in re-ranking
14:14:46 21 process within Alexa's NLU?
14:14:49 22 A. That's correct.
14:14:49 23 Q. And this re-ranking process happens before the NLU
14:14:53 24 outputs its result; right?
14:14:56 25 A. Well, the re-ranking process, there can be a run

14:15:04 1 through the system where it has -- where it ranks it, then
14:15:08 2 there is an issue, there is something that happens, and
14:15:12 3 there is a signal of some sort and then it re-ranks and then
14:15:16 4 it outputs the re-ranked version.
14:15:19 5 Q. You said that, but what really happens is that within
14:15:24 6 the NLU, the re-ranker will re-rank based on information
14:15:31 7 that had already been provided to the NLU when it was
14:15:34 8 interpreting the current utterance, right?
14:15:37 9 A. Well the timing of when it gets that information
14:15:43 10 isn't particularly discussed. It's -- it has -- the idea is
14:15:48 11 that the NLU produces an NBest list, and then other
14:15:53 12 information can be brought in to re-rank it.
14:15:56 13 Q. But the NBest list is not provided to a speechlet or
14:16:07 14 a program that would respond to the user until after this
14:16:10 15 re-ranking process has been done, right?
14:16:12 16 A. Not necessarily, under system features that feed the
14:16:16 17 re-ranker, one of the inputs is the ASR. So you could have
14:16:21 18 -- you could have things feeding into the -- the decision to
14:16:25 19 re-rank that come in at various points in time.
14:16:28 20 Q. Well, let's be clear about this. When it says ASR,
14:16:32 21 that's the ASR results that were -- that was produced, these
14:16:39 22 are the words from speech recognition that produced or that
14:16:44 23 are being used to generate the intents we're looking at
14:16:48 24 here; right?
14:16:49 25 A. I mean, it doesn't -- it doesn't say that, and the

14:16:51 1 ASR was already provided as input to the NLU.
14:16:55 2 Q. Let's be clear. You don't have any evidence that
14:16:59 3 re-ranker somehow operates because a user indicates that
14:17:05 4 Alexa made a mistake, right, you just made that up?
14:17:09 5 A. That the user said it? I don't think I said that
14:17:13 6 because the user said it made a mistake.
14:17:15 7 Q. You don't identify any actual predetermined event
14:17:18 8 that would cause Alexa's NLU to reinterpret the words or
14:17:25 9 phrases, correct?
14:17:26 10 A. It says that their system features and various kinds
14:17:30 11 of input including context that would cause it to re-rank.
14:17:33 12 Q. That wasn't the question. You haven't identified a
14:17:37 13 single predetermined event that would could cause Alexa to
14:17:40 14 somehow reinterpret the words or phrases, correct?
14:17:43 15 A. I haven't picked one out, no.
14:18:07 16 Q. Let's look at slide PDX3-42. This one that interpret
14:18:14 17 the natural language utterance based to the advertisement
14:18:16 18 and responsive to the existence of a pronoun in the natural
14:18:20 19 language utterance, determine whether the pronoun refers to
14:18:22 20 one or more of the product or service, and you understand
14:18:27 21 that for Alexa's NLU to interpret a pronoun, a speechlet has
14:18:36 22 to first provide a dialog act, right?
14:18:42 23 A. I'm sorry, say that again.
14:18:43 24 Q. Sure, for Alexa's NLU to actually determine the
14:18:48 25 meaning of a pronoun if it's going to do that, it has to

14:18:51 1 first receive a dialog act from the speechlet, right?
14:18:59 2 A. Dialog act from the speechlet. I'm not sure I'm
14:19:03 3 following what you mean in this context here.
14:19:05 4 Q. Sure. So you actually have no evidence other than
14:19:10 5 what you thought you heard from Alexa, whether or not the
14:19:15 6 NLU actually determines whether the pronoun refers to one or
14:19:21 7 more of a product or service, correct?
14:19:23 8 A. So, I don't know whether it's -- I don't know what
14:19:28 9 it's done with the pronoun, I don't know that -- what it
14:19:31 10 thinks about the pronoun, what I know and what I have shown
14:19:36 11 here, is that it is responsive to the existence of the
14:19:42 12 pronoun in the natural language utterance. So the pronoun
14:19:46 13 exists in one utterance, doesn't exist in another, and there
14:19:49 14 are different responses.
14:19:50 15 Q. Right. But you don't know whether in either of these
14:19:54 16 cases the pronoun is actually used by the NLU, is used to
14:20:01 17 identify a product or service, right?
14:20:03 18 A. I don't know -- in this particular case, I didn't
14:20:07 19 analyze any deeper than seeing that it clearly is doing
14:20:12 20 exactly what's in the claim element.
14:20:15 21 Q. Well it's not, right, because this claim element
14:20:20 22 requires that the system, right, Alexa, do a determination,
14:20:25 23 which is whether the pronoun refers to a product or service.
14:20:29 24 And we both know from your deposition and the evidence in
14:20:34 25 this case, that when a user says buy it now, Alexa ignores

14:20:41 1 the it, right, you get the same result when you say buy now,
14:20:46 2 right?
14:20:46 3 A. Buy it now is a different example than this.
14:20:49 4 Q. I understand. I understand. We agree on that, don't
14:20:52 5 we?
14:20:52 6 A. I believe it is true if you say buy now or buy it
14:20:57 7 now, the behavior is the same.
14:20:59 8 Q. Right.
14:21:00 9 A. Clearly here what color produces a different result
14:21:05 10 from what color is it.
14:21:06 11 Q. Sure. And the reason what color gets a different
14:21:09 12 result, is that the intent is generated from what color gets
14:21:14 13 routed through a different speechlet in this example in the
14:21:18 14 shopping speechlet, right?
14:21:20 15 A. It would seem that what color gets routed to
14:21:24 16 something that's looking up on the internet for a
14:21:27 17 definition, but it also, it's looking up the color peach,
14:21:33 18 which I don't know where that came from. It just seems
14:21:35 19 clear that the existence of the pronoun tells Alexa that
14:21:40 20 we're talking about the purchase that we were making
14:21:44 21 previously.
14:21:45 22 Q. Well, the existence of the pronoun may make it more
14:21:50 23 likely that the intent gets routed to the shopping
14:21:54 24 speechlet, it doesn't show that the NLU resolves the pronoun
14:21:59 25 to anything, right?

14:22:00 1 A. It's clearly showing that when I use the pronoun,
14:22:08 2 that I'm talking about the product that I'm trying to
14:22:11 3 purchase.
14:22:11 4 Q. Well that's not true. As far as the NLU knows, this
14:22:15 5 could be just like buy now or buy it now, the only
14:22:21 6 difference in this example is one of those utterances got
14:22:24 7 routed to something other than shopping; right?
14:22:27 8 A. I think the behavior of the system speaks for itself.
14:22:32 9 Q. Well, let's be clear, you haven't identified any
14:22:37 10 Amazon engineer testimony that supports that the pronoun is
14:22:40 11 being resolved in this example, right?
14:22:45 12 A. You said the pronoun is being resolved --
14:22:48 13 Q. Sure, you have no evidence from this slide from any
14:22:52 14 Amazon engineer that this pronoun is being resolved in the
14:22:55 15 way the claim requires; right?
14:22:57 16 A. I have no testimony from an Amazon engineer about it.
14:23:03 17 Q. And you have no document from Amazon --
14:23:05 18 THE COURT: I think he was --
14:23:07 19 You can finish your answer.
14:23:11 20 THE WITNESS: Thank you, Your Honor. I think
14:23:13 21 that the claim doesn't say resolve, it's clear that the
14:23:19 22 presence of the pronoun tells Alexa that we're talking about
14:23:24 23 the thing that I'm buying it, and you can tell that by the
14:23:27 24 fact that it's giving me the color stone, which is the color
14:23:30 25 of the thing that is being purchased.

14:23:32 1 **BY MR. HADDEN:**

14:23:33 2 Q. **Let's be clear. You do not have any Amazon document**

14:23:36 3 **that says that Amazon's NLU determines whether the pronoun**

14:23:41 4 **it refers to anything; correct?**

14:23:43 5 A. **I am not presenting -- I am not presenting Amazon**

14:23:49 6 **technical documents on this.**

14:23:50 7 Q. **You are not presenting Amazon source code that would**

14:23:54 8 **show that the pronoun is determined to refer to anything;**

14:23:58 9 **right?**

14:23:59 10 A. **I'm not presenting source code here.**

14:24:02 11 Q. **Let's look at the '703 patent. Now, the utterance**

14:24:12 12 **that you used when you were explaining this example was**

14:24:19 13 **Alexa, I want to buy an iPhone case, right?**

14:24:22 14 A. **Yes, I believe so.**

14:24:24 15 Q. **And when you say Alexa, I want to buy an iPhone case,**

14:24:31 16 **you don't have any evidence to show that Alexa would obtain**

14:24:38 17 **your shipping information; right?**

14:24:45 18 A. **Well it shipped it to my home, so --**

14:24:48 19 Q. **Let's be clear. Let's look at slide PDX3-55, please,**

14:25:04 20 **Mr. Patterson. So if we look at the element obtain without**

14:25:10 21 **further user input after the receipt of the user input**

14:25:13 22 **shipping information, do you see that?**

14:25:15 23 A. **Yes.**

14:25:16 24 Q. **Right. So this stuff about without further user**

14:25:20 25 **input means that the user can't say anything else after they**

14:25:25 said in your example Alexa, I want to buy an iPhone case;
14:25:29 right?
14:25:57 A. So, when I say buy it now --
14:26:01 Q. That's not the utterance that got you started. You
14:26:04 got to have the same utterance, right, we start receive a
14:26:07 user input, comprising a natural language utterance. And
14:26:11 what you said was, I want to buy an iPhone case, right?
14:26:14 A. I did say that, that's right.
14:26:16 Q. So that's the user input. Now without any other user
14:26:21 input, we have to obtain shipping information. Right?
14:26:27 Alexa doesn't obtain your shipping information when you say
14:26:32 I want to buy an iPhone case, right?
14:26:38 A. I believe when I say buy it now, it then goes through
14:26:46 and --
14:26:46 Q. Sir you're not reading the claim. It says you have
14:26:50 to obtain without further user input after the receipt of
14:26:54 the user input. And the user input was I want to buy an
14:26:58 iPhone case, right?
14:27:00 A. And the user --
14:27:01 Q. Let me finish the question, please, so it has to
14:27:04 obtain the shipping information at that point without any
14:27:08 further user input. So you haven't shown and it's a true
14:27:12 fact that Alexa does not obtain the shipping information
14:27:16 when you say Alexa I want to buy an iPhone case; right?
14:27:20 A. That may well be, when I say Alexa buy it now, at

14:27:26 1 that point it obtains my shipping information and my payment
14:27:29 2 information and it process all the way through.
14:27:31 3 Q. Right and when --
14:27:32 4 THE COURT: Hold on, you're still cutting him
14:27:36 5 off when he's still talking, you finish and then you start.
14:27:39 6 Go ahead, sir.
14:27:40 7 MR. HADDEN: Sorry, Your Honor.
14:27:42 8 THE WITNESS: When I say Alexa buy it now, at
14:27:45 9 that point, it obtains my shipping information, it obtains
14:27:47 10 my payment information and it completes the purchase without
14:27:50 11 any further user intervention.
14:27:52 12 BY MR. HADDEN:
14:27:52 13 Q. Sure. But if we look at what the claim requires, it
14:27:55 14 requires that the shipping information has been obtained
14:27:59 15 after you say in your example Alexa, I want to buy an iPhone
14:28:04 16 case because it explicitly requires without further user
14:28:09 17 input. If you were going to say buy it now, that would be
14:28:13 18 further user input; right?
14:28:16 19 A. I think buy it now is the user -- is the initial user
14:28:25 20 input.
14:28:26 21 Q. That's not what you told us this morning. This
14:28:28 22 morning you told us, it was I want to buy an iPhone case.
14:28:32 23 So let's say instead you said buy it now. If you say buy it
14:28:41 24 now, the rest of the claim doesn't work because you don't
14:28:47 25 identify the product based on buy it now; right? You need

14:28:59 1 to pick one input and stick with it through the claim.
14:29:16 2 A. I think when I said I want to buy an iPhone case,
14:29:21 3 that was establishing various information, short-term
14:29:24 4 information, and I said buy it now, that was the utterance
14:29:30 5 that's being talked about here, and from that it completed
14:29:34 6 the purchase.
14:29:38 7 Q. Well, the problem is if you look at the identifying
14:29:43 8 stuff, right, identify without further user input after the
14:29:48 9 receipt of the user input, the product or service to be
14:29:54 10 purchased, right? Now the product you're trying to buy, the
14:29:58 11 iPhone case, it had already been identified before you said
14:30:02 12 buy it now. That's why in your presentation to this jury
14:30:05 13 this morning, you said instead, Alexa, I want to buy an
14:30:09 14 iPhone case, right?
14:30:10 15 A. I think you're misinterpreting here. The identify,
14:30:15 16 obtain and obtain steps, there is an identify, and then the
14:30:20 17 next step is obtain, and the next was obtain, those are all
14:30:24 18 internal. So the obtain, I have already provided payment
14:30:27 19 information to Amazon, so when it says obtain, I have
14:30:31 20 already provided that on the website. And when it says
14:30:34 21 obtain payment information, it's getting it from the
14:30:39 22 database. And when I say obtain -- obtain without further
14:30:43 23 user input shipping information, it's doing that from my
14:30:49 24 account, the long-term information that it has with Amazon.
14:30:52 25 So those identify, obtain, obtain, are internal.

14:30:57 Q. Understood. I'm going back to what your input was.
14:31:02 Your input, Alexa I want to buy an iPhone case, is that
14:31:06 where we start this claim, or do we start with the claim
14:31:10 with the buy it now?
14:31:11 A. I think for the purposes of this analysis, it's Alexa
14:31:14 buy it now.
14:31:14 Q. So if it's buy it now, nowhere did you show the
14:31:18 identify without further user input receipt of the user
14:31:22 input the product or service to be purchased. Right?
14:31:27 That's not identified by buy it now, right?
14:31:31 A. That information is already part of the context of
14:31:33 the conversation. I've already established it and the
14:31:38 identify is that Alexa is getting it from the context.
14:31:47 Q. I didn't follow that. But we'll move on.
14:31:50 Let's look at the last element of this claim,
14:31:52 please, Mr. Patterson. Now, this required that the system,
14:31:59 Alexa, request, provide a request for user confirmation to
14:32:08 use the payment information and the shipping information for
14:32:12 a purchase transaction, correct? And you didn't show that
14:32:17 at all; correct? You showed this slide. Nowhere did you
14:32:25 show Alexa saying Dr. Polish, please confirm this payment
14:32:31 information and the shipping information, right?
14:32:41 A. I think that's right, I think the receipt was sent --
14:32:44 was sent via e-mail.
14:32:46 Q. And you mentioned this morning that though you didn't

14:32:54 show it to the jury, that you could somehow have a
14:32:58 single-turn utterance that would itself include short-term
14:33:05 knowledge about itself, right?
14:33:10 A. I don't think I said exactly those words, but I did
14:33:13 say that you could have -- you could have -- it's possible
14:33:16 to have a single -- a single utterance that would have
14:33:20 short-term information in it.
14:33:23 Q. But you didn't give the jury any example of one of
14:33:26 those, did you?
14:33:27 A. That's correct.
14:33:28 Q. You never testified that such a thing actually --
14:33:31 well, in all of the interaction that you talked about with
14:33:37 Alexa, you did not provide an example of such an utterance,
14:33:41 correct?
14:33:41 A. That's right.
14:33:43 MR. HADDEN: No further questions. Thank you.
14:33:46 THE COURT: Redirect.
14:33:48 REDIRECT EXAMINATION
14:33:49 BY MR. YOON:
14:33:49 Q. So, Dr. Polish, a couple of things. First of all, in
14:33:54 the questioning by counsel for Amazon, counsel did not ask
14:34:02 you about the speaker ID or Voice ID, did he?
14:34:05 A. That's correct.
14:34:06 Q. And the speaker ID and Voice ID you showed in your
14:34:10 demonstration, what is that?

14:34:12 1 A. So, the Speaker Voice ID is a -- its a component of
14:34:20 2 the Alexa system that you have to enroll, when you speak a
14:34:25 3 number of sentences, I think it was in my example, and it
14:34:29 4 learns your voice so that later on when you use the system,
14:34:35 5 it identifies you as you, and can customize the operation of
14:34:39 6 it for you.
14:34:39 7 Q. And is that long-term shared information that's used
14:34:45 8 in each communication with Alexa thereafter?
14:34:48 9 A. Yes, absolutely.
14:34:49 10 Q. Now, there was a discussion regarding -- if we could
14:34:54 11 have slide PDX3-55, on the screen. Do you recall there was
14:35:10 12 a discussion, if we could just focus on the right side for a
14:35:13 13 moment. There was a discussion regarding the I want to buy
14:35:18 14 an iPhone case and buy it now. Do you recall that?
14:35:21 15 A. Yes.
14:35:21 16 Q. And this is after you had already done the speaker
14:35:24 17 ID?
14:35:24 18 A. Yes.
14:35:25 19 Q. And so at the time you did this transaction because
14:35:28 20 of the speaker ID, what did Alexa know?
14:35:31 21 A. It knew it was me, and it had my payment and shipping
14:35:35 22 information.
14:35:35 23 Q. Okay. If we go to the bottom on there. The
14:35:39 24 statement buy it now, do you see that we had some discussion
14:35:44 25 about pronouns, do you recall that, sir?

14:35:45 A. Yes.
14:35:46 Q. And when Alexa heard the word it, what did it
14:35:50 understand it would refer to?
14:35:51 A. Well, it was referring to the previously established
14:35:54 iPhone case that I had asked for.
14:35:56 Q. So when you said buy it now, that phrase, did Alexa
14:36:04 know what it was going to buy?
14:36:06 A. Yes.
14:36:06 Q. Did it have your shipping information?
14:36:08 A. Yes.
14:36:09 Q. Did it have your financial information?
14:36:11 A. Yes.
14:36:11 Q. Did it execute the transaction without further input?
14:36:14 A. It did.
14:36:15 Q. Now, let's now go to PTX-226. And this is the
14:36:26 PTX-226, this is the natural language deep dive that has
14:36:31 been shown to a large number of people at Amazon?
14:36:34 A. Yes.
14:36:35 Q. And this was what Mr. Vanee testified about?
14:36:38 A. Yes.
14:36:38 Q. And both counsel for VoiceBox and Amazon have asked
14:36:43 you about this document; correct?
14:36:45 A. Yes.
14:36:46 Q. Now, looking at this one, let's now go to the example
14:36:50 given in the Amazon, actual Amazon technical document, shall

14:36:55 1 we. And if we could turn to page 12. Slide 32, I'm sorry,
14:37:03 2 slide 32.
14:37:04 3 This is the multi-turn example about setting an
14:37:09 4 alarm clock.
14:37:11 5 A. Yes.
14:37:11 6 Q. And prior to this conversation, Alexa would have the
14:37:17 7 long-term information from your speaker ID?
14:37:19 8 A. Yes.
14:37:20 9 Q. And if, for example, you had previously set the
14:37:24 10 alarm, Alexa would also have the long-term information about
14:37:27 11 what was previously set as an alarm?
14:37:30 12 A. Yes.
14:37:30 13 Q. Let's walk through what if actual Amazon technical
14:37:34 14 document says. The first phrase there is what the user
14:37:39 15 would say to Alexa, Alexa set alarm. That's an utterance,
14:37:43 16 correct?
14:37:43 17 A. Yes.
14:37:44 18 Q. And if we could have in a split screen this slide and
14:37:49 19 Claim 13 of the '681 patent on the screen. Let's walk
14:37:53 20 through the element.
14:37:58 21 So, looking at Claim 13, the first one is
14:38:03 22 receive an utterance at a voice input device?
14:38:07 23 A. Yes.
14:38:08 24 Q. If we can just highlight the first part of that one.
14:38:11 25 In this example, has Alexa received an utterance

14:38:15 1 at a voice input device?
14:38:17 2 A. Yes.
14:38:18 3 Q. Now, let's go to see what happened when you say that,
14:38:23 4 Alexa says for what time? Do you see that?
14:38:26 5 A. Yes.
14:38:26 6 Q. And now the user says six. Do you see that?
14:38:30 7 A. Yes.
14:38:31 8 Q. Now, let's go to what is shown in the Alexa technical
14:38:41 9 document. If we can go to slide 36. At slide 36, the
14:38:50 10 statement six is made, do you see that?
14:38:53 11 A. Yes.
14:38:53 12 Q. And at the time the word six is made in this
14:38:57 13 conversation with Alexa, has short-term knowledge, shared
14:39:02 14 knowledge been accumulated about the current conversation?
14:39:05 15 A. Yes.
14:39:06 16 Q. And what is that?
14:39:07 17 A. Well, that we're working on an alarm and that we're
14:39:13 18 setting a notification.
14:39:15 19 Q. Right. And in fact, the previous intent was set
14:39:19 20 notification intent, that's to set the alarm?
14:39:22 21 A. Right, that's at least in part the context and it is
14:39:25 22 the short-term information.
14:39:26 23 Q. If we can highlight short-term shared knowledge about
14:39:30 24 the current conversation. And if we can highlight
14:39:46 25 accumulate, too.

14:39:47 Let's talk about the next element, accumulate
14:39:50 long-term shared knowledge about the user. Do you see that?
14:39:53 A. Yes.
14:39:54 Q. And an example of that was the speaker ID that you
14:39:58 showed in your example?
14:39:59 A. Yes.
14:39:59 Q. And another example would be a previous alarm?
14:40:03 A. Yes.
14:40:05 Q. Now, let's go back to the multi-turn example in the
14:40:10 Amazon technical document. Can we go back to slide 32. And
14:40:14 you see that in response to six, Alexa says is that six in
14:40:19 the morning or evening. Do you see that?
14:40:21 A. Yes.
14:40:21 Q. And so now if we could go to slide 38, at this time,
14:40:34 from previous conversations would Alexa know that this was
14:40:39 for an alarm to be set?
14:40:40 A. Yes.
14:40:40 Q. Would that be the previous intent?
14:40:42 A. Yes, it would.
14:40:43 Q. And would Alexa know in the short-term information
14:40:47 what time was previously offered in the build slot?
14:40:52 A. It has time equals 0600 and 1800.
14:40:57 Q. When you say morning, what does Alexa do having known
14:41:02 the context of the conversation in the previous short-term
14:41:05 information?

14:41:06 1 A. Then it changes the slot to be time equals 0600.
14:41:11 2 Q. Okay. And so looking at the next element, it says
14:41:15 3 identify a context associated with the utterance, did Alexa
14:41:22 4 do that?
14:41:24 5 A. Yes.
14:41:24 6 Q. Did it use both the short-term and long-term
14:41:27 7 knowledge?
14:41:27 8 A. Yes.
14:41:27 9 Q. For example, with regards to the long-term knowledge,
14:41:30 10 would that be the speaker ID?
14:41:32 11 A. Yes.
14:41:33 12 Q. And if you were changing the alarm, would the
14:41:36 13 long-term knowledge include the previous alarm?
14:41:39 14 A. Yes.
14:41:40 15 Q. Now, if you look at the next element, establish an
14:41:44 16 intended meaning for an utterance within the identified
14:41:46 17 context, as shown in slide 38, does Alexa do that?
14:41:53 18 A. Yes.
14:41:54 19 Q. And if you go to the last element, generate a
14:41:57 20 response to the utterance, if we can now go back to slide
14:42:01 21 32, after the user says morning, what does Alexa do?
14:42:08 22 A. Well, it has figured out the slot now and it comes
14:42:13 23 back with alarm set for 6:00 a.m. tomorrow.
14:42:16 24 Q. So does Alexa practice that element in the example
14:42:20 25 that you provided?

14:42:20 A. Yes.
14:42:22 Q. Now, there was some discussion about whether or not
14:42:29 Alexa might ask the user to help it resolve ambiguity, do
14:42:35 you recall that?
14:42:35 A. Yes.
14:42:36 Q. You can take that down, Mr. Smith. If we could go to
14:42:39 slide 24.
14:42:40 Actually, go to slide 23, please. This is on
14:42:56 PTX-226, slide 23, I believe, PTX, the actual exhibit. Next
14:43:09 page. And next page. There it is.
14:43:13 Are there situations where for example in the
14:43:17 utterance play Hunger Games that Alexa would find what they
14:43:22 call true ambiguity?
14:43:24 A. Yes.
14:43:24 Q. In this situation here, the scores were close, .82
14:43:30 and point .81.
14:43:32 A. Right.
14:43:32 Q. Because the scores were so close, Alexa couldn't make
14:43:36 a call as to what would be the blue arrow, is that fair to
14:43:39 say?
14:43:39 A. That's right.
14:43:40 Q. If you go to the next page. When Alexa faces true
14:43:45 ambiguity, does it ask the user to help it correct it?
14:43:49 A. Yes, this is an example of it, where it's asking the
14:43:52 user for some information to help it out.

14:43:55 Q. And would that be part of a multi-turn conversation
14:44:00 that Alexa would have with the user?
14:44:02 A. Yes. This is a multi-turn conversation.
14:44:04 Q. In that conversation, would the NLU do multiple
14:44:09 turns?
14:44:09 A. Yes, it would cycle through multiple times to resolve
14:44:16 it.
14:44:16 Q. And if we could now go back to slide 22. And sir,
14:44:26 there were some questions about whether or not Alexa
14:44:31 interpreted certain things. But I wanted to point out one
14:44:36 thing here, actually. If we go back to the question that
14:44:40 you were asked -- I apologize, I'm going to move to a
14:44:45 different slide and come back to this. It was -- do you see
14:44:51 on slide 33, please. Do you remember counsel for Amazon
14:45:00 asking you about whether the NLU models are currently aware
14:45:05 of a dialog context, do you recall that?
14:45:13 A. Yes.
14:45:14 Q. If we could go back to slide 11. And do you see that
14:45:19 this is the recognition for the NLU core pipeline?
14:45:25 A. Right.
14:45:25 Q. And here is the model, do you see that, sir?
14:45:29 A. Yes.
14:45:29 Q. And the slide we're talking about is the models might
14:45:33 not be at the time using the dialog context, is that fair to
14:45:38 say?

14:45:38 A. That's right.
14:45:38 Q. And you recall that in your response to counsel, you
14:45:43 indicated there was also rules and FST, do you see that?
14:45:47 A. Right.
14:45:47 Q. And this input, this would be coming from the ASR?
14:45:52 A. That's right.
14:45:53 Q. And so the match, the rules, the FST and the models
14:45:58 would receive the output from the ASR, correct?
14:46:02 MR. HADDEN: Objection, leading Your Honor.
14:46:03 THE COURT: Could you make this a little less
14:46:06 leading.
14:46:08 MR. YOON: Yes, Your Honor, my fault.
14:46:10 BY MR. YOON:
14:46:10 Q. Do you see this in this document here just the
14:46:13 reference to the rules, FST and models. How does that
14:46:17 receive the input from the ASR?
14:46:19 A. So the ASR provides input to the NLU and that's --
14:46:25 according to this diagram, it's sent to the rules on
14:46:30 pathway, the FST pathway and the models pathway.
14:46:34 Q. And now going back to slide 34. Now, do you recall
14:46:43 you were asked some questions about this slide in the
14:46:47 context of the models, the DNN models, do you recall that?
14:46:52 A. Yes.
14:46:52 Q. And you were asked questions, particularly I think
14:46:56 you discussed six?

14:46:58 A. Yes.
14:46:59 Q. Now, do you see the title of this slide is a two-step
14:47:02 approach?
14:47:02 A. Yes.
14:47:03 Q. And what was Step 1?
14:47:06 A. So, Step 1 it says that recognizer generate the yes
14:47:14 intent, no intent, or slots for the current turn utterance.
14:47:19 Q. And that's only the first step of the two-step
14:47:22 approach, correct?
14:47:22 A. Yes.
14:47:23 Q. And if you go to the next slide. What occurs in
14:47:31 Step 2?
14:47:31 A. In Step 2, the context interpreter merges the yes,
14:47:40 no, multi-turn intent with previous intent and slots so it
14:47:49 takes the previous intent that was somehow ambiguous and
14:47:54 merges it with the results here to disambiguate.
14:47:56 Q. If we go to the next -- page 36. And that Step 2 is
14:48:02 what we were discussing with regard to the clock setting
14:48:05 example, correct?
14:48:05 A. Yes.
14:48:06 Q. Take that down.
14:48:08 Just two more questions, Dr. Polish. Now, you
14:48:15 answered multiple questions regarding source code during the
14:48:19 cross-examination, do you recall that?
14:48:20 A. Yes.

14:48:21 Q. And how did you work with Mr. Peck?
14:48:24 A. So Mr. Peck was doing his analysis of Amazon's source
14:48:33 code, he was being communicated with by the lawyers. I saw
14:48:37 the questions that were being provided to Mr. Peck and I
14:48:42 added my own questions and consulted on what he was -- what
14:48:47 he was tasked with looking for. And then I saw his results
14:48:52 and I saw his reports as a result of his investigation.
14:48:58 Q. And did you read Mr. Peck's report?
14:49:00 A. Yes.
14:49:01 Q. And did you look at the source code that Mr. Peck
14:49:04 identified in his report?
14:49:05 A. Yes.
14:49:05 Q. And you understand that Mr. Peck will be a witness in
14:49:08 this case?
14:49:09 A. Yes.
14:49:09 MR. YOON: No further questions, Your Honor.
14:49:11 THE COURT: All right. Thank you. Thank you,
14:49:13 sir. You may step down.
14:49:14 We're going to take our afternoon break.
14:49:18 COURTROOM DEPUTY: All rise.
14:49:19 (Jury exiting the courtroom at 2:49 p.m.)
14:49:37 THE COURT: All right.
14:49:46 MR. YOON: Your Honor, we can -- Dr. Polish is
14:49:49 excused and we can then recall him in rebuttal?
14:49:52 THE COURT: Excuse me?

14:49:53 MR. YOON: I apologize, Your Honor, I just
14:49:55 wanted to confirm that we can now talk to him that he's off
14:49:59 the stand and everything.
14:50:00 THE COURT: Yes.
14:50:00 MR. YOON: Thank you.
14:50:01 (A brief recess was taken.)
15:10:33 COURTROOM DEPUTY: All rise.
15:10:38 (Jury entering the courtroom at 3:10 p.m.)
15:10:50 THE COURT: All right everyone, welcome back.
15:10:58 Please be seated.
15:10:59 What's next?
15:11:02 MR. SMITH: Your Honor, the plaintiff calls John
15:11:06 Peck.
15:11:17 COURTROOM DEPUTY: Please raise your right hand.
15:11:23 Please state and spell your full name for the record.
15:11:26 THE WITNESS: My name is John Charles Peck
15:11:29 junior, J-O-H-N, Charles, C-H-A-R-L-E-S, Peck, P-E-C-K,
15:11:38 junior.
15:11:40 CHARLES JOHN PECK, JR., having been duly sworn
15:11:44 was examined and testified as follows:
15:11:47 DIRECT EXAMINATION
15:11:49 BY MR. SMITH:
15:11:54 Q. Good afternoon, Mr. Peck.
15:11:56 A. Good afternoon.
15:11:57 Q. Could you please introduce yourself to the jury?

15:12:00 1 A. Yes. My name is John Charles Peck, Jr.
15:12:04 2 Q. And generally what's your role in this case?
15:12:08 3 A. On this case I have been reviewing Amazon source
15:12:13 4 code, Alexa source code.
15:12:14 5 Q. And what's your educational background, briefly?
15:12:18 6 A. Yes, so I have a bachelor of science in computer
15:12:22 7 engineering in Carnegie Mellon University and I have a
15:12:26 8 master of science from computer science from the University
15:12:30 9 of California.
15:12:33 10 Q. Mr. Peck, what's your work experience be like since I
15:12:36 11 graduated college?
15:12:37 12 A. I have been a practicing computer engineer for over
15:12:41 13 twenty-five years, I have developed a lot of software and
15:12:44 14 hardware products.
15:12:46 15 Q. And do you have any experience with voice related
15:12:53 16 technology, like at issue in this case?
15:12:57 17 A. I do. I had a client several years ago who was
15:13:01 18 interested in ASR technology. They wanted me to investigate
15:13:06 19 and profile a software package called Sphinx, it performs
15:13:13 20 ASR.
15:13:13 21 Q. Do you have any experience with big computer systems
15:13:15 22 with lots of servers?
15:13:17 23 A. I do. Part of my prior employment at Foxfire
15:13:24 24 Technologies Corporation, and Transarc Corporation which
15:13:27 25 made distributed file systems, and also with a company

15:13:31 1 called Adion Systems, a client that interacted with servers
15:13:35 2 on a network.
15:13:36 3 Q. And during your professional career, have you ever
15:13:38 4 been awarded patents?
15:13:41 5 A. Yes, I have. I am the named inventor on ten U.S.
15:13:46 6 patents issued.
15:13:46 7 Q. And in what kind of expertise are you bringing to
15:13:54 8 this case?
15:13:54 9 A. In this case I performed a source code review of the
15:13:58 10 Amazon Alexa source code.
15:14:00 11 Q. And have you done source code review like this in
15:14:05 12 other litigation matters?
15:14:08 13 A. Yes, I have. I have performed source code review on
15:14:11 14 several other litigation related matters in the past.
15:14:16 15 MR. SMITH: Your Honor, at this point, VoiceBox
15:14:19 16 offers Mr. Peck as an expert witness in the field of
15:14:22 17 computer source code analysis.
15:14:23 18 MR. HADDEN: No objection.
15:14:24 19 THE COURT: All right. He will be so
15:14:26 20 recognized.
15:14:27 21 MR. SMITH: Thank you, Your Honor.
15:14:28 22 BY MR. SMITH:
15:14:28 23 Q. Mr. Peck, first of all, what is source code?
15:14:32 24 A. So source code are a set of commands that are given
15:14:35 25 to the computer to execute, to perform a function. Text

15:14:39 1 code can be textual form typed in by a person.
15:14:49 2 Q. And when you did you start looking at source code in
15:14:53 3 this case?
15:14:53 4 A. I began looking at source code in this case after
15:14:57 5 July 2020, when I joined the case.
15:15:00 6 Q. And what generally was your assignment for you when
15:15:06 7 you joined the case?
15:15:07 8 A. Generally I was asked to examine Amazon Alexa source
15:15:13 9 code to identify certain features in the source code that
15:15:19 10 perform the system function and report back on that.
15:15:22 11 Q. How did you go about reviewing the code, for example,
15:15:27 12 was it -- did you go somewhere to do that?
15:15:29 13 A. That's right. So this source code was produced by
15:15:34 14 Amazon on to a computer that I understand was located in
15:15:40 15 Amazon's counsel's offices, and I was able to access that
15:15:43 16 computer by a remote computer and examine the source code
15:15:47 17 that way.
15:15:48 18 Q. And was there any way for you to get any evidence
15:15:54 19 from the source code that you were looking at that you could
15:15:58 20 bring to the trial?
15:15:59 21 A. Yes, that's right. There was a process where I could
15:16:03 22 print out source code that I thought was of relevance, and
15:16:08 23 make it available for this trial.
15:16:10 24 Q. And did you hear -- strike that. Were you here for
15:16:16 25 Dr. Polish's testimony earlier?

15:16:18 A. Yes, I was.
15:16:19 Q. And were you asked to compare the source code to the
15:16:25 patent claims like -- or to the patent claims in this case?
15:16:30 A. No, I was not.
15:16:31 Q. What was your role as you understood it relative to
15:16:37 Dr. Polish?
15:16:37 A. I understood that Dr. Polish would be performing that
15:16:40 comparison between the source code and the claims.
15:16:43 Q. And aside from looking at source code, did you look
15:16:48 at any other types of evidence in the case?
15:16:51 A. Yes. I attended depositions from Amazon's witnesses.
15:16:57 I read several deposition transcripts. And I reviewed
15:17:02 Amazon technical documents.
15:17:04 Q. And those depositions, what sort of employees were
15:17:07 you going to depositions for?
15:17:09 A. These were Amazon engineers who worked on the Amazon
15:17:14 Alexa system. There were several witnesses who worked on
15:17:19 different parts of the Amazon Alexa system, including the
15:17:23 NLU, the ASR, and something called ranking arbitration as
15:17:30 well as the shopping.
15:17:31 Q. Based on your analysis of the source code, would you
15:17:35 say that the diagram that Amazon has been showing that they
15:17:38 say indicates Alexa, was that accurate based on the source
15:17:44 code analysis that you did?
15:17:45 A. No, I don't think that's entirely accurate.

15:17:48 Q. Okay. Well, for one thing, did you take a look at a
15:17:54 source code for what we have been referring to as the ASR in
15:17:57 this case?
15:17:58 A. Yes, I did. I reviewed source code for two different
15:18:03 versions of Amazon Alexa ASR, there was an older version and
15:18:08 a newer version that was put to use later.
15:18:10 Q. And did you review any--strike that.
15:18:14 Did you find any source code for an NLU?
15:18:17 A. Yes, I did.
15:18:19 Q. And did you review that code for the NLU, too?
15:18:23 A. Yes, I reviewed that source code for the NLU.
15:18:27 Q. And were you able to identify source code for reusing
15:18:34 information from a prior utterance?
15:18:37 A. Yes, I did identify a source code that uses
15:18:41 information from a prior utterance.
15:18:44 Q. Now, Mr. Peck, I'm going to show you a document here
15:18:49 on the screen here. My understanding is it's sensitive
15:18:56 information from Amazon. So we don't have extra copies.
15:19:00 But I'm just going to put it up on the Elmo for you to take
15:19:04 a look.
15:19:05 And first of all, could you take a look at the
15:19:11 bottom here and just identify the exhibit there, exhibit
15:19:14 number there in the middle?
15:19:15 A. Yes. So this source code is stamped with a Bates
15:19:20 stamp that reads AMZ_VB_ SC print_000074.

15:19:32 Q. Was this some of the evidence that you were able to
15:19:37 collect from the Amazon Alexa source code computer?
15:19:42 A. Yes, it was. And you can see at the very top, it's
15:19:47 hard to see but there is a patent name that's at the very
15:19:53 top and that patent name includes the NLU. Then also in the
15:20:01 code you can see there is a number of statements that say
15:20:04 import and they say Amazon dots blue shift dot NLU for
15:20:10 example, several lines, 11 and 19, for example.
15:20:13 Q. The first thing you were referencing was that name up
15:20:17 here, you said?
15:20:17 A. Yes, that's right, I can see that there now. You can
15:20:21 see there is an NLU in the package name in the beginning, a
15:20:27 name of this particular source code at the file is at the
15:20:30 end, rule based context interpreter dot JAVA.
15:20:35 Q. At this point the plaintiff VB Assets would move into
15:20:42 evidence PTX-541(a)?
15:20:45 MR. HADDEN: No objection.
15:20:52 THE COURT: It's admitted.
15:20:54 (PTX Exhibit No. 541(a) was admitted into
15:20:55 evidence.)
15:20:55 BY MR. SMITH:
15:20:55 Q. At a general level, what does the source code in
15:21:00 Exhibit 541(a) do based on your review?
15:21:05 A. So generally it adds information to current
15:21:08 interpretations from previous interpretations.

15:21:11 Q. And I think part of the file name here is called it's
15:21:17 called a rule based context interpreter, is that right?
15:21:21 A. Yes, that's correct.
15:21:23 Q. And is this code part of the NLU?
15:21:29 A. Yes, it's part of the NLU.
15:21:32 Q. And why don't we go to line 26. And do you see that
15:21:44 kind of here, line 26, there is a text that's kind of green
15:21:49 and there is kind of some asterisk next to it. Do you see
15:21:53 that?
15:21:53 A. Yes.
15:21:53 Q. What is that type of text in the source code?
15:21:56 A. Yes. So this is what's called a comment. And a
15:22:01 comment is a description that's left by a source code
15:22:04 developer. To convey information to one who is reading and
15:22:08 writing the source code. In this case, it reads class merge
15:22:12 current turn NBest with discourse context using rules.
15:22:17 Q. So you're saying this comment would be something the
15:22:22 Amazon engineers would put into the code?
15:22:25 A. Yes, that's correct.
15:22:26 Q. Do you see here at the very end it says context using
15:22:29 rules. Do you see that?
15:22:30 A. I see that.
15:22:31 Q. And so based on the comment here, is the source code
15:22:38 that you looked at that follows consistent with how it's
15:22:42 being described by the Amazon engineers with the comment?

15:22:46 A. Yes, it is. And this source code file is making use
15:22:51 of rules as a way to accomplish what I described earlier as
15:22:57 its function.
15:22:58 Q. Is this an example of Alexa using manual rules?
15:23:03 A. Yes, it is.
15:23:05 Q. Now, Mr. Peck, in the course of your analysis of this
15:23:13 code, were you able to figure out approximately when this
15:23:20 source code file was added to the Amazon computer that you
15:23:26 looked at?
15:23:27 A. Yes, I recorded the date in my report. That date was
15:23:33 August 31st, 2020.
15:23:34 Q. So you found examples of manual rules on the Amazon
15:23:38 computers in August 2020?
15:23:42 A. That's correct.
15:23:47 Q. Now, why don't we go to line, the next page here, you
15:24:03 see here line 79?
15:24:07 A. I see that.
15:24:07 Q. And then if you go over there is some text here.
15:24:12 When you see this text here on line 79, what is that text?
15:24:16 A. So this text is actual source code that's executed by
15:24:22 the computer when it runs. And this is one statement, one
15:24:29 operation that takes place in the source code.
15:24:31 Q. And so this is how -- this is code that Alexa would
15:24:36 actually use when it's running?
15:24:38 A. Yes, that's correct.

15:24:39 Q. And do you see where it says get rules, do you see
15:24:43 that?
15:24:43 A. I do see that.
15:24:45 Q. What does that refer to?
15:24:46 A. So this is what's called a function call. It's
15:24:52 retrieving a set of rules to use in part of this operation.
15:24:58 Q. And does this operation relate to what we heard
15:25:03 earlier from Dr. Polish about context and using past
15:25:11 information to fill in the gap for what the user wants?
15:25:17 A. Yes, it does. This component is named the
15:25:20 rules-based context interpreter. It's taking information
15:25:24 from previous interpretations and adding to current
15:25:29 interpretations.
15:25:31 Q. Now, Mr. Peck, was this code file that we're looking
15:25:36 at the only example of Amazon Alexa using manual rules?
15:25:41 A. No, there were many other examples.
15:25:45 Q. Let me put on the document viewer the next document.
15:25:54 And do you see here, if we go down to the very bottom,
15:25:58 Mr. Peck, do you see it's labeled PTX-541(b)?
15:26:03 A. I see that.
15:26:05 Q. And are you able to read the real small text at the
15:26:10 very top?
15:26:11 A. Yes, I can read that.
15:26:13 Q. Okay. Sorry this is an eye exam test. Is this one
15:26:17 of the files that you were able to obtain from the Amazon

15:26:22 Alexa computer when you did your analysis?
15:26:24 A. Yes, it was.
15:26:26 Q. Okay. And --
15:26:29 MR. SMITH: At this point, Your Honor, we would
15:26:31 move PTX-541(b) into evidence.
15:26:40 MR. HADDEN: No objection.
15:26:41 THE COURT: Thank you. It is admitted.
15:26:41 (PTX Exhibit No. 541(b) was admitted into
15:26:43 evidence.)
15:26:43 BY MR. SMITH:
15:26:48 Q. Mr. Peck, is this code that we're looking at here in
15:26:53 Exhibit 541(b), is this part of the NLU of Alexa?
15:26:58 A. Yes, it is. In the patent name, NLU indicating its
15:27:06 an NLU. The name of this file is the contextual content
15:27:14 carry over dot JAVA. As well as I see references in the
15:27:19 actual code informing that, indicate that it's from the NLU.
15:27:23 Q. Does this indicate that Alexa uses this content?
15:27:31 A. Yes, that's correct.
15:27:32 Q. And were there other indications here in the code
15:27:35 that it's from the NLU?
15:27:37 A. Yes. For example, at line 10, there is an import
15:27:45 statement, you can see that it says
15:27:51 comment.amazon.Blushift.NLU, that indicates to you that this
15:27:53 is NLU source code.
15:27:59 Q. Let me go the line 63 of this. Do you see some is

15:28:17 something that is public class, contextual intent carrier
15:28:25 IPMPL. Do you see that?
15:28:25 A. Yes.
15:28:26 Q. What does that refer to?
15:28:30 A. This is the name of this class.
15:28:34 Q. And what does this class do generally.
15:28:39 A. So this is class is deciding which intents are to be
15:28:45 carried forward into the current set of intents being
15:28:48 considered.
15:28:58 Q. Why don't we go down here do you see line 71 where it
15:29:02 says intent carryover engine, rules engine, do you see that?
15:29:06 A. I see that.
15:29:07 Q. And based on your analysis, did you determine that
15:29:12 Alexa as an entire engine devoted to rules?
15:29:16 A. Yes, in this particular model, there is a particular
15:29:21 engine put in use that uses rules to perform its function.
15:29:24 Q. Is this file that we just looked at Exhibit 541(b)
15:29:30 another example of Alexa using manual rules in 2020 or
15:29:38 later?
15:29:39 A. Yes, it is.
15:29:45 MR. SMITH: Thank you, Mr. Peck. Pass the
15:29:47 witness.
15:29:48 THE COURT: Cross.
15:29:55 MR. HADDEN: You can leave that up.
15:29:57 CROSS-EXAMINATION

15:29:58 1 BY MR. HADDEN:
15:30:02 2 Q. Good afternoon, Mr. Peck.
15:30:04 3 A. Good afternoon.
15:30:06 4 Q. Now, you're not an expert in Natural Language
15:30:10 5 Understanding, are you?
15:30:12 6 A. I'm not an expert on Natural Language Understanding,
15:30:17 7 I'm a practicing engineer of over 25 years with experience
15:30:20 8 in software and hardware development.
15:30:22 9 Q. Thank you, sir.
15:30:23 10 And you're not offering any opinion to the jury
15:30:27 11 here today that Amazon infringes any of these patents, are
15:30:31 12 you?
15:30:31 13 A. No, I am not.
15:30:33 14 Q. And you're not offering any opinion to the jury today
15:30:36 15 that these patents are valid, are you?
15:30:38 16 A. No, I am not.
15:30:40 17 Q. Thank you.
15:30:41 18 Your assignment, if I understand it correctly,
15:30:47 19 was to answer questions from VB Assets's lawyers, is that
15:30:53 20 right?
15:30:53 21 A. Yes, that's correct.
15:30:54 22 Q. And Dr. Polish didn't ask you any questions, did he?
15:31:00 23 A. I received communication from counsel about what
15:31:04 24 parts to look at. Dr. Polish testified today about that
15:31:08 25 process.

15:31:09 Q. Okay. But when you were doing your analysis, you
15:31:11 weren't talking to Dr. Polish, correct?
15:31:13 A. No, I didn't talk to Dr. Polish directly. My
15:31:18 instructions came from counsel.
15:31:22 Q. Can we look at JTX-001, please, Mr. Patterson. Claim
15:31:30 13.
15:31:37 A. I have that in front of me.
15:31:48 Q. Now is it fair to say that between you and
15:31:55 Dr. Polish, you're the one who knows most about how Alexa
15:32:00 source code operates?
15:32:01 A. I would say I reviewed the source code for Amazon
15:32:06 Alexa.
15:32:06 Q. If we look at this Claim 13, did VB Assets' counsel,
15:32:12 did they ask you to go find out if there was Alexa source
15:32:17 code that accumulated short-term shared knowledge about the
15:32:20 current conversation?
15:32:24 A. So I recall my report addressing questions that
15:32:28 counsel presented me, those particular questions are in my
15:32:32 report. I don't recall this point being a particular
15:32:34 question being asked.
15:32:36 Q. Okay. So even though this claim requires computer
15:32:40 code that accumulates short-term shared knowledge, VB
15:32:45 Assets's lawyers didn't ask you to try to find that in
15:32:49 Alexa; correct?
15:32:50 A. No, the questions that I was asked, I don't think

15:32:55 1 relate to the claim language here. They were another terms
15:33:01 2 that were questions.
15:33:01 3 Q. If we look at the next element C, which requires
15:33:06 4 accumulating long-term shared knowledge, is it also the case
15:33:10 5 that VB Assets's lawyers didn't ask you to find out whether
15:33:14 6 Alexa has source code that accumulates long-term shared
15:33:17 7 knowledge?
15:33:17 8 A. I don't remember a question with those exact terms in
15:33:21 9 it.
15:33:21 10 Q. Okay. If we go to the next element, identify a
15:33:24 11 context associated with the utterance wherein the
15:33:30 12 conversational speech engine identifies the context
15:33:32 13 associated with the utterance from the short-term shared
15:33:37 14 knowledge and the long-term shared knowledge, is it also the
15:33:41 15 case that VB Assets's lawyers didn't ask you to find out
15:33:46 16 whether or not the Alexa source code does that?
15:33:48 17 A. Yeah, I don't remember a question posed in those
15:33:52 18 terms.
15:33:52 19 Q. Okay. And the same with element E, establish an
15:33:55 20 intended said meaning for the utterance, and then continues,
15:33:59 21 do you see the text there, sir, is it true that VB Assets's
15:34:02 22 lawyers didn't ask you to try to find out whether the Alexa
15:34:06 23 source code does this?
15:34:08 24 A. That's correct, I don't recall such a question being
15:34:11 25 asked.

15:34:12 Q. Okay. And if we go to the last element, generating a
15:34:15 response to the utterance, is it also true that the VB
15:34:19 Assets's lawyers didn't ask you to find out whether or not
15:34:22 the Alexa source code does this?
15:34:25 A. Yes, I don't recall a question like this being asked.
15:34:30 Q. Let's look at JTX-002, sir, the '703 patent, Claim
15:34:42 25. If we look here, sir, in particular there is an element
15:35:00 -- I can use this as well, thank you, Mr. Patterson.
15:35:02 There is an element here E, determine without
15:35:05 further user input after the receipt of the user input, a
15:35:09 context based at least on the one or more words or phrases,
15:35:14 do you see that element, sir?
15:35:16 A. I see that.
15:35:17 Q. Did VB Assets's lawyers ask you to find out whether
15:35:22 or not the Alexa source code did that?
15:35:25 A. I don't recall the question being put in these terms.
15:35:31 Q. Okay. If we go to the next element, F, sir, identify
15:35:34 without further user input after the receipt of the user
15:35:38 input, the product or service to be purchased on behalf of
15:35:41 the user based at least on the determined context. Did VB
15:35:46 Assets 's lawyers ask you to try to find out whether or not
15:35:49 the Alexa source code does that?
15:35:51 A. I don't remember a question like this being posed to
15:35:55 me.
15:35:56 Q. Okay. Let's go to the '176 patent, please,

15:36:01 1 Mr. Patterson. This is JTX-003, Mr. Peck. Look at Claim
15:36:11 2 40. Claim 40 has this piece at the end, an adaptive
15:36:16 3 misrecognition engine configured to determine that the
15:36:18 4 conversational language incorrectly interpreted the words or
15:36:22 5 phrases, and it continues, do you see that language,
15:36:24 6 Mr. Peck?
15:36:24 7 A. I see that.
15:36:26 8 Q. Did VB Assets's lawyers ask you to try to figure out
15:36:29 9 whether or not the Alexa source code performs that step?
15:36:33 10 A. No, I don't recall a question like that being posed
15:36:35 11 to me.
15:36:36 12 Q. Let's look at JTX-005, sir, this is the '097 patent.
15:36:43 13 If we could go to Claim 23. This claim it requires the last
15:36:55 14 element, requires interpreting the natural language
15:36:58 15 utterance based on the advertisement and responsive to the
15:37:01 16 existence of a pronoun in the natural language utterance,
15:37:05 17 determine whether the pronoun refers to one or more of the
15:37:08 18 product or service, or a provider of the product or service?
15:37:11 19 Did VB Assets's lawyers ask you to try to find
15:37:15 20 out whether or not the Alexa source code did that.
15:37:22 21 A. No, I don't recall a question put in those terms to
15:37:25 22 me.
15:37:26 23 Q. Thank you, sir. From your review of the Alexa NLU
15:37:32 24 source code, it's true, isn't it, that NLU generates
15:37:36 25 multiple interpretations at a given utterance, right?

15:37:41 A. Yes, that's correct.
15:37:42 Q. Thank you, sir.
15:37:44 And that context interpreter code that you were
15:37:48 shown earlier by counsel, that code takes the interpretation
15:37:54 that has already been part of the NBest list of
15:37:58 interpretation and reorders and re-ranks them, is that
15:38:04 correct?
15:38:04 A. Yes, that's one of the things that takes place within
15:38:07 that piece of code.
15:38:09 MR. HADDEN: Thank you for your time, sir. No
15:38:12 further questions.
15:38:12 THE COURT: All right. Redirect?
15:38:14 MR. SMITH: No redirect, Your Honor.
15:38:16 THE COURT: Thank you, sir. You're excused.
15:38:19 What's next?
15:38:20 MR. YOON: Yes, Your Honor, at this time, VB
15:38:23 Assets will play by videotape deposition, the deposition
15:38:28 testimony of Philippe DiCristo, who is the former senior
15:38:32 science manager at Amazon from 2016 to 2022.
15:38:37 THE COURT: Thank you.
15:38:41 (Videotape deposition of Philippe DiCristo:)
15:38:59 Q. Dr. DiCristo, could you please state your full name
15:39:02 for the record?
15:39:02 A. Yes. My full name is Philippe Jack Joel DiCristo.
15:39:08 Q. Doctor, you were the chief scientist at VoiceBox

15:39:12 1 Technologies, correct?
15:39:13 2 A. This has been my title at VoiceBox Technologies.
15:39:16 3 Q. And you were the VP of speech and NLU R & D at
15:39:24 4 VoiceBox Technologies, correct?
15:39:24 5 A. That is also the title I own at VoiceBox.
15:39:27 6 Q. Okay. And NLU is that an abbreviation for Natural
15:39:33 7 Language Understanding?
15:39:33 8 A. Yes, it is.
15:39:35 9 Q. And you were at VoiceBox from July 2003 to
15:39:40 10 September 2016.
15:39:43 11 Is that fair to say?
15:39:45 12 A. Yes. I don't remember if it's the -- the dates
15:39:49 13 specific there, but that -- that sounds right. Yeah, that
15:39:52 14 sounds accurate.
15:39:54 15 Q. And you were employee number nine at VoiceBox?
15:39:58 16 A. That's what I recall.
15:39:59 17 Q. And you were at VoiceBox for over 13 years?
15:40:02 18 A. I was at VoiceBox for a long time, yes.
15:40:08 19 Q. And over 13 years; correct, sir?
15:40:11 20 A. Over 13 years, probably.
15:40:13 21 Q. Now, at VoiceBox, you led a team of 100 engineers and
15:40:17 22 scientists designing and implementing new speech recognition
15:40:23 23 and natural understanding algorithms, is that fair to say?
15:40:28 24 A. Yeah, we were building a product that leveraged
15:40:34 25 technologies such as speech recognition and natural link --

15:40:40 1 language understanding.
15:40:41 2 Q. Dr. DiCristo, during the time at VoiceBox, you did
15:40:45 3 lead a large team of engineers and scientists; correct?
15:40:49 4 A. I did lead a large team of people, which included
15:40:52 5 engineers, scientists, program managers, TPM, and so on,
15:40:57 6 technical program managers and so on.
15:41:00 7 Q. What was VoiceBox Cybermind?
15:41:02 8 A. Cybermind was a product that we developed and that
15:41:13 9 leveraged existing technologies such as speech recognition
15:41:16 10 and Natural Language Understanding, and allowed people to --
15:41:19 11 to communicate with the machine and turn some command into
15:41:25 12 actions.
15:41:26 13 Q. When you prepared your LinkedIn profile, you believed
15:41:31 14 that the Cybermind was an Amazon-like system; correct?
15:41:37 15 A. Well, I think the "like" -- you know, when we say
15:41:41 16 like, of course you can take into different interpretation
15:41:45 17 here. It's open for interpretation.
15:41:47 18 But on the surface, they have like appearance.
15:41:50 19 Of course, their implementation could be quite different.
15:41:57 20 Q. But you described Cybermind as the Amazon Echo-like
15:42:02 21 system; correct?
15:42:04 22 A. Again, I think like is very -- it's Echo-like.
15:42:09 23 Right. Like, is there -- there is different flavors of
15:42:13 24 like.
15:42:13 25 And as -- as a person, as a customer without any

15:42:17 1 technical background, you could consider that they look
15:42:19 2 alike.
15:42:20 3 Q. And they operated alike, correct?
15:42:22 4 A. That would -- that's actually very vague. They are
15:42:26 5 -- I think they are -- they are very different on the -- on
15:42:29 6 the way they are implemented in terms of code is different.
15:42:32 7 The system that they use to recognize speech is different.
15:42:35 8 There a lot of differences that we could go into.
15:42:39 9 Q. But the Cybermind, the first version of it came out
15:42:42 10 in 2006, correct?
15:42:43 11 A. So the Cybermind product was actually never
15:42:46 12 productized, but it was demoed. And I definitely -- I
15:42:51 13 definitely think I remember even from the video that in
15:42:56 14 2006, we had something that looked like that.
15:42:58 15 Q. And the -- to your knowledge, the Amazon Echo product
15:43:04 16 was publicly announced in the 2014 time frame; correct?
15:43:11 17 A. I don't recall when exactly the Amazon system came
15:43:15 18 out, so I think it's hard for me to say. 2014 or 15 or
15:43:20 19 around that time frame.
15:43:23 20 Q. In 2013, IEEE ranked VoiceBox Technologies number 13
15:43:30 21 in patent power for the computer software industry. Do you
15:43:34 22 see that?
15:43:34 23 A. Yes, I do.
15:43:35 24 Q. And you put that statement in your LinkedIn profile;
15:43:39 25 correct?

15:43:39 A. I did put that statement in my LinkedIn profile.
15:43:43 Q. And so you understood that at least in the 2013 time
15:43:48 frame, the IEEE had ranked VoiceBox Technologies as having
15:43:53 the 13th most significant patent portfolio in the computer
15:43:59 software industry; correct?
15:44:00 A. So I remember actually seeing a document that was
15:44:03 actually provided by Mike Kennewick. That was a link where
15:44:07 the IEEE had ranked different companies of something they
15:44:12 call patent power. And that seems like a good thing to put
15:44:16 on your LinkedIn profile.
15:44:19 Q. So sir, you -- you began working after VoiceBox for
15:44:23 Amazon in October 2016. Is that fair to say?
15:44:26 A. I began to work for Amazon in October 2016.
15:44:31 Q. And when did you stop working for Amazon, sir?
15:44:35 A. I stopped working for Amazon on July 1st, 2022, --
15:44:40 January 1st. I'm sorry. I don't know why I said July 1st.
15:44:46 January 1st, 2022.
15:44:51 Q. Okay. Dr. DiCristo, looking at Exhibit 48, do you
15:44:56 see the top e-mail is from Rich Kennewick, dated April 30th,
15:45:03 2012?
15:45:03 A. Yes, I can see the E -- I can see that.
15:45:07 Q. And the subject there is: "Next
15:45:14 steps-VoiceBox/Intel." Do you see that?
15:45:15 A. Yes, I see that.
15:45:16 Q. And it says: "Bob, Michael, and Philippe; can we

15:45:21 meet today around 4:30 to discuss the two docs Intel has
15:45:28 sent us." Do you see that?
15:45:30 A. I can see that.
15:45:31 Q. And when you were working with Intel for VoiceBox,
15:45:34 was it your understanding that -- that Intel had an interest
15:45:37 in a speech-enabled system?
15:45:39 A. So I don't fully recall the -- the Intel project
15:45:44 itself. Maybe it's a bit far. But that's -- usually
15:45:48 customers were interested in some type of assistant-type
15:45:52 system based on NLP.
15:45:56 Q. If you'll look at the Intel ideal system document.
15:46:01 If you'll look at page 91949. One of the things that Intel
15:46:09 identified was: "Use of context."
15:46:17 Do you see that?
15:46:18 A. Okay. So can you repeat the question?
15:46:21 Q. Do you see the Intel document, in terms of
15:46:25 identifying the ideal system, number seven included: "Use
15:46:29 of context."
15:46:31 Do you see that?
15:46:32 A. Yes, I do.
15:46:34 Q. And that includes two types of history: "History of
15:46:38 the current conversation." And "history of prior
15:46:42 conversations by current user." Do you see that?
15:46:45 A. I can see that.
15:46:46 Q. And it says:

15:46:48 "Useful for personalization/preference setting
15:46:53 over time."
15:46:54 Do you see that?
15:46:55 A. I can see that -- that text from Intel, yeah.
15:46:58 Q. In VoiceBox, is agents use context that included both
15:47:03 the history of the current conversation and prior
15:47:07 conversations; correct?
15:47:10 A. It's a bit -- it's a bit vague. I mean, there are --
15:47:14 there are some implementation detail of it. I think I
15:47:17 mentioned that the agents had -- that the context basically
15:47:20 is all about prior information provided, which could be a --
15:47:25 a private conversation, and they -- they could take
15:47:28 advantage of it at that time.
15:47:31 Q. And so VoiceBox agents could take advantage of using
15:47:35 context that included prior conversations from the user;
15:47:38 correct?
15:47:40 A. Like I said, I think an agent had the
15:47:44 responsibilities to determine if a -- a request -- they can
15:47:47 respond to a request. And they would use information
15:47:50 available, which that we call these contexts.
15:47:55 Q. And that information available could also include the
15:47:59 prior conversations; correct?
15:48:01 A. Like, I don't remember exactly the whole
15:48:04 implementation detail of how -- what it would entail what
15:48:08 the context structure was. And as I said, it's more -- at a

15:48:12 1 high level, the context is general information of prior
15:48:15 2 event, basically, at the time of the request.
15:48:18 3 Q. And those prior events could include prior
15:48:22 4 conversations, correct?
15:48:23 5 A. Again, it's prior information that was available
15:48:26 6 before that -- that I know for -- that -- that conversation.
15:48:29 7 Q. And so prior information could include information
15:48:34 8 that came from prior conversations, correct?
15:48:39 9 A. I'm not sure I understand the question, because I
15:48:42 10 guess, like I said, it depends on how we define context. As
15:48:45 11 I mentioned, I define context as prior information that was
15:48:50 12 either provided by the system or by a person. And an agent
15:48:53 13 could leverage that information if they were designed to do
15:48:59 14 so, or if that was relevant.
15:49:02 15 Q. And looking at Item 10 in the Intel documentation, it
15:49:06 16 says: "Learning from interactions-wants system to learn
15:49:12 17 from previous conversation."
15:49:14 18 Do you see that?
15:49:15 19 A. I can see that.
15:49:18 20 Q. And one of the capabilities that the VoiceBox system
15:49:21 21 had and its agents had was the capability of learning from
15:49:25 22 previous conversations, correct?
15:49:26 23 A. So I guess it's hard for me to answer, because it's
15:49:31 24 -- like you see, learn is in quote marks, so learn has a lot
15:49:35 25 of vague notion of it. But I can say that actually, there

15:49:39 1 were no artificial intelligence in that -- in that -- at
15:49:44 2 that time.
15:49:44 3 Q. But the operation of the VoiceBox system could adapt
15:49:48 4 over time based on prior conversations, correct?
15:49:52 5 A. I don't recall that capability.
15:49:57 6 Q. And Dr. DiCristo, why don't you just take a moment to
15:50:01 7 look at Exhibits 55, 56, and 57. And when you've had a
15:50:06 8 chance to look at those text messages, I'll ask you a
15:50:09 9 question. Okay?
15:50:11 10 A. Okay.
15:50:12 11 Q. Dr. DiCristo, these are texts that you exchanged with
15:50:17 12 Kelley Price in the January 2006 to January -- sorry --
15:50:23 13 January 6, 2017, to January 9, 2017, time frame; correct?
15:50:30 14 A. Yes, this is correct.
15:50:35 15 Q. And these are screen shots of the texts that you
15:50:42 16 exchanged, correct?
15:50:43 17 A. This is correct.
15:50:45 18 Q. And sir, with regards to this, Ms. Price was a
15:50:49 19 VoiceBox employee; correct?
15:50:50 20 A. Kelley Price was a -- dealing with human resources,
15:50:55 21 as I previously stated.
15:50:57 22 Q. And so, she was the human resources at VoiceBox;
15:51:02 23 correct?
15:51:03 24 A. She was a human resources at VoiceBox. She was a
15:51:07 25 human resources employee at VoiceBox, yeah.

15:51:12 Q. And with regards to this text here, there is a
15:51:18 reference on Exhibit 57; right? There is a reference here
15:51:28 to -- looking at Ms. Price, it says: "Hey, can I forward an
15:51:34 invite to the Amazon event tomorrow night to folks who were
15:51:38 laid off, and maybe some who weren't?" Do you see that?
15:51:43 That's at 12:26 p.m.
15:51:46 A. Yes, I do.
15:51:49 Q. And I'd now like to mark from Tab 6 as Exhibit 58 a
15:51:55 document that shows eventbrite.com, bearing Bates numbers
15:52:02 VoiceBox 8212.
15:52:09 And Dr. DiCristo, do you see Exhibit 58 is a
15:52:14 document where the top says: "Eventbrite.com, do you see
15:52:20 that?
15:52:23 A. Yes, I do.
15:52:24 Q. And then it says, the description it says the
15:52:26 executive leadership team for Amazon's voice and advanced
15:52:29 shopping hosts an informal event to discuss our vision for
15:52:33 the future, with drinks and appetizers." Do you see that?
15:52:37 A. Yes, I do.
15:52:40 Q. Understood, sir. What I wanted to ask you was, you
15:52:44 understand that this event is the event that Ms. Kelley is
15:52:48 referring to in Exhibit 1 -- in Exhibit 57; correct?
15:52:53 A. Kelley Price is -- is -- probably was aware of that
15:52:57 event and was probably interested in sending some of the
15:53:00 people there.

15:53:01 Q. I'm just saying that the event -- the -- the Amazon
15:53:06 event that Ms. Kelley Price was referring to in Exhibit 57
15:53:10 was the event described in Exhibit 58 that lists you as an
15:53:15 attendee, but you did not attend; correct?
15:53:18 A. Well, as I mentioned, this was an event organized,
15:53:23 like, by my team at that time. I did not participate. And
15:53:27 I -- Kelley Price was probably interested in sending people
15:53:31 to that event.
15:53:33 (End of videotape.)
15:53:36 MR. YOON: Your Honor, at this time we would
15:53:40 like to move into evidence exhibit PTX-42 and 48.
15:53:45 MR. HADDEN: No objection.
15:53:47 THE COURT: Thank you. They're admitted.
15:53:49 (PTX Exhibits No. 42 and 48 were admitted into
15:53:49 evidence.)
15:53:49 MR. YOON: Your Honor, now at this time we would
15:53:51 like to play the video testimony of Rino Caruccio, who is
15:53:57 the vice-president of financial analysis for devices at
15:54:01 Amazon.
15:54:02 MR. HADDEN: I'm sorry, we have other exhibits
15:54:05 for that depo designation we would like to move in.
15:54:09 MR. YOON: I can do that for you.
15:54:11 MR. HADDEN: That would be great.
15:54:12 MR. YOON: Your Honor, this was the one we had
15:54:14 to add. The parties would also like to move into evidence

15:54:19 1 exhibit PTX-42, 48, 55, 56, 57, and 58.
15:54:28 2 MR. HADDEN: No objection.
15:54:31 3 THE COURT: All right. Thanks. They are
15:54:32 4 admitted.
15:54:33 5 (PTX Exhibits No. 42, 48 55, 56, 57 and 58 were
15:54:34 6 admitted into evidence.)
15:54:34 7 MR. YOON: And now, Your Honor, we would like to
15:54:36 8 play to deposition testimony of Rino Caruccio, who is the
15:54:41 9 vice-president of financial analysis for devices at Amazon.
15:54:46 10 THE COURT: All right.
15:54:48 11 (Videotape deposition of Rino Caruccio:)
15:54:50 12 Q. Mr. Caruccio, Exhibit 173 is an e-mail from someone
15:54:55 13 named Andy Vu dated September 8, 2021, with the subject
15:55:02 14 line, Re: OP1 outputs review, ten-year P&L, LTVLF and
15:55:11 15 IBVLF. Do you see that?
15:55:11 16 A. Okay. I have it, yes. I see the e-mail.
15:55:14 17 Q. Mr. Caruccio, are you familiar with an OP1, what an
15:55:21 18 OP1 is at Amazon?
15:55:23 19 A. Yes, I am.
15:55:24 20 Q. What is that?
15:55:25 21 A. OP1 is a mechanism within the Amazon company usually
15:55:30 22 held in September/October time frame whereby businesses,
15:55:40 23 products, review their strategy including investment needed,
15:55:45 24 financial projections, financial forecast for the next 12 to
15:55:51 25 18 months.

15:55:53 Q. What are some of the ways that this strategy benefits
15:55:58 the whole company system?
15:56:00 A. Again, in my personal opinion, when a customer
15:56:03 engages with a -- an Amazon device such as an Echo, there
15:56:09 are -- there is value in terms of music that they may
15:56:14 purchase from our Amazon music business. It would be
15:56:20 products that they may buy, you know, on our Amazon
15:56:24 marketplace and our consumer business. It could include
15:56:28 advertising that we would gain from advertising--
15:56:36 advertisers being willing to advertise on our devices.
15:56:40 There are sometimes some subscriptions that we provide or
15:56:45 garner. There are skills or third-party Alexa-enabled
15:56:56 devices that we might be able to engage with. There are
15:57:06 some modernization opportunities with, you know, partners,
15:57:15 stake holders, so there are -- those are some of the ones
15:57:23 off the top of my head that we would garner.
15:57:26 Q. Are you familiar with the term get big, with respect
15:57:29 to Alexa or EFD?
15:57:33 A. Yes, I'm familiar with the term get big.
15:57:36 Q. And what do you understand about it?
15:57:39 A. As -- as kind of a strategic overview of the Alexa
15:57:48 business and as a way of categorizing some of the key themes
15:57:52 that the Alexa business needs to think about, they decided
15:57:58 their -- categorized their strategy into three buckets, get
15:58:09 big, get close, get fit. And as -- and used that as kind of

15:58:15 1 a theme in terms of determining where to invest and where to
15:58:19 2 prioritize.
15:58:25 3 (End of videotape.)
15:58:28 4 MR. YOON: Your Honor, at this time, we would
15:58:30 5 seek to move into evidence exhibit PTX-173.
15:58:34 6 MR. HADDEN: No objection.
15:58:35 7 THE COURT: Thank you. It is admitted.
15:58:35 8 (PTX Exhibit No. 173 was admitted into
15:58:39 9 evidence.)
15:58:39 10 MR. YOON: At this time, Your Honor, the
15:58:40 11 plaintiff will now present the video deposition testimony of
15:58:45 12 Vinod Krishnan, who is the director of engineering and
15:58:49 13 science for Alexa Shopping at Amazon.
15:58:54 14 (Videotape deposition of Vinod Krishnan:)
15:59:00 15 Q. All right. In your folder, for the record,
15:59:03 16 Exhibit 78 is a document about with Bates numbers ending in
15:59:09 17 58040. Does the Alexa Shopping team use Wikis to document
15:59:15 18 the Alexa Shopping technology?
15:59:21 19 A. As a team, Alexa shop be uses many -- many
15:59:27 20 technologies to capture the thought process. Wikis could be
15:59:41 21 one of those.
15:59:42 22 Q. And so Exhibit 78, we're looking at a Wiki for a
15:59:46 23 document in the technology of Alexa Shopping, right?
15:59:50 24 A. It looks like this Wiki, or this document, captures a
15:59:55 25 version of -- of Alexa Shopping.

16:00:04 Q. And for the record, Exhibit 80 is an e-mail and
16:00:08 attachments, the Bates number on the first page ends in
16:00:15 342249.
16:00:17 Mr. Krishnan, do you recognize Exhibit 80 to be
16:00:22 an e-mail?
16:00:23 A. Exhibit 80, the first part of Exhibit 80 is an
16:00:28 e-mail, but then later on it turns into a document.
16:00:36 Q. Okay. And do you see that you received Exhibit 80,
16:00:42 your name is on there in the recipients?
16:00:45 A. Yes. A lot of names. Yeah.
16:00:56 Q. And you received the e-mail around August 30th of
16:01:01 2018?
16:01:02 A. Yes.
16:01:04 Q. And if you look at -- so Michelle Solano sent the
16:01:09 e-mail and she says please find attached the docs for
16:01:14 today's 2:00 p.m. PT 2019 NA consumer OP1: Alexa Shopping,
16:01:22 do you see that?
16:01:23 A. I see that.
16:01:24 Q. And there are 1, 2, 3 attachments that all have OP1
16:01:30 Alexa shopping, and 2019 OP1 in the file name?
16:01:36 A. I see all three of them.
16:01:38 Q. Okay. If you could now go to Exhibit 84, which has
16:01:43 an ending Bates number of 85508. And do you recognize this
16:01:52 to be a Wiki page from the Alexa Shopping Wiki that is
16:01:57 entitled "Dialog Platform"?

16:02:00 A. Yes.
16:02:02 Q. And Exhibit 87 is -- Exhibit 87 is the -- this is an
16:02:09 e-mail attaching the 2022 OP1; correct?
16:02:15 A. That's correct.
16:02:18 Q. An e-mail you received?
16:02:21 A. Yes.
16:02:26 Q. And just like the other OP1's, the 2022 OP1 was
16:02:34 presented to Amazon senior leadership?
16:02:39 A. It was presented to Amazon consumer and Alexa senior
16:02:44 leadership.
16:02:47 (End of video.)
16:02:50 MR. YOON: Your Honor, at this time we would
16:02:53 like to move into evidence exhibit PTX-78, PTX-80, PTX-84,
16:03:01 and PTX-87.
16:03:03 MR. HADDEN: No objection.
16:03:10 THE COURT: Is it clear what the corresponding
16:03:12 exhibit is with the script.
16:03:15 MR. YOON: I believe so, Your Honor.
16:03:17 THE COURT: I have been trying to follow it, I
16:03:19 just want to make sure that some day we don't have a
16:03:23 question as to what exhibit was talked about versus what
16:03:26 exhibit you entered.
16:03:27 MR. YOON: Yes, the exhibits that were talked
16:03:29 about in the transcript are the videos, are the ones that
16:03:33 are put into evidence.

16:03:34 THE COURT: With the same number? My problem is
16:03:37 there is a deposition number that's mentioned, no one knows
16:03:40 what the correlation is between the deposition exhibit and
16:03:43 the exhibit you just put in.
16:03:45 MR. YOON: Yes, Your Honor. It's our practice
16:03:47 in our deposition numbers, our exhibits and our exhibit
16:03:51 numbers are one in the same. We see sequentially numbered
16:03:54 our deposition exhibits and they should be the same as our
16:03:57 exhibit numbers.
16:03:58 THE COURT: Well, if at any point they're not,
16:04:00 make sure we have on the record, what the correlation is.
16:04:04 MR. YOON: Will do, Your Honor.
16:04:05 THE COURT: Thank you.
16:04:06 MR. YOON: And lastly, we are now playing the
16:04:08 deposition testimony of Scott Hayden who is the
16:04:10 vice-president of intellectual property at Amazon.
16:04:13 (Videotape deposition of Scott Hayden:)
16:04:19 Q. I think earlier in your testimony you mentioned you
16:04:21 looked at the Comverse patent purchase agreement, is that
16:04:28 right?
16:04:28 A. Yes.
16:04:28 Q. And why don't we -- figure out where that is, what
16:04:33 was the motivation to enter a patent purchase agreement with
16:04:40 Comverse?
16:04:40 A. So knowing that Alexa was coming and that we really

16:04:43 1 didn't have time to secure patents on our own technology, we
16:04:47 2 wanted to make sure that we had some assets that we could
16:04:51 3 counter claim or use in discussions or negotiations with
16:04:56 4 others. And we bought a small patent portfolio with
16:05:05 5 Comverse.
16:05:05 6 Q. Exhibit 158 is a document entitled objections and
16:05:10 7 responses of defendants to plaintiff VB Assets LLC's first
16:05:15 8 set of interrogatories, numbers 1 through 14?
16:05:20 9 A. I see that.
16:05:20 10 Q. And have you seen Amazon's interrogatory responses
16:05:24 11 before today?
16:05:26 12 A. I may have seen one subsection of one of these, but I
16:05:31 13 don't recall this one.
16:05:32 14 Q. Why don't we go to page 21 of Exhibit 158.
16:05:40 15 A. Okay.
16:05:41 16 Q. And the bottom of page 21, do you see there is a
16:05:44 17 sentence that says "Amazon became aware of U.S. Patent Nos.
16:05:49 18 '8,886,536; 8,073,681; and 9,015,049 in 2015." Do you see
16:06:05 19 that?
16:06:05 20 A. Yes.
16:06:06 21 Q. And do you understand that the three patents I just
16:06:10 22 quoted from the interrogatory response have been referred to
16:06:14 23 in this case as the '536, the '681, and the '049 patents?
16:06:22 24 A. I do.
16:06:24 25 Q. Do you know how Amazon became aware of the those

16:06:29 1 three patents in 2015?
16:06:32 2 A. My understanding is those were cited in prosecution
16:06:36 3 of our own patent applications.
16:06:40 4 Q. And does -- and do you have any other facts about the
16:06:43 5 context for how those patents came to be cited in Amazon
16:06:50 6 patents?
16:06:50 7 A. A patent examiner would have thought they needed to
16:06:54 8 be cited and cited them. Often times it's got the right
16:06:58 9 words in a search result. Sometimes it's something they are
16:07:02 10 aware of. I don't know why or how that came about.
16:07:06 11 Q. Do you know if the citation was from the patent
16:07:10 12 examiner or if it was from the patent prosecution counsel of
16:07:14 13 Amazon?
16:07:15 14 A. I don't. I don't know.
16:07:17 15 (End of videotape deposition.)
16:07:22 16 MR. YOON: Your Honor, at this time we would
16:07:24 17 like to move into evidence PTX-158.
16:07:27 18 THE COURT: So you got a forty-page document of
16:07:31 19 their interrogatory responses that has a whole bunch of
16:07:34 20 stuff in there that's not related to what he just talked
16:07:37 21 about. Why are we doing this?
16:07:39 22 MR. YOON: I apologize. That was what was used
16:07:42 23 at deposition. We can excerpt it so it's just one or two
16:07:45 24 pages.
16:07:46 25 THE COURT: I'm not admitting this. You can

16:07:49 1 figure out what you want to excerpt it to, talk with them
16:07:52 2 and we'll talk about it on Monday.
16:07:54 3 MR. YOON: Understand, Your Honor.
16:07:56 4 At this time we'll be calling Mr. Brett Reed.
16:07:59 5 THE COURT: How about we give the jury a break,
16:08:01 6 let you go a little early. I think we're making good time.
16:08:05 7 We'll be fine. And it's Friday afternoon and some of us are
16:08:09 8 getting tired. So I'll just remind you don't talk to anyone
16:08:13 9 about the case, which you probably want to, and have safe
16:08:18 10 travels and we'll see you on Monday morning.
16:08:20 11 COURTROOM DEPUTY: All rise.
16:08:22 12 (Jury exiting the courtroom at 4:08 p.m.)
16:08:44 13 THE COURT: All right. So I don't think that's
16:08:47 14 going to be a problem, but I'm not putting in a bunch of
16:08:52 15 stuff that I don't know how it's going to be cited some day
16:08:55 16 as evidence, so let's figure out how to deal with that one.
16:08:59 17 Anything we need to talk about.
16:09:01 18 MR. YOON: Only procedurally, Your Honor, two
16:09:04 19 things. One is that I had a discussion with the local
16:09:10 20 counsel. I think we were trying to have a discussion this
16:09:12 21 weekend of how to make things more efficient for the Court
16:09:16 22 on Monday, as counsel indicated we don't have a claim
16:09:19 23 construction issue anymore, and so hopefully this weekend we
16:09:23 24 can come back with a proposal to you to make Monday evening
16:09:26 25 or Tuesday much more efficient.

16:09:28 THE COURT: All right. I appreciate that.

16:09:31 Anything else?

16:09:31 MR. HADDEN: No, Your Honor.

16:09:32 COURTROOM DEPUTY: Court is adjourned.

(Court adjourned at 4:09 p.m.)

I hereby certify the foregoing is a true and accurate transcript from my stenographic notes in the proceeding.

/s/ Dale C. Hawkins
Official Court Reporter
U.S. District Court