08:06:45

IN THE UNITED STATES DISTRICT COURT

FOR THE DISTRICT OF DELAWARE

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

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

844 King Street
Wilmington, Delaware

BEFORE: THE HONORABLE MARYELLEN NOREIKA
United States District Court Judge

APPEARANCES:

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

Counsel for the Plaintiff

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

- - - - - - - - - - -

08:34:34 THE COURT: All right, everyone. Please be seated. So we have some objections. I know you guys gave some stuff to Mr. Buckson, but I haven't looked at it, you can do that all on the clock. Okay. What do we have?

08:34:50 MR. YOON: Your Honor, there are some issues that have been made, but we did want to report to the court that the parties did work through all the issues with Dr. Polish.

08:34:58 THE COURT: We're wasting the jury's time, I do

08:35:02 appreciate that and thank you, but let's not waste time.

08:35:06 MR. SALMASTLIAN: Good morning, Your Honor,

08:35:09 Vigen Salmastlian. There is an issue with a source code

08:35:14 exhibit that --

08:35:15 THE COURT: I have this, this is the only

08:35:16 document that I have right now. So do you have an objection

08:35:19 to this?

08:35:21 MR. WARE: Yes.

08:35:22 THE COURT: Let's deal with it.

08:35:24 That was the one that was actually timely made,

08:35:29 so let's deal with that first.

08:35:30 MR. WARE: Jeffrey Ware for defendant.

08:35:33 And here Amazon is using -- sorry, let me just

08:35:39 put this -- Your Honor, may I put this on the Elmo?

08:35:43 THE COURT: Yes.

08:35:46 MR. WARE: So we are re-raising Amazon's Daubert

08:35:51 to the portion of the Daubert motion that was denied with

08:35:56 leave to re-raise at trial. It may be appropriate, or more

08:36:02 appropriate for this Court to consider this issue after the

08:36:06 damages experts have testified and the Court has gotten more

08:36:11 information regarding the specifics of the methodology

08:36:14 Amazon is objecting to, and so for the appreciating the

08:36:21 issue of time here, Amazon simply re-raises--

08:36:25 THE COURT: No. No. You're not re-raising

08:36:28 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 re-raise what I did, you know, weeks ago or even a week ago, I have no idea what that is, okay? So if you want to make an objection, you're -- it's worth your time and you make it. Because I don't know what the problem is with it.**

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

Amazon still objects to the underlying cost savings methodology as raised in the motion and specifically --

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

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

compared it -- which right now they accuse the '681 patent of infringing, and then he took Alexa Shopping's turn rate which plaintiff accuses the remaining patents of infringe -- or Amazon of infringing the remaining patents, and because at one point in time in 2018, there was a 1.5 turn difference, and he further opines that there was a certain thing that happened with Amazon Shopping, a simple CX program that, at that time, made it so that he could support the valuation for his 1.5 turns benefit analysis he compared Alexa Shopping with Alexa active customers and so the basis of his calculation if you look here on this -- on the demonstrative is that he said because Alexa Shopping had poor customer experience prior to this time, you can see when Alexa Shopping, which is in red, goes way up in late 2018, that the benefit of 1.5 turns could be calculated based on the difference here pre the September 18 of Alexa and Alexa Shopping user growth.

So Amazon's first objection to cost saving methodologies is no cost saving methodology at all except by the Federal Circuit, because here at this time, plaintiffs have accused Amazon of infringing both Alexa, and Alexa Shopping, you can't take the difference of growth rates in both to determine the cost savings of using either the '681 patent for Alexa, or the remaining asserted patents for Alexa Shopping.

08:40:23 So that is the first basis of Amazon's objection that cost savings is supposed to be as recognized by the Federal Circuit, the difference between infringing and a non-infringing use, usually a non-infringing alternative or sometimes the alternative of actually doing nothing. In this case, they're comparing two infringing products --

THE COURT: But infringing of different patents, so it's an infringing of the '681, and non-infringing of the '681, but infringing of the others?

MR. WARE: Well they accused '681 actually of infringing all -- or they accused both Alexa and Alexa Shopping of infringing all of the asserted patents, but only specifically Alexa Shopping for infringing the other asserted patents.

THE COURT: So they're comparing the red line is alleged infringing '681, and the blue line is a different product that also is alleged to infringe the '681?

MR. WARE: Yes, it's the Alexa Shopping growth rate. I'm sorry, the blue is Alexa, and red is Alexa Shopping.

THE COURT: Okay. What's your next basis?

MR. WARE: The next basis is it's an apples to oranges comparison, and in order to effectively assert that because Alexa worked well at the time because of the patents, and even though Alexa Shopping wasn't working well

08:42:06 before this change in 2018, even though again it was using
08:42:10 -- it was practicing the other patents, that the 1.5 turn
08:42:14 rate --
08:42:15 THE COURT: You keep saying the other patents.
08:42:17 So --
08:42:22 MR. WARE: So the voice ad patents and the '703
08:42:26 patent.
08:42:26 THE COURT: Yes, I know. Where I'm getting
08:42:29 confused is what is -- you just -- when I thought that's
08:42:33 what you were talking about before, '681 versus the other
08:42:36 patents, you said no, no, everything infringes, is alleged
08:42:41 the infringe the '681, so now we're talking about something
08:42:44 different, right? Now you're saying one thing is the '681
08:42:47 and another thing is the other patents?
08:42:49 MR. WARE: Well, Your Honor, the '681 patent
08:42:53 claims the more -- just most generally may call it
08:43:00 contextual understanding using short and long-term
08:43:03 knowledge. And so, they assert that all Alexa interactions
08:43:09 infringe the '681 based on that, whether it's in the Alexa
08:43:14 Shopping domain, or in any other Alexa domain.
08:43:17 With respect to the voice ad patents and the
08:43:20 '703 patent, they only accuse Alexa Shopping functionality
08:43:27 of infringing those patents.
08:43:30 THE COURT: Okay. So with that understanding,
08:43:32 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 **MR. WARE: And the third ground is that he is**
08:45:30 **pointing to a single Alexa Shopping document, OP1 and the**
08:45:35 **purported improvement of Alexa Shopping after certain**
08:45:40 **improvements were made, but he's applying the value of those**
08:45:45 **improvements to Alexa overall. Where again, specifics of**
08:45:51 **Alexa Shopping and Alexa are very different. And on top of**
08:45:58 **that, he fails to show that even when you see the red line**
08:46:03 **moving up, he fails to show that the red line moving up is**
08:46:07 **actually caused by reduction of turns in the patent, or**
08:46:10 **reduction of turns in dialogue which is the purported**
08:46:15 **benefit that Mr. Reed claims, and is trying to monetize for**
08:46:20 **plaintiff.**
08:46:21 **And so again, that's -- it makes no logical**
08:46:26 **sense, and we would object under Rule 702 and Daubert for**
08:46:31 **that.**
08:46:31 **THE COURT: Now, when you said before that this**
08:46:36 **might make more sense to raise after the experts testify,**
08:46:41 **wouldn't the expert already have testified about this, so**
08:46:44 **doesn't that suggest that this is not really a reliability**
08:46:48 **issue if you're willing to let the expert talk about this**
08:46:50 **before I rule on it?**
08:46:52 **MR. WARE: Well, Your Honor, we believe that you**
08:46:56 **already considered this issue when you did rule on the**
08:46:58 **initial Daubert motion, and you had suggested that this may**
08:47:01 **be an issue for cross-examination. But you granted Amazon**

leave to re-raise it at trial. We think and know that if a damages verdict issues by the jury in this Court, that the Federal Circuit will overturn because again, this makes no -- first of all makes no logical sense, and it also fails to follow any recognized cost savings methodology that's ever been accepted by the Federal Circuit.

THE COURT: I guess we'll have to find out. Let's hear. And I really want to hear mostly on the first ground which is how can this -- how can this have anything to do with infringing the benefits of the patent when you are comparing infringing versus infringing.

MR. MACDONALD: Thank you, Your Honor. Matt Macdonald on behalf of VB Assets. Let me just get straight to Your Honor's question.

THE COURT: That would be good without even the present preface of telling me, just get there.

MR. MACDONALD: The answer is, we cannot run an experiment inside Alexa that compare what would happen in the but for world, where there is infringement and non-infringement. We can't do that, we don't have that ability because our position is that Amazon continuously infringes, and we don't control their system.

So what our expert did was look through Amazon data and Amazon documents to attempt to find evidence that would allow him to proxy for the effect of the infringing

alternative.

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

The second piece of evidence --

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

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

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

08:50:11 patent, right?

MR. MACDONALD: Yes, Your Honor.

08:50:12 THE COURT: The one-and-a-half turn, is that a patented feature?

08:50:13 MR. MACDONALD: No, the one-and-a-half turns is what Dr. Polish will testify is the consequence of using the patent.

08:50:15 THE COURT: But using which patent?

08:50:16 MR. MACDONALD: All of the patented technology.

08:50:20 THE COURT: Okay. So that's where I'm missing this. You get the benefit of the one-and-a-half turns from that, then how -- don't the blue things also have the benefit of the patents because you're accusing them of infringement?

08:50:22 MR. MACDONALD: They absolutely do. So the question that Dr. Reed -- excuse me, that Mr. Reed attempted to answer was what evidence in Amazon's files that I have, that would allow me to quantify the impact of this 1.5 turns. And we can't run an experiment --

08:50:23 THE COURT: I don't understand how you can say the effect of infringing the patents is you get this benefit when you're accusing things of infringing the patents that don't apparently have that benefit.

08:50:25 MR. MACDONALD: I think the answer is that we don't have -- we don't have -- we don't have real world

08:50:27

08:50:29

08:50:33

08:50:37

08:50:40

08:50:41

08:50:43

08:50:47

08:50:51

08:50:56

08:50:58

08:51:00

08:51:03

08:51:06

08:51:10

08:51:12

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 non-patented in this slide, and you're not.

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

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

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

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

08:53:57 having that extra turn gives.

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

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

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

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

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

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

08:54:17 on that.

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

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

08:54:28 Alexa Shopping --

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

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

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

08:54:40 right?

08:54:40 MR. WARE: Yes.

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

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

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

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

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

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

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

jury has already heard this, also lends credence to the fact that this is an issue for cross rather than an issue of reliability.

All right. So what's the next issue that we have?

MR. SALMASTLIAN: Your Honor, the issue, the next issue is related to the document AMZ_VB_SCPRINT 74 to 77, and the issue here is this is a source code document that Mr. Peck raised for the first time in his reply report, but it was not for any issue related to any patent that's currently at issue in this case. He restricted his opinion to be a response to paragraph 459 of Mr. Johnson non-infringement report. And that part of Mr. Johnson's opinion relates to '536 patent, Claim 3. That claim has been long dropped, and that patent has been dropped from this case. And so we have no idea what Mr. Peck is going to say about this source code document with respect to any patent that's currently at issue in this case.

MR. SMITH: Your Honor, Mr. Peck is a source code expert and all he did was opine about the operation of the source code. His opinion is agnostic with respect to any particular patents --

THE COURT: So where did he say he's going to rely on this document with respect to any issue that is currently in the case?

09:01:11 MR. SMITH: In his reply report --

09:01:14 THE COURT: Show me where, because he's saying it's not relevant to any patent left in this case.

09:01:20 MR. SMITH: So in his reply report on page -- towards the bottom of page 10, if we can put that on the screen. It's exactly -- probably 20 percent of his report is just on this particular file. So of the ten page report, two pages are about this file. And all he's going to testify to is exactly what he says in his report. It's not agnostic as to any particular patent, it's just Dr. Johnson is saying that the Amazon Alexa works in a certain way, and he was disputing based on code.

09:01:57 THE COURT: But when you said it's agnostic, he said this is response to '459. What was '459?

09:02:04 MR. SMITH: So if we go to '459, this was about, then I'll quote, '459, Johnson says "Polish also accuses "a dialog management feature that determines a context by passing." And he goes on from there. And then he talks about the context interpreter. And the context interpreter comes up again and again and again with respect to all of the VoiceBox patents that are being asserted. So Dr. Johnson has a particular paragraph. I'll say that his report was 500 pages, and every single argument incorporates everything else. So it's basically a spiderweb of opinions that cover everything seemingly.

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

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

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

09:05:35 THE COURT: I am going to let this go, to the
09:05:37 extent that it's just saying look, there is this rules-based
09:05:40 context interpreter, if he starts getting into much more
09:05:44 specifics that weren't disclosed, then you can object again
09:05:48 and I'll reconsider. But I think it's fair to say that with
09:05:52 respect to the source code, it was disclosed that he was
09:05:55 going to note that there is a rules-based context
09:06:00 interpreter in the device.

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

09:06:06 MR. SALMASTLIAN: There is just one other issue
09:06:07 Your Honor, and it might be resolved at this point, but
09:06:10 there was a late add source code exhibit this morning which
09:06:14 is Amazon -- go ahead.

09:06:16 MR. SMITH: We'll withdraw that in light of your
09:06:19 ruling.

09:06:19 THE COURT: All right.

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

09:06:22 THE COURT: All right. Mr. Buckson, do you want
09:06:24 to get the jury.

09:06:26 MR. YOON: Your Honor, just one procedural point
09:06:27 just to let you know Your Honor, we're going to be starting
09:06:30 with Mr. Freeman. The parties agreed to add some additional
09:06:34 testimony to Mr. DiCristo, so we'll start with Mr. Freeman
09:06:39 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 the jury?**

09:10:46 A. **Sure. My name is Tom Freeman, I was born in Seattle Washington, that's the other Washington, and that I lived there my entire life with a short stint in Anchorage, Alaska. I was a co-founder of VoiceBox technologies. And I worked there for nine years. And the reason I am here today, I guess, is that I was responsible for the invention of some of the important patents in this particular case.**

09:11:19 Q. **Mr. Freeman, if we could just go over your background 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 value of the technology to the customer was extremely important, so I think you would call this technical sales**

and technical product management.

Q. And just generally, what sort of products were you selling back then?

A. Everything from database publishing to relational databases, to holographic beam forming metamaterial techniques, it was quite a broad world of technologies.

Q. Do you have any formal engineering training?

A. Oh, no, I do not, sir.

Q. Did you learn technology on the job?

A. I did indeed, sir, to be able to interpret it is extremely important.

Q. And Mr. Freeman, do you know the Kennewick brothers?

A. I do know the Kennewick brothers, I know them well.

Q. How did you first come to meet them?

A. Rich and I worked together in a company called Interleaf, which produced technical publishing software for the '757, '767, and eventually the '787, and in that he worked for me, and eventually I worked for him at VoiceBox.

Q. And you mentioned VoiceBox. How did it come about that you guys founded that company?

A. So the idea of being able to talk to your computer, talk to the internet, be able to do information retrieval I found extraordinarily compelling. And so much so that when we eventually began to form the company, it became clear that, you know, early money is like yeast, and you have to

be able to, be able to address the marketplace, and in such I invested the first $400,000, which I might add caused a little distress in the marriage.

Q. For you was that a lot of money?

A. That was a significant amount of money.

Q. And when you were going to decide to found VoiceBox with the Kennewick brothers, and put all that money in of your own money, were you aware of any other voice-based computer solutions that were out there?

A. There were a lot of people attempting there, there was command and control systems, they would follow the menu and say exactly the right words or nothing good would happen.

Q. What do you mean by command and control?

A. So if I wanted to talk to a pharmacist, you had to be able to say the exact words on the menu tree in order to get to that particular destination. As opposed to just saying one shot and be able to say I would like to talk to a pharmacist about my -- about the drugs that are being assigned to me.

Q. When you're saying talk to a pharmacist, is this something on a computer or something like a phone?

A. I'm sorry, this would be a telephone tree, the primary form that this took was a telephone, telephone based 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?**

A. **Yes, I'm seeing it.**

Q. **Okay. Well before we play this. Mr. Freeman, are you familiar with the video of PTX-543?**

A. **I am indeed, sir.**

Q. **Did you participate in the filming of that video?**

A. **I was what they call in the business a talking head, yes, sir.**

MR. SMITH: Your Honor, at this point plaintiff would like to introduce into evidence PTX-543.

MR. HADDEN: No objection.

THE COURT: Thank you. It is admitted.

(PTX Exhibit No. 543 was admitted into evidence.)

MR. SMITH: Mr. Smith, could you start playing 543, please?

(PTX-543 video played.)

MR. SMITH: Can you pause that?

So Mr. Freeman, who is that individual talking there in the video?

A. **That individual would be me, sir.**

Q. **And here in this -- in the video, what were you talking to?**

A. **So that was the Cybermind, what we call the head, the whole system together was called Cybermind because it included the software, but that was the Bluetooth speaker**

and mic.

Q. Do you recall approximately when this video of you and the Cybermind head was taken?

A. Yeah, it was about 2006 that the NBC affiliate approached us and said this would be useful to do.

MR. SMITH: Mr. Smith, if you could continue playing this.

(Video played.)

BY MR. SMITH:

Q. So Mr. Freeman, I think at the end of that video, the newscaster mentioned that there was going to be the hand held thing out by Christmas. And what was -- was VoiceBox able to do that?

A. We were not at the time.

Q. Why not?

A. There were compliments of issues, one was what's called the bottom, the buildable materials to build that thing, was expensive. Number two, there wasn't like a post sales revenue model to be able to bring it down. And three, Toyota, wonderful amazing company, approached us with a $5 million deal and said we want that technology in our cars, so would you please defer your resources in order to do it. Very difficult to turn down that kind of offer.

Q. What did VoiceBox do then?

A. What we did was, we said all right, the NLU and the

core technology we will continue to develop over time, and we will make better but we will figure out how to shrink this, in order to fit it into what's called the head unit of an automobile which is the navigation unit, the infotainment unit, the radio, heater, all that kind of stuff.

Q. At this point, when you had decided to work with Toyota and put the Cybermind technology into the Toyota cars, at that point, did you give up on Cybermind?

A. No. Cybermind was the, the image that we were chasing, the notion of why this would be compelling, again, why would you cross the street. We can make it this good in an automobile, but we can make it this good in a living room.

Q. Did you work on any technical improvements for Cybermind in that time period?

A. The way software works, as you well know, is you just don't put a formula into -- into software, instead you iterate, and you iterate, and you get better, and you get better, they're called drops, they're called -- they're called release candidates, but yes, to answer your question, completely.

Q. Did you work on any potential patents at that point?

A. We were working on ways, we believe very much in the patent process, so we were working on ways of improving our technology both from a recognition point of view, from an

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

09:24:03 **primary idea here.**

09:24:06 Q. **Do you see here on the Item 22, there it says**

09:24:11 **October 16, 2006, do you see that?**

09:24:13 A. **I do.**

09:24:14 Q. **And is that consistent with your recollection of when**

09:24:17 **you filed for this patent?**

09:24:18 A. **It's consistent with the video that we were listening**

09:24:25 **to.**

09:24:27 Q. **And at this time, had you thought about the idea of**

09:24:29 **using long-term shared knowledge?**

09:24:33 A. **Absolutely. It turns out that when you and I have a**

09:24:38 **conversation, two things happen. One is we understand**

09:24:41 **what's happening in the conversation right now. But**

09:24:45 **secondly, and importantly, you have a basis of past**

09:24:49 **experience with me that makes a lot more sense. You say I**

09:24:54 **need to get to the freeway, what freeway are you talking**

09:24:58 **about, are you talking about I95, or Interstate 5, there is**

09:25:02 **a big difference between those two. And that by having a**

09:25:06 **long-term memory, we understand what freeway means in this**

09:25:10 **particular context.**

09:25:11 Q. **At the time that you filed the patent, were you aware**

09:25:14 **of any other voice system that had that kind of long-term**

09:25:17 **memory?**

09:25:18 A. **Not to my knowledge, sir.**

09:25:20 Q. **And what was your contribution for the '681 patent,**

what role did you play?

A. So my role was to sit there and give it the litmus test, would this be used, why would it be used. Are we just having a technology chasing an idea, or is this actually useful in the different use cases that we were talking about, in the car, in the living room and different places.

Q. And Mr. Smith, if could go to column 11, start around line 10, please.

And Mr. Freeman, I moved to a section of the patent, the '681 patent. Do you see that on your screen there?

A. I do.

Q. And in this section of the patent, it says, for example, a user utterance of quote, what about traffic? Close quote, may have a distinct meaning in various contexts. Do you see that?

A. I do.

Q. Is this part of the patent that you worked on?

A. Yes, it is.

Q. And what was the idea behind this example about traffic?

A. Well, if you were to look at me and say completely cold, what about traffic, I would look at you and say I think you're crazy. But to be able to establish a context about what traffic means, are we talking about music, are we

talking about a movie, did this request come from a movie agent, did this request come from a smart television, did this request come from a navigation system, that helps narrow down what traffic means. It doesn't mean you wouldn't have a follow-up question because follow-up questions are useful -- humans do this all the time. But you can make certain assumptions based on where it's coming from, and what the user is doing, what "traffic" means in this situation.

Q. And would the long-term shared knowledge be helpful for the system to figure out what was meant by traffic?

A. Oh, yes, very much so. How have you used this word before, and how were you using this word before. And so you have not only the history of those -- of the usage of the word traffic, but also of the different elements within traffic itself.

Q. And at that time when you invented the '681 patent, what was it going to do for Cybermind in your view?

A. It was -- so the power of the Cybermind was to make -- to take the friction out of a conversation with a human. So it's -- a lot of this user interface stuff is, I don't want to repeat everything I just said, I want you to be smart enough to understand what traffic means. What is traffic like on the way to the hospital? Well if it understands me, it understands the hospital I go to, every,

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

since that particular technology. But the aspect is completely the same.

Q. And so and what was -- so you're saying as an example, you could tell Cybermind to play a song, and then Cybermind might suggest that you might want to buy the ring tone?

A. Or I could ask it, right? So one of the key notions here, and how advertising expanded in the course of the time that we were dealing with this, was that instead of it being a one-way flow from the source -- from the radio, if you will, into the consumer, if the world switches around where I could say gosh, could I get a ring tone, that would be pretty cool, and that advertising becomes interactive.

Q. Why don't we now pull up JTX-005, please, Mr. Smith. And Mr. Freeman, do you see that JTX-005, the upper right-hand corner has the last three digits '097?

A. I do.

Q. And this one also lists you as a co-inventor with Mike Kennewick, do you see that?

A. I do.

Q. What was the '097 patent about?

A. The '097 was an -- there is a term of art you folks use, an expansion, no, it was an expansion of our patents that we had in the earlier ones, so it's a continuation, that was the word I was looking for, it was a continuation

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

of the elements that go into Bluetooth, and all the elements that go into the head, and all of the elements that you have to require in terms of your R & D, and we were looking at what is called a BOM cost, building and materials cost, that once we put it in the channels was going to make it prohibitively expensive for your average consumer, for your above average consumer to be able to purchase. So the notion of being able to subsidize that by being able to have an ad-free, or an ad-based system, being able to have people be able to buy products in which we would get a percentage of the opportunity, allowed us to have a post-sales business model that would allow us to sell for less money than we were currently thinking about doing.

Q. So would the version of Cybermind that had ads on it be less expensive, or more expensive than the version of Cybermind without the ads?

A. It would be less expensive. And consumers became used to this kind of model for this time period that you can buy yourself down to a non-ad environment.

Q. When did you leave VoiceBox?

A. 2010.

Q. And when you left VoiceBox in 2010, had you been able to implement the '681 patent with the long-term memory that we talked about, and the '176 and '097 patents with the advertising in a physical product at that point when you

left?

A. Okay. I'll respond carefully, because you said in a physical product to be released. The answer to that is no. But the answer is, did we have point releases in engineering that was getting us there, that software development is not magic, sir, it is a lot of I wonder if that worked, let's try this and you do it again and you do it again. And that's -- that iterative process is the process by which software is realized.

MR. SMITH: Thank you, Mr. Freeman. Appreciate your time. I'll pass the witness.

THE WITNESS: Thank you.

THE COURT: Cross-exam.

CROSS-EXAMINATION

BY MR. HADDEN:

Q. Good morning, Mr. Freeman.

A. Good morning.

Q. Just to go back a little to your background, I understand you went to the University of Washington. What was your degree in?

A. So I had a liberal arts degree as an undergraduate, and a masters of business administration from the Foster School which is a prestige school at the University of Washington.

Q. And just to be clear, so you quit VoiceBox in May of

09:36:15 **2010; is that correct?**

09:36:18 A. **That is correct.**

09:36:18 Q. **And over the ten years you spent at VoiceBox before**

09:36:23 **you quit, it's correct that VoiceBox spent more money than**

09:36:28 **it received; correct?**

09:36:30 A. **I think it -- I'm not a subject matter expert in this**

09:36:35 **field, but I would say -- I'm not a subject matter expert in**

09:36:40 **that field.**

09:36:40 Q. **So you don't know?**

09:36:42 A. **It would -- I would answer yes to your question.**

09:36:48 Q. **Now, when you left VoiceBox in 2010, VoiceBox didn't**

09:36:55 **have -- had never enabled eCommerce transactions through**

09:37:01 **voice, correct?**

09:37:04 A. **Not true.**

09:37:06 Q. **Excuse me?**

09:37:07 A. **Not true.**

09:37:09 Q. **So can --**

09:37:10 A. **It depends what you mean, was it commercialized,**

09:37:13 **absolutely true, it was never commercialized.**

09:37:16 Q. **Do you have a copy of your deposition, it should be**

09:37:19 **that binder there, do you remember when you were deposed in**

09:37:22 **this case, sir? Were you able to find it there?**

09:37:29 A. **Deposition transcript? Sure.**

09:37:32 Q. **I would ask you to look at page 47, lines 9**

09:37:37 **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 enabled eCommerce transactions through voice, and I interpreted your previous question is, were we actually 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 to actually do it before you left in 2010, and your answer was, you could not, right?**

09:38:22 A. **Do it is a very uncomfortable word. In that were we able to demonstrate it? Yes. Were we able to commercialize 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 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 this patent, did you?**

09:38:57 A. **So the answer to your question is technical contributions of -- well, did I develop an algorithm by which engineers would at some point take and build a typical solution, the answer is yes. Was I an engineer who built a**

09:39:03

09:39:10

09:39:18

09:39:22

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

word.

Q. Let's look at the patent again, please, Mr. Patterson, JTX-003. And let's look at Claim 40.

Did you invent a conversational language processor?

A. Did I invent a conversational language processor?

Q. Yes, sir.

A. I did not.

Q. Did you invent an adaptive misrecognition engine, sir?

A. So did I outline what the algorithm would be to make it useful. The answer is yes. Did I write the code, the answer is no.

Q. Can you describe for me what the algorithm is that you developed for an adaptive misrecognition engine?

A. So it is about where the utterance comes from and what the purpose of the utterance is. For instance, if you're in a distributed environment, and you have a navigation system and that you're picking up the words about traffic from a navigation system, that's telling you in an enormous amount of information about what the word traffic means in that particular context, poison IV and Poison Ivy are phenetically extremely close to each other, but one comes from medical, and one comes from a gardening app, then you have a clue as to what that information means. So even

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

09:46:08 **BY MR. HADDEN:**

09:46:09 Q. **So this is a paper that you wrote with Mr. Baldwin,**

09:46:14 **another inventor on the '681 patent. Is that correct?**

09:46:17 A. **That is correct.**

09:46:18 Q. **And you wrote this paper with Mr. Baldwin, it's a**

09:46:25 **marketing material for VoiceBox, correct?**

09:46:27 A. **It was a vision piece, correct.**

09:46:34 Q. **If we look at page 6 of this paper. Can you move to**

09:46:41 **that Mr. Patterson?**

09:46:42 **At the top is a heading cooperative**

09:46:45 **conversation. Do you see that, sir?**

09:46:46 A. **I do.**

09:46:48 Q. **And then it goes on and it says what is a cooperative**

09:46:53 **conversation, and you wrote in this paper that the**

09:46:56 **cooperative principle was first described by Paul Grice in**

09:47:01 **1975. Do you see that?**

09:47:04 A. **I do.**

09:47:04 Q. **And the cooperative principle is what you were**

09:47:07 **talking about with your counsel, where people need to work**

09:47:12 **together to get to some goal in their conversations,**

09:47:14 **correct?**

09:47:15 A. **Correct.**

09:47:15 Q. **And that idea that you're talking about with your**

09:47:19 **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 -- so if we highlight the paragraph that begins a cooperative conversation, and the material just below it, it would be the second paragraph, Mr. Patterson.**

09:47:40 **There is --**

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

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

09:47:52 **THE COURT: Don't talk to each other, talk to me.**

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

09:47:58 **THE WITNESS: No offense taken.**

09:48:03 **BY MR. HADDEN:**

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

09:48:12 A. **Correct.**

09:48:16 Q. **And that included this notion of the parties in a conversation using shared knowledge, right?**

09:48:17 A. **Correct again.**

09:48:18

09:48:25

09:48:28

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**

that, correct?

A. We were talking about implementing a lot of these ideas in a machine, and that as such, that was less important.

MR. HADDEN: I have no further questions. Thank you for your time.

THE WITNESS: Thank you.

THE COURT: All right. Thank you.

Redirect.

MR. SMITH: No redirect. Thank you.

THE COURT: All right. Thank you, sir. You are excused.

THE WITNESS: Thank you.

THE COURT: What's next?

MR. YOON: Your Honor, at this time, we are going to call by deposition, Kelly Vanee, who was the former Vice-President of Learning and Engineering at Amazon.

(Videotape deposition of Kelly Vanee:)

Q. -- PowerPoint presentation with a Bates number ending in 44736?

A. Oh, there we go. Yep, it's open now.

Q. Well, do you recognize Exhibit 226 to be a slide deck that you presented at Amazon?

A. Yes, this is a slide deck that I helped author and present.

Q. And do you recall when you presented this slide deck?
A. This was presented numerous times throughout the -- the history of Alexa. And it was -- the presentation was recorded and watched by a large number of individuals, as well.

(End of videotape.)

MR. YOON: Your Honor, at this time, we move into evidence PTX-226.

MR. HADDEN: No objection.

THE COURT: All right. It's admitted.

(PTX Exhibit No. 226 was admitted into evidence.)

MR. YOON: And Your Honor, at this time, VoiceBox will call its expert, Nathaniel Polish.

THE COURT: Okay.

COURTROOM DEPUTY: Please raise your right hand. Please state and spell your full name for the record.

THE WITNESS: Nathaniel Polish, N-A-T-H-A-N-I-E-L, P-O-L-I-S-H.

NATHANIEL POLISH, having been duly sworn, was examined and testified as follows:

DIRECT EXAMINATION

BY MR. YOON:

Q. Good morning, Dr. Polish.

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 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. Dr. Polish, could you briefly describe your educational background?**

09:56:36 A. **Sure. I have a bachelor's in physics from Columbia, and I have a Ph.D. in computer science from Columbia. My Ph.D. work was in computer speech.**

09:56:47 Q. **And Dr. Polish, could you briefly describe your professional experience?**

09:56:51 A. **Sure. Outside of -- as soon as I got done with high school, I started with a consulting firm called Daedalus Technology Group, it's a research and development outsource for other companies, companies coming to me and ask me to help them create new technologies, and we and a team of engineers produce prototypes and early stage concepts for the clients.**

09:57:16 Q. **And Dr. Polish, could you briefly describe your professional experience?**

09:57:20 A. **Yeah. So in addition to Daedalus, I from time to time start companies that commercialize technologies that I invent. I have probably done that seven or eight times over the last forty years. Two of the companies relevant to 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**

that goes on for some particular purpose, in some cases, involving eCommerce.

Q. And if we could have PDX slide four. Dr. Polish, what have you been asked to do in this case?

A. So I was asked to look at the patents, look at the accused products, and determine whether I believe that there was infringement of those -- of those patents by those products.

Q. And what did you do to prepare yourself to offer such an opinion?

A. I looked at a lot of documents. I reviewed the patents. I looked at the prosecution histories. Prosecution histories are the interaction between the patentee and the Patent Office as part of getting the patent. I looked at the Court's claim construction order. I looked at PTAB rulings, I used Alexa extensively. I reviewed many technical documents provided to me in the case relating to Alexa, and then I confirmed my analysis with source code. And I ultimately compared the claims of the patents to the Alexa system.

Q. And if I could have PDX slide five.

Dr. Polish, as part of your analysis, did you review the Court's claim construction of certain claim terms with respect to the four patents?

A. I did.

Q. **And for example, do you see there is a reference there to speech recognition engine/speech recognition?**

A. **Yes.**

Q. **And for illustration, do you see that it's software or hardware that recognizes the words or phrases in the natural language utterance?**

A. **Yes.**

Q. **And in your analysis, did you apply the Court's claim construction to your infringement analysis?**

A. **I did.**

Q. **Now, Dr. Polish, if a claim term was not expressly interpreted by the Court, did you apply the ordinary meaning of claim terms to a person of ordinary skill in the field?**

A. **Yes, I did.**

Q. **If we could have slide six on the screen.**

And Dr. Polish, in your opinion, what's the level of ordinary skill in the art with respect to the four patents?

A. **So, a person of ordinary skill in my view, is somebody who has a bachelor's degree in computer science or related field, plus two years of practical experience in various kinds of speech systems. You could trade some of that off with more education or more experience, but this is approximately the level of skill that you would need to understand the patents.**

Q. **And is it your understanding that the patents were written for people of such skill?**

A. **Yes.**

Q. **And that would be the background of someone reading the patent?**

A. **Yes, that's right.**

Q. **And if we could now have on the screen JTX-1, the '681 patent.**

And Dr. Polish, what conclusions have you reached regarding Claim 13 of the '681 patent?

A. **My conclusion is that the Alexa products infringe Claim 13 of the '681 patent.**

Q. **And if we could now have on the screen the '176 and '097 patents.**

And Dr. Polish, with respect to the asserted claims to each of the patents in this case, what are the opinions that you have reached with respect to whether or not Alexa infringes those patents?

A. **So the asserted claims of each of those patents are infringed by the Alexa systems.**

Q. **In particular Alexa Shopping?**

A. **Yes.**

Q. **And let's go specifically, if we could have just the '176 patent on the screen. And what conclusions have you 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 reached with regard to Claim 23 of the '097 patent?**

10:03:31 A. **That Claim 23 is infringed by the Alexa Shopping systems.**

10:03:36 Q. **And if we could have the '703 patent on the screen. And sir, what conclusions have you reached regarding Claim 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 by element, your infringement analysis for those patents shortly, but let's talk a little bit about the background for each patent.**

10:04:00 **If we could have on the screen the '681 patent. If we could turn to column one, lines 7 to 8 of the '681 patent.**

10:04:16 **And -- I'm sorry, column one lines 5 to 9, I apologize.**

10:04:22 **What was the field of invention for the '681 patent?**

10:04:25 A. **Well the invention relates to a cooperative conversation model for human to machine voice.**

10:04:34 Q. **And the '681 patent you understand relates to a**

machine voice user interface, correct?

A. That's right.

Q. And what is a cooperative conversation model, sir?

A. A cooperative conversation model is some kind of model that represents an expected dialogue between a user and a computer. So it's something that represents what interaction would be expected between the human and the computer for some particular purpose.

Q. And if we could now have Figure 1 of the '681 patent. And column six lines 64 and 65. Dr. Polish, can you describe the system disclosed in the '681 patent?

A. Sure, so this block diagram, a very high level diagram of how this all works. In the upper left is what's labeled as input, that's the human speech that comes in. The next box is labeled ASR, Automatic Speech Recognition that converts it, converts the speech into some form of text. Then there is a Conversational Language Processor which is, also can be an NLU or Natural Language Understanding processor that seeks to understand what the text is talking about. And that, in combination with applications and databases, tries to understand what the user's intentions or goals are, and ultimately constructs an output which goes out on the left as generated speech.

Q. If we could now have Figure 3 of the '681 patent, and 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.**

Q. **And what does it say?**

A. **So short-term shared knowledge is accumulated during a single conversation. So if you're engaged in a conversation with a computer that's a multiple interaction with the computer, from interaction to interaction, that's short-term knowledge.**

Q. **And if we look at lines 63 to 65, do you see there is a reference there to build long-term shared knowledge?**

A. **Yes.**

Q. **How does the '681 describe how long-term shared knowledge can be built?**

A. **So long-term shared knowledge, is knowledge about the user that extends over multiple conversations, so it would be things about the user that presumably don't change, it might involve where their device is, it might involve the acoustics of their room, it might involve their profile in some other way that would help with the conversation.**

Q. **And if we could now look at Figure 3 and have column 13, lines 31 to 36 on the screen.**

And does the '681 patent describe how long-term and short-term shared knowledge can be used?

A. **Yes, so it talks about that for long-term and short-term shared knowledge, it accumulates over a single conversation to help build user profiles, other information about the user.**

Q. **And what is the benefit of using long-term and short-term shared knowledge?**

A. **There is a lot of benefits to it. It generally streamlines the conversations. It means that the user doesn't have to keep repeating themselves because you have already set a context of what you're looking for, so you don't have to keep saying the context over and over again. It streamlines the conversation, it makes errors less likely because you're interpreting things from the point of view of what's come before. With long-term knowledge, you might have information about whether someone is an expert or a novice, or whether they're a native English speaker or something of that nature.**

Q. **Would the '681 patent reduce the number of turns or interactions a user would have to have in order to get something done?**

A. **Yes, it would streamline a number of things including the number of times you would have to go around.**

Q. **And if we could now discuss the '176 and '097 patent together. And for convenience sake, Dr. Polish, you understand that the '176 patent and '097 patent share the same specification?**

A. **Yes, I do.**

Q. **And so I'll just for convenience, direct you to where it is in the '176 patent, but I'll be asking you about both.**

10:11:03 A. **Okay.**

10:11:04 Q. **If I could go to column one, lines 8 to 10 of the**
10:11:08 **'176 patent. What is the field of invention for the '176**
10:11:13 **and '097 patents?**

10:11:15 A. **So the field of the invention of these patents is**
10:11:18 **selecting and presenting advertisements based on natural**
10:11:22 **language processing of voice-based input.**

10:11:26 Q. **And if we could have Figure 1, column two, lines 50**
10:11:30 **to 53 on the screen. And looking at Figure 1 of the '176**
10:11:37 **and '097 patent, can you describe the system disclosed in**
10:11:41 **those patents?**

10:11:43 A. **Sure, similar to what we saw before, the input, which**
10:11:47 **is human speech, is on the upper left connecting to an ASR,**
10:11:55 **which connects to a Conversational Language Processor or**
10:12:02 **NLU. This then connects to applications including an**
10:12:06 **advertising application and connects to databases. From**
10:12:12 **that, output is generated and sent to the user in the form**
10:12:17 **of vocalized speech.**

10:12:23 Q. **If we could also have Figure 1, column 4, lines 17 to**
10:12:29 **34, if you could move it down a little it. Column 17 is cut**
10:13:11 **in half.**

10:13:12 **Dr. Polish, does the '176 and '097 patent**
10:13:16 **explain how the Conversational Language Processor operates?**

10:13:20 A. **Yes, it describes it here, it talks about long-term**
10:13:23 **and short-term knowledge. And using that to understand the**

current utterance.

Q. And looking at line 25, do you see the statement "for example, Conversational Language Processor 120 may use context determination model 130 to establish context for a current utterance by having domain agents." Do you see that?

A. Yes.

Q. What is a domain agent?

A. A domain agent is some piece of code that is associated with a particular topic area. So the domain might be shopping, it might be weather. It might be travel. But it's some topic area where the interactions with the users make sense to group them together within a single module.

Q. Do you see there is a reference there to scoring possible interpretations and selecting a highest scoring interpretation?

A. Yes.

Q. What do you understand that to be part of?

A. So these systems are not -- they're not ever completely sure about themselves, I mean, sometimes they may be, but usually they operate with competing hypotheses. So someone is talking to it, and it thinks they're probably talking about the weather, but maybe they're talking about travel. So it will often maintain a collection of competing

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

7 to 12 on the screen. Does the '176 patent provide examples of the types of ads that may be used, or how ads may be selected?

A. Yes, it talks about ways in which ads could be selected based upon any number of things, including user profiles, demographics, so that things would be served to a user that would be appropriate to their needs or might be more likely to be purchased than some other offering.

Q. And if we could have column two, lines 21 to 25 on the screen.

Does the '176 and '097 patents provide examples of the types of advertisements that can be used by the system?

A. Yes. They talk, for example, about the advertisements could include sponsored messages, purchase opportunities, other marketing communications. It would be -- it could be a pretty broad range of things, but importantly, it could be just any kind of message that is offering a purchase.

Q. If we can now go to Figure 3, column two, lines 55 to 57.

Dr. Polish, what's shown in Figure 3 of the '176 and '097 patents?

A. So this is a flow chart that is showing roughly what happens when voice input is received where voice input comes

in at the top, it's identified and then -- and then an ad is chosen, and it's then presented to the user in the form of speech. And then -- and then the user can interact with that, with that advertisement, presumably to purchase something or ask a follow-up question or something like that. And then action is taken based on that. And it can loop where it continues to present output and get interaction from the user, and ultimately the behavior of the user is being tracked.

Q. And given the interface between the user and the system, do the '176 and '097 patents streamline the communications between the user and the system such as the number of turns or interactions?

A. Yes, as you go through these loops, things become more focused as you're interacting with the system in response to this ad. And more information is being obtained. So this approach to sort of focusing the interaction reduces the amount of cycles through the system you would have to do.

Q. If we could go to JTX-2, the '703 patent. And if we could now have column one, lines 11 to 17 on the screen.

And what is the field of invention for the '703 patent?

A. So the '703 patent is related to voice commerce, and particularly to systems and methods around streamlining

checkout. So it makes the checkout process more efficient.

Q. And if we could now have Figure 1, column three, lines 20 to 21 of the '703 patent on the screen.

And what's shown there, sir?

A. All right. So this is a diagram that's showing various components of the system. It's showing the voice commerce application that has various components within it that are managing profiles and managing the user interface, that are then talking -- are working with storage devices and processors, and have databases associated with them. And then this whole system communicates to various third-party external systems which would include search engines and retailers and service providers.

Q. And if we could now have Figure 2 and column three, lines 20 to 23 on the screen.

And what's shown in Figure 2, Dr. Polish?

A. So this is showing block diagrams for different components of the systems and how they communicate, it's actually showing something fairly similar to things we've seen before. On the upper left are input devices such as an Alexa, which would go into a speech recognition engine. That would feed a Natural Language Processing system or NLU and also talk to a database, and also talk to applications. And ultimately the result of all that goes to output devices where the system speaks.

10:22:57 Q. **And if we could now have Figure 3, column 3, line 24 to 26.**

10:23:03 **Can you explain what's shown in Figure 3 of the '703 patent?**

10:23:06 A. **This is a flow chart that is showing the operation of the system from where it receives a natural language utterance, and then the steps it takes to automatically without further user input, complete the purchase and ultimately complete the transaction for the user.**

10:23:13 Q. **And if you look in Step 304 and 310, there is reference to without further user input. Do you see that?**

10:23:38 A. **Yes.**

10:23:43 Q. **What does that indicate?**

10:23:46 A. **Well that means that it's not waiting for some sort of step from the user, that when it gets to that point, it just effectively falls through and gathers the information it needs from other sources besides the user. So it's an ultimate streamlining of a purchase.**

10:23:47 Q. **Does that reduce the number of turns or interactions the user has with the system?**

10:23:48 A. **Yes, if it can operate without further user input, then there is no further turns at all, and it just completes the transaction.**

10:23:52 Q. **If we could now have Figure 3, column eight, lines 19 to 24 on the screen.**

10:23:58

10:24:01

10:24:06

10:24:08

10:24:11

10:24:13

10:24:17

10:24:20

10:24:25

10:24:29

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 **analysis to determine whether or not Alexa infringed the**
10:26:28 **VoiceBox patents, what did you review and base your opinion**
10:26:31 **on?**
10:26:31 A. **I reviewed technical documents provided to me in the**
10:26:35 **case. I looked at the Court's construction order. I used**
10:26:41 **Alexa. And I confirmed what I saw by looking at source**
10:26:48 **code. And then I looked at the patents and compared it to**
10:26:51 **what I was seeing in Alexa.**
10:26:54 Q. **If you would open your binder, sir, and do you**
10:26:58 **recognize exhibit PTX-68?**
10:27:05 A. **Yes.**
10:27:06 Q. **And is that the document entitled ASR overview and**
10:27:11 **workflows?**
10:27:11 A. **That's right.**
10:27:12 Q. **Is that one of the technical documents from Amazon**
10:27:15 **that you based your opinion on?**
10:27:16 A. **Yes, it is.**
10:27:17 **MR. YOON: I would like to move into evidence**
10:27:19 **now, Your Honor, exhibit PTX-68.**
10:27:22 **MR. HADDEN: No objection.**
10:27:24 **THE COURT: Thank you. It is admitted.**
10:27:26 **(PTX Exhibit No. 68 was admitted into evidence.)**
10:27:26 **BY MR. YOON:**
10:27:27 Q. **And so do you see here sir on the bottom right corner**
10:27:32 **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 10:27:48 10:27:49 A. **Yes, my understanding.**

10:27:51 Q. **And do you see that in this document here it's entitled ASR overview and work flow?**

10:27:56 10:27:59 A. **Yes.**

10:28:00 Q. **That's for Alexa, correct?**

10:28:02 A. **That's correct.**

10:28:02 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?**

10:28:07 10:28:12 10:28:15 A. **Yes, that's right.**

10:28:17 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?**

10:28:19 10:28:25 10:28:34 A. **Yes, it does.**

10:28:37 Q. **And where is that indicated?**

10:28:38 A. **So the box on the upper right of this diagram is labeled ASR, that's where the speech recognition takes place.**

10:28:39 10:28:43 10:28:48 Q. **Does the Amazon Alexa include an NLU engine?**

10:28:48 A. **Yes.**

10:28:53

Q. **Where is that?**

A. **That's just in the box below, it's labeled NLU on it.**

Q. **Now, if we can now expand that document back.**

And if you could now expand the bottom there that says flow description. -- oh, actually, would go back Mr. Smith, I apologize. If we can just expand out that top part.

Dr. Polish, what's an ASR?

A. **It's a speech recognizer, it's something that takes speech and does some kind of recognition on it, turning it into text or some other kind of more compact representation.**

Q. **And what is an NLU, sir?**

A. **An NLU is a Natural Language Understanding system, which is some kind of -- can be referred to as Artificial Intelligence, it tries to make sense of speech in the context of a particular purpose. So it puts the speech into some kind of context, you think of it as understanding or assigning meaning.**

Q. **If we could have a split screen with this figure and the flow description below. If we can just focus on the flow description through steps 1 to 5 for a moment.**

A. **Okay.**

Q. **Mr. Smith, can you expand it more? And I'm only doing this, Dr. Polish, to try to make it a little easier to read.**

A. **Yes.**

Q. **Now, does this flow description and document describe how Alexa operates?**

A. **Yes, in part it does.**

Q. **Could you walk through that for us?**

A. **Sure, it starts off at Item 1 where the examiner asks Echo or Alexa what's the weather today. So you then have, it notices the wake word, which is Alexa, and it streams the rest of it, what is the weather today to the ASR. So that goes there. The ASR takes it and does further processing on it, and it comes up with what it's referring to as the NBest utterances, which is what its NBest guesses are as to what the audio stream contains.**

Q. **Is that Step 2?**

A. **Yes.**

Q. **And what happens next?**

A. **So the NBest gets sent to a, what is referred to as a speechlet which is some application code that is oriented to a particular task such as shopping, or getting the weather. And ultimately, the NLU detects what is referred to as the domain which is "Weather", in this case, and the intent is "get the weather", and it says here no slot is detected. And that means that the intent "GetWeather" would have slots associated with it which are pieces of information that would be needed to fulfill the request such as location and**

time.

Q. If you look at 4(a), it says it goes in parallel through both FST's, rule-based and exact match and statistical models paths and then goes to a range to choose the best output, do you see that?

A. Yes.

Q. We discussed in this case about rules-based and AI and machine learning, do you recall that?

A. Yes.

Q. Does the NLU of Amazon Alexa use rules to determine?

A. Yes, my understanding is that there are rules-based components to the NLU among other things that it does. So it's definitely using rules and that's indicated here.

Q. If we can now go to and re-expand it, and focus on steps 6 through 9. What happens after the NLU detects the domain which is "Weather," the intent, which is "GetWeather," and no slot is detected?

A. So what it is talking about here is that it's noticing that it doesn't have all the information that it needs, so it's saying, if you look at 6(b), where it says that the "Weather" speechlet detects a missing slot, this case location, and you can have a pre-configured rule that says that it should now perform an action which is ask for location in that case, you can have a different rule which says use the location where the device is located, or some

other rule, but in this case, the rule is set to if you don't have a location, ask for it. So it then goes off and prompts the user for the location.

Q. And would the slot eventually be filled?

A. Yes, if the user cooperates and provides the location, the slot would eventually be filled and that would be satisfied and they could fulfill the request.

Q. If we could now turn in your binder to PTX-78.

A. Okay.

Q. And do you recognize exhibit PTX-78?

A. Yes, I do.

Q. And what is the title?

A. The title is history of Alexa Shopping and Alexa, and it's dated December 17th, 2020.

Q. And this is -- was one Amazon technical documents you reviewed to base your opinion on?

A. Yes.

MR. YOON: I would now like to move into exhibit PTX Exhibit 78.

MR. HADDEN: No objection.

THE COURT: Thank you. It is admitted.

(PTX Exhibit No. 78 was admitted into evidence.)

BY MR. YOON:

Q. If we could now turn to page Amazon 8041. Generally, 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, it's showing the connection of the device, in this case the Echo on the far left, passing information to this operating system called DeeSOS, which communicates with the ASR and communicates with the NLU and communicates with speechlets, 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, 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 intents and slots, each of them having a domain associated with them. The output of the NLU is a collection of the hypotheses of what the user's intent is, and how different 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 labeled intent.**

10:37:30 Q. **Now, if we expand the page back. Let's go to Section 2.1, the glossary, how Amazon defines the terms. 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 that you based your opinion on?**

10:39:23 A. **Yes, it is.**

10:39:24 **MR. YOON: I would like to now move into evidence, PTX-84.**

10:39:25 **MR. HADDEN: No objection.**

10:39:27 **THE COURT: Thank you. It is admitted.**

10:39:29 **(PTX Exhibit No. 84 was admitted into evidence.)**

10:39:30 **BY MR. YOON:**

10:39:32 Q. **Do you see on this document here, it has the reference to dialogue platform?**

10:39:34 A. **Yes.**

10:39:36 Q. **And do you see there is a primary owner of this Amazon technical document is Amazon Shopping?**

10:39:38 A. **Yes, I do.**

10:39:47 Q. **And do you see that it provides -- if we can go back. If you can go to the section there, it provides an overview, do you see that, Section 1?**

10:39:48 A. **Yes.**

10:39:55 Q. **And that includes an overview on conversation run time, what's new to Alexa Shopping, conversation management, do you see that?**

10:39:59 A. **Yes.**

10:40:00
10:40:00
10:40:04
10:40:09
10:40:09

Q. **If we can now turn to page 5509. And if we could expand conversation runtime.**

Dr. Polish, what are the three primary components of the Alexa Shopping dialogue platform?

A. **There is the Shopping FrontDoor, the Shopping Orientation, and Shopping Conversation State Management.**

Q. **If we could expand shopping conversation state management, let's start with the word "state." What is state?**

A. **State generally refers to the state of a system or all the different parts of the system at the current moment. So in a person it might be, you know, weight, height, heart rate, things like that, things that might change, but at a given moment, it has a particular state.**

Q. **Does Alexa use states?**

A. **Yes, it does, it uses states as part of taking somebody through a conversation to make something happen.**

Q. **Now, with regards to the Alexa Shopping platform, what does shopping conversation state management do?**

A. **Well, what it does, what it is explaining here is that it's keeping track of what's been done across time so that it can help -- it can help a user make progress through their purchase, or make progress through their transaction and keep track of what they've done and what they need to do.**

10:42:15 Q. **Do you see the statement "we enable a smarter conversation with customers by paying attention to what they have said and done to make better decisions later." Do you 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 Honor, was already admitted during the deposition of 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 Natural Language Understanding Deep Dive?**

10:42:54 A. **Yes.**

10:42:54 Q. **And you were here, you understand that one of the authors of this document is an Amazon Alexa engineer, Kelly Vanee?**

10:43:04 A. **Yes.**

10:43:05 Q. **And was this one of the documents you based your 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 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 also has the -- it's interpreted tomorrow to mean that date.
10:44:59 Q. And if could go to the next page? And does this
10:45:07 slide indicate the stages of the Alexa NLU engine?
10:45:12 A. Yes, it gives you a little more detail, it's showing
10:45:15 there is an intent classification where it comes up with the
10:45:19 notion that this is a get weather forecast intent, and then
10:45:22 it has a separate step which is slot resolution where it's
10:45:26 trying to figure out how to fill in the date slot which it's
10:45:31 doing here with the word tomorrow.
10:45:33 Q. Do you see at the bottom the reference there to BIF?
10:45:36 A. Yes.
10:45:37 Q. That's the output of the NLU right?
10:45:39 A. Yes, that's the same BIF we saw before.
10:45:41 Q. So the NLU engine of Alexa would go through a
10:45:45 recognition step, and then a resolution step before it
10:45:50 outputted the BIF?
10:45:51 A. That's right.
10:45:52 Q. What is recognition?
10:45:54 A. I'm sorry, say that again.
10:45:55 Q. What does recognition refer to?
10:45:59 A. Recognition refers to that it's taking the recognized
10:46:04 speech, and it's breaking it up into pieces that it can then
10:46:09 use to further process and determine the intent.
10:46:13 Q. And then what's resolution, sir?
10:46:16 A. So resolution is where it's taking any number of

possible words that the user could have used to indicate time, and turning it into a recognizable time entry.

Q. So that, for example, when you ask about the weather, Alexa, before it gave you the weather would want to know what date you wanted the weather associated with?

A. That's correct.

Q. If we could now go to page 10?

What is illustrated on this page?

A. This is another view on this process, this is where somebody says "what is the weather in Seattle," recognition takes place, and it turned it into domain is weather, and the intent is get weather forecast.

Q. And this would be done by the NLU?

A. Yes.

Q. If we go to page 11. Now, we have been talking about the recognition in the NLU. Do you recall that, sir?

A. Yes.

Q. And what's shown on this page?

A. So this is a particular view on what's taking place in the NLU, particularly around recognition. It's showing that a number of things are happening simultaneously. So the input comes in on the left where it says input request, and it's going to, in this case, four different places. The first one it goes to is labeled exact match rules. So those 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

learning and AI, they're talking about the model there.
Q. What would be the FST step?
A. The FST step is something involving something referred to as parsing, where it's trying to, in a more deterministic rules-based approach, try to take apart the words and understand what they mean using rules rather than statistics.
Q. Let's talk about one possible operation in the Alexa system. Okay, sir?
THE COURT: At the end of this, let's take our break.
MR. YOON: Yes, let me just finish this paragraph.
BY MR. YOON:
Q. In this example here, an utterance is received at the ASR, it interprets it and detects and it provides that to the recognition engine of the NLU core pipeline. Okay, sir?
A. Yes.
Q. And if that input satisfies the rules in the system, it would reach resolution, correct?
A. Yes.
Q. And so, in the Alexa system when the NLU is interpreting intent, there are situations where it will only use rules and will not use AI or machine learning?
A. Yes, if a rule can be satisfied, it will satisfy the

10:51:03 rule preference over the other things because it's faster.
10:51:07 Q. And you have been -- I will?
10:51:09 MR. YOON: I'll stop there, Your Honor.
10:51:11 THE COURT: All right. Let's take our 15-minute
10:51:14 morning break.
10:51:15 COURTROOM DEPUTY: All rise.
10:51:17 (Jury exiting the courtroom at 10:51 a.m.)
10:51:44 THE COURT: All right. Anything we need to
10:51:46 discuss?
10:51:47 MR. YOON: No, Your Honor.
10:51:47 MR. HADDEN: No, Your Honor.
10:51:48 (A brief recess was taken.)
11:10:52 COURTROOM DEPUTY: All rise.
11:10:54 (Jury entering the courtroom at 11:10 a.m.)
11:11:14 THE COURT: All right, everyone welcome back.
11:11:20 Everyone else, you can be seated.
11:11:23 Please continue.
11:11:24 MR. YOON: Yes, Your Honor.
11:11:25 BY MR. YOON:
11:11:26 Q. Dr. Polish, do you recall earlier this morning we
11:11:29 were talking about how the NLU would score, come up with
11:11:35 hypotheses and provide scores?
11:11:37 A. Yes.
11:11:37 Q. And if we could now go to page 18 of Exhibit 226.
11:11:46 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 utterance play Hunger Games, so the user has presumably said play Hunger Games, and the system is trying to figure out what that means, so it's come up with, in this case, four different intent hypotheses. Two of them are music related, one of them is books, one of them is video. So those are potentially different domains. And it scored them with these different ranked values on the right, and the blue arrow means that it's picked the first one as the best hypothesis, so it's assuming that when the user said play 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 the same utterance play Hunger Games, is there information that the system could use that would result in the NLU selecting something different than the scoring of just Hunger Games?**

11:13:08 A. **Yes, so what this is showing at the top is in addition to saying play Hunger Games, the device that that request came from was a Fire TV, which is a video play back device, so in that case, it has decided that the best one was actually the last one, which was play a video called Hunger Games.**

11:13:33 Q. **So, in the first example, just the words play Hunger Games, it ranked play the song Hunger Games number one, with 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 TV, it altered the selection because it came from a TV, to video?**

11:13:52

11:13:56

11:13:56 A. **Right, that was additional information that was available that changed the score, that changed the results of its decision making.**

11:13:56

11:13:59

11:14:02

11:14:03 Q. **If we go to the next page. Does information stored by the user, could that affect what's selected?**

11:14:10

11:14:13 A. **Right. So in this case, it's saying that intent resolution, which is a component of the system which is described in the document found that the user had an album called Hunger Games in their cloud player. So in that case, the user had previously purchased, or somehow obtained an album called Hunger Games and stored it. So now when it goes through this recognition process, the hypothesis that you mean the album Hunger Games, that becomes the best hypothesis.**

11:14:20

11:14:23

11:14:27

11:14:30

11:14:35

11:14:39

11:14:44

11:14:50

11:14:51 Q. **Could that purchase of the album Hunger Games come from buying it over Alexa?**

11:14:56

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 determine what the intent interpretation is?**

11:15:06

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?**

A. .6.

Q. Could context also be used by the Alexa system to cause a re-ranking or readjustment?

A. Yes.

Q. And if we go to the previous page to this one. And what's shown here on this Alexa technical document?

A. This is just some text on this presentation that says based on additional signals or features from the rest of the system, the recognition results can be rescored and re-ranked. So the signals are features is a very broad statement of pretty much any kind of input could be taken to influence what the ranking is on these different hypotheses.

Q. One of those inputs could be context?

A. That's correct.

Q. If we could now go to PTX-472. And Dr. Polish, do you recognize exhibit PTX-472?

A. Yes, I do.

Q. Is that one of the documents you relied on in forming your opinion and understanding regarding the design and operation of Alexa?

A. Yes.

MR. YOON: I would like to now move into evidence exhibit PTX-472.

MR. HADDEN: No objection.

THE COURT: Thank you. It is admitted.

(PTX Exhibit No. 472 was admitted into evidence.)

BY MR. YOON:

Q. And do you see there is a reference there to speaker ID cloud. Do you see that?

A. Yes.

Q. And what is that, sir?

A. So speaker ID is where the system identifies different speakers from within a household, so there is a speaker registration process where you can say a few sentences, and it learns your voice so that particularly if you have multiple people in the same household, it can distinguish those different people within the household and do different things depending on who they are.

Q. If could go to page 2 of this document, and if you would expand speaker ID cloud, speaker identification service. And what's that, sir?

A. So this is just -- this is talking about what I just said, which is that it's using this service to distinguish individual users of the device in a single household. They would use the speaker ID service.

Q. And is this -- what type of information is this in the system?

A. So this would be long-term information, so your registration with the system when you enroll yourself in

this, would retain some piece of information that would be retained from conversation to conversation, presumably over the course of months or years, that would be used by the system to customize your experience depending upon who you are.

Q. Looking at the last sentence there, it says "the incoming audio stream is forked to the speaker ID service, and this service helps in determining who in the household is currently using the device." Do you see that?

A. Yes.

Q. What does that indicate?

A. When they say forked, what they mean is that the stream, audio stream is split, so it's going to both the speaker ID service, as well as, the speech recognizers, this is happening without causing some sort of disruption, so it's happening in parallel if you will to the regular speech recognition.

Q. Does Alexa, ASR and NLU use this speaker ID information in interpreting what is said?

A. Yes.

Q. How so?

A. Well, it knows, it learns from this -- who the user is, and loads various profiles and information about the user that is then used to direct the conversation.

Q. And if we could now expand back and do you see there

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

retained by the lawyers in the case to look at Amazon's source code, and when I had questions about things I wanted to understand about the source code, I told lawyers about it, they instructed Mr. Peck, and he went off and did the research. And I got reports back from him, including a report that he filed in this case of what he found.

Q. Now, Dr. Polish, have you helped prepare a video illustrating a conversation with Alexa?

A. Yes, I have.

Q. And if we could have PDX-3-7 on the screen.

And when you were having your actual conversation with Alexa, what features did you look for?

A. So I went through this conversation with Alexa particularly looking for Voice ID, the presence of long-term information, the presence of short-term information, I showed voice purchasing, I showed pronoun use which is important in one of the patents, and I showed a voice advertisement. And when you watch the video, you can see all these -- all these elements come out in the video.

Q. Why don't we play the first video now.

(PDX-3-7 video played.)

MR. YOON: If we can play the next video.

(Video played.)

MR. YOON: Your Honor, at this time, we would like to move into physical evidence PTX-674.

11:28:13 **MR. HADDEN: No objection, Your Honor.**

11:28:15 **THE COURT: Thank you. It is admitted.**

11:28:15 **(PTX Exhibit No. 674 was admitted into evidence.)**

11:28:21 **BY MR. YOON:**

11:28:27 Q. **Now, Dr. Polish, let's now go into the specifics of your infringement analysis. If we could have PDX3-9 on the screen.**

11:28:45 **And focusing on the '681 patent, Dr. Polish, what's your opinion with respect to Claim 13 of the '681 patent?**

11:28:55 A. **It's my opinion that Claim 13 is infringed by the Alexa system.**

11:29:00 Q. **And now let's discuss Claim 13. If I could have PDX3-10 on the screen. Do you see on that screen is the full language of Claim 13 of the '681 patent?**

11:29:14 A. **Yes.**

11:29:15 Q. **Dr. Polish, what's a nontransitory computer readable medium containing computer-executable instructions?**

11:29:24 A. **That's patent language for some piece of storage computer code, so the computer executable instructions are just computer code, and nontransitory means that it's stored somewhere.**

11:29:41 Q. **So focusing on the first part of Claim 13, what we call the preamble, if we can just expand that Claim 13**

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

11:31:22 A. **Well, technical documents that I've seen, as well as, 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 meets the receive an utterance requirement?**

11:31:53 A. **Yes. This is showing an Alexa device receiving speech and that the speech, the utterance is being sent to the Alexa voice service.**

11:32:07 Q. **And when someone speaks to Alexa, can we speak in one or more words that have different meanings and different contexts?**

11:32:15 A. **Yes, they can.**

11:32:16 Q. **If we now go back to the next element of the claim, PDX3-10. If we could have the next one on the screen, accumulate short-term. This element recites "accumulate short-term shared knowledge about the current conversation, wherein the short-term shared knowledge includes knowledge about the utterance received at the voice during the current 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 own use -- and my own use of the system.**

11:32:57 Q. **If we go to the next slide, please.**

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

response from the user, and it goes out and makes a request of the user and then tries to interpret what the user says as a response to that disambiguous request.

Q. So in the example here, when it hears the number six, it could be either an intent to set the volume, it could be a song name, or it could mean this is a put multi-turn?

A. That's correct.

Q. In the example you gave in your video, that was multi-turn?

A. That was 6:30, followed by a request for morning or evening, followed by my saying morning.

Q. And PTX-226, that's the deep dive document that Mr. Vanee was the author of what we discussed, correct?

A. That's right.

Q. Go to the next element.

Can we go to PDX3-15, please. And what is -- does that document that you have been discussing provide further support for your opinion about whether the accumulate short-term knowledge is met?

A. Yes, it does. This is talking -- this is again from that same document. This is talking about a component called the context interpreter that merges the results of the disambiguation request of what came before so as to come up with the intent.

Q. If could go to PDX3-16. And this is also an

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

of it including from that demo you saw.

Q. And if we could now go to PDX3-17. And this is another document that we've -- that you've relied on. What's shown here in PDX3-17?

A. Right, so this is a document, it's a technical document that's talking about among other things, the speaker ID system that we were talking about before where it talks about that if people in a household can register themselves or enroll themselves in this program which allows the system to differentiate people in the household. And this would be a form of long-term shared knowledge.

Q. And in order to do the speaker ID, you had a conversation with Alexa, correct?

A. Yes, it was done within Alexa, as you saw in the demo, it's part of Alexa.

Q. If we can now go to the next slide. And how did your video here illustrate that -- how did your video illustrate how Alexa accumulates long-term shared knowledge about the user?

A. Well, in this case, you can see that I'm changing an alarm. So the long-term information was the old setting of the alarm, and I am changing it to a new time. So it has -- it's accumulated long-term information that informs the conversation and makes the conversation more streamlined.

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 from the communication?**

11:39:57

11:39:58 A. **Yes, it was.**

11:40:00 Q. **If we could now go to slide PDX3-19. And how do the documents here that you based your opinion on illustrate the use of long-term shared knowledge?**

11:40:07

11:40:11

11:40:13 A. **So this is a document that talks about what they call an enrollment profile, mutable model, which as it's described here, this is something that would listen to your voice, understand some of the acoustic properties of your voice and your environment and make changes to the, to this model that it stores that would cause the ASR to do a better job over time.**

11:40:18

11:40:28

11:40:31

11:40:34

11:40:39

11:40:44

11:40:48 Q. **And now let's go to the next element. And we could have that next element on the screen identified. The next element of Claim 13 refers to identify a context associated with the utterance, wherein a conversational speech engine identifies the context associated with the utterance from the short-term shared knowledge and the long-term shared knowledge." Do you see that?**

11:40:52

11:41:07

11:41:12

11:41:16

11:41:19

11:41:24

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 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. And this element says "establish an intended meaning for an utterance within the identified context, wherein the conversational speech engine establishes the intended meaning within the identified context to disambiguate an intent that the user had in speaking the one or more words that have the different meanings in the different contexts." 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, my own use of it.**

11:42:30 Q. **If we can now go to PDX3-21. How did your opinions, how were they supported by PTX-226?**

11:42:53 A. **We talked about this before, what's going on here is where you have some kind of ambiguity that the system needs to resolve and it has, and it gets more information and then this context interpreter needs to run to combine the result 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 **meaning for the utterance within the identified context,**
11:46:27 **wherein the conversational speech engine establishes the**
11:46:30 **intended meaning within the identified context to**
11:46:34 **disambiguate an intent that the user had in speaking the one**
11:46:40 **or more words that have different meanings in the different**
11:46:43 **contexts." Do you see that?**
11:46:45 A. **Yes.**
11:46:45 Q. **Is that element met by Alexa?**
11:46:47 A. **It is.**
11:46:48 Q. **What's the basis for your opinion?**
11:46:50 A. **Technical documents and my own use of the system.**
11:46:52 Q. **And if we could now go, I want to go to PDX3-22,**
11:46:58 **we're skipping 21, go to 22. 22. PDX3-22.**
11:47:37 **Could you just advance it to 22, please. If we**
11:48:11 **could just go back up with regards to, put PDX3-10 up again.**
11:48:22 **And sir, can you explain how Alexa does this, please?**
11:48:27 A. **So this is talking about the disambiguation of words**
11:48:34 **that have one or more meanings, that have more than one**
11:48:40 **meaning. So we've seen this already with when I said 6:30**
11:48:47 **which could mean two different things, and it went off and**
11:48:52 **established which it meant. There was also, we saw in the**
11:49:00 **technical document, we saw Hunger Games being interpreted,**
11:49:05 **there were different possible meanings for that, a song, an**
11:49:09 **album, a video, and Alexa went off and figured out,**
11:49:13 **disambiguated which it meant, and it used short-term**

information to do that to establish what the meaning was in the different topics.

Q. Going back to Exhibit 226, how is this demonstrated in terms of the establishing the intended meaning?

A. Right. So this is the example of Hunger Games, and so Hunger Games can mean in this case, I guess four different things, a song, an album, a book and a video. And it was able to use, in this case long-term information which was whether somebody had purchased the album before to disambiguate it and decide that it meant the album.

Q. We can keep that language up and then turn to slide 22 of 226. Advance two pages from the right. Thank you.

And we looked at this during your testimony. How does this support the basis for your opinion that the identified context is to disambiguate an intent that the user had in speaking one or more words?

A. So, this is showing that you can re-rank, or in this case, that would mean disambiguate the intent of the user and the meaning of the words based upon the ASR, the ER, or the content, and that can be long-term information, short-term information, it could be further information from the user.

Q. And did that support your opinion, sir?

A. It does.

Q. If we could now go back to PTX3-10 and go to the last

element of Claim 13 of the '681 patent. You see this is "generate a response to the utterance where in the conversational speech engine, grammatically or syntactically adapts the response based on the intended meaning established within the identified context." Do you see that?

A. **Yes.**

Q. **Is that element met by Alexa?**

A. **It is.**

Q. **What is the basis of that opinion?**

A. **Documents that I have seen, and how I have used the system.**

Q. **We now have PDX3-24. And is that element met, the last element of Claim 13 met by Alexa and supported by the technical documents?**

A. **Yes. So this technical document is showing after the utterance has gone through the system, ultimately what comes out is a voice response that is played back through the speaker.**

Q. **And if we could go to Exhibit 226, slide 32. This is a multi-turn example for Alexa, correct?**

A. **Yes.**

Q. **And this example is very similar to what you used in your video, correct?**

A. **Yes, it is.**

11:53:20 Q. **And this is the example set forth in the Amazon technical documents that was the deep dive?**

11:53:24 A. **Yes, it is.**

11:53:27 Q. **And would Claim 13 of the '681 patent be practiced during the multi-turn example?**

11:53:34 A. **Yes, absolutely.**

11:53:36 Q. **Now, there is some discussion in this case about like a single-turn conversation. Are you familiar with that?**

11:53:39 A. **Yes, generally.**

11:53:46 Q. **And would single-turn conversations include short-term shared knowledge?**

11:53:47 A. **It certainly could. It wouldn't -- not all would necessarily, but you certainly could have a single-turn that would have short-term information in it.**

11:53:53 Q. **And where the single-turn have short-term information, would Alexa also use long-term information?**

11:53:58 A. **The long-term faster information is used with speaker ID, there is lots of long-term information that would be used all the time, so in a situation where a single-turn is using short-term, it would be using both.**

11:54:02 Q. **And in the case where a single-turn utterance included short-term shared knowledge, would Claim 13 be practiced?**

11:54:05 A. **Yes.**

11:54:09 Q. **Now, let's turn to your opinion as to the '176**

11:54:13

11:54:16

11:54:20

11:54:25

11:54:27

11:54:30

11:54:34

11:54:34

11:54:35

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

11:56:30 A. **Yes, this is showing the input device and it's showing the user's utterance being sent to the Alexa voice service.**

11:56:37 Q. **If we could now go back to PDX3-10 -- PDX3-26. Thank you. Looking at the next element, "a speech recognition engine coupled to the input device, wherein the speech recognition engine recognizes one or more words or phrases in the natural language utterance." Is that element met by 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 the technical documents show that Alexa practices this element?**

11:57:26 A. **Right. So these documents are showing that the speech comes in on the input device and is sent to the ASR which is in the upper right, the Automatic Speech Recognition system which is also shown in Figure 1 of the '176 patent as the ASR in the Figure 1.**

11:57:46 Q. **Going to the next element. This element of the claim**

recites "a conversational language processor coupled to the speech recognition engine, wherein the conversational language processor is configured to." Do you see that?

A. Yes.

Q. Is that present in Alexa?

A. Yes, it is.

Q. And what's the basis for your opinion?

A. From the technical documents that I have seen.

Q. And how does Alexa meet this?

A. The conversational language processor is the NLU and some other components, it's also including -- it's also including speechlets as well, as part of that, that we talked about, but in this diagram that's contained within the NLU.

Q. And if we can go to the next slide. And how does this support your opinion regarding whether or not the Alexa Shopping system has a conversational language processor coupled to the speech recognition engine?

A. So this is from the technical document we have seen before, the coupling to the speech recognition system is via this thing on the left labeled SIRF, S-I-R-F, that's what's coming from the speech recognition, and goes into this recognition engine within the NLU.

Q. And so there is a recognition engine within the NLU 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 advertisement in the context established for the natural 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 shop for something, that is -- that's recognized, and that information is sent to a shopping speechlet which is a piece of application code that is connected to the NLU, and it selects a device or selects a product which is then presented. So the example that I gave in my demo was that I said "Alexa, I want to buy an iPhone case," and the speechlet went off, shopping speechlet went off and figured out that it wanted to sell me this case.**

12:02:17 Q. **And stay on this slide, stay on this slide, but I notice here that this also has the next element recited in the claim, press a selected advertisement via an output 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**

causes it to rescore, or to reevaluate what -- which intent is the right one. So that can be done based upon any number of signals that are passed into it to change the scoring of the intent.

Q. If we could have Exhibit 226, page 22 on the screen. This is the NLU deep dive document that we were discussing. And does re-ranking support your opinion with regard to the last element?

A. Yes, so this re-ranking could happen in response to getting some information that indicated that you had gotten it wrong, that the top scoring intent was wrong, and then it would re-rank it and try a different intent.

Q. And one of the events that could cause that, an input would be context, correct?

A. Yes, that's one of the inputs listed here.

Q. All right. Let's now turn to the '097 patent. Looking at Claim 23 of the '097 patent, starting with the preamble, do you see that this recites a system providing natural language processing based on advertisements, the system comprising. Do you see that?

A. Yes.

Q. Is that present in Alexa Shopping?

A. Yes, it is.

Q. And what's the basis for that?

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 instruction; correct?**

12:07:09 A. **Yes.**

12:07:11 Q. **Can we go to the next element? This one recites "provide an advertisement associated with a product or service for presentation to a user." Is that element present in Alexa Shopping?**

12:07:17 A. **Yes, it is.**

12:07:20 Q. **And why is it?**

12:07:25 A. **From the technical documents and from my own use of it.**

12:07:27 Q. **And how did your own use illustrate providing an advertisement associated with a product or server for presentation to the user?**

12:07:31 A. **When I asked Alexa for an iPhone case, it came up with one to offer me. It offered it to me, it presented it, it gave me its price. That offering is an advertisement, certainly as it was discussed in the patent.**

12:07:33 Q. **And what the Alexa advertisement offered is Amazon's choice is a Lago compatible with iPhone 13 Pro case liquid silicone case, right?**

12:07:35 A. **That's right.**

12:07:41 Q. **And that was selected by Alexa Shopping, correct?**

12:07:44 A. **Yes, that was selected by Alexa Shopping, I had nothing to do with choosing it.**

12:07:45

12:07:50

12:07:54

12:08:01

12:08:03

12:08:13

12:08:18

12:08:20

12:08:21

12:08:26

12:08:30

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.**

Q. **Can we go to the next element?**

This element recites, "one or more physical processors programmed with computer program instructions which, when executed, cause the one or more physical processors to." Do you see that?

A. **Yes.**

Q. **Is that element present in Alexa Shopping?**

A. **Yes, it is.**

Q. **And why is it?**

A. **Same as we had before, these are -- this is running on servers, the servers are running in AWS which has physical processors, so there are physical processors running the computer programs embodying Alexa Shopping.**

Q. **Going to the next element. Receive a user input comprising a natural language utterance. Does Alexa receive a user input comprising a natural language utterance?**

A. **Yes.**

Q. **Can we go to the next slide. Did your demonstration illustrate how Alexa receives a user input comprising a natural language utterance?**

A. **Yeah, certainly when I said, "Alexa I want to buy an iPhone case," that was the user input was natural language.**

Q. **Can you go to the next element. This element requires "provide, without further user input after the 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 would provide without further user input after the receipt of user input the natural language utterance that's an input to the speech recognition engine?

12:13:47 A. What goes on is I speak to the Echo, the Echo receives it, and sends it straight to the ASR, there is no 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 user input after the receipt of the user input, one or more words or phrases recognized from the natural language utterance as an output of the speech recognition engine. Is 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 the ASR goes straight to the NLU without my having to do anything further.

12:14:41 Q. So the natural utterance that was said is received by 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 "determine, without further user input after the receipt of the user input, a context based at least on the one or more 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 that intent is part of the context. And that's done without any further input from the user.**
12:15:33 Q. **And do the Amazon technical documents support the fact that it's a context based on one or more words or phrases?**
12:15:42 A. **Yes, that's -- we have seen it in technical documents so far including PTX-226.**
12:15:48 Q. **And one of the things that the Alexa NLU has is a context interpreter?**
12:15:52 A. **That's right.**
12:15:54 Q. **Is that illustrated in the red box here from PTX-226?**
12:15:56 A. **Yes the context interpreter is there, it's a 2 step**
12:16:03

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

12:18:23 Q. **Next slide. Let me just ask you. Sir, in your example of the video, when you said buy it, did Alexa 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 without requiring any input with respect to either financial 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 **York, 10025?**

12:20:37 A. **Yes.**

12:20:37 Q. **That's your address?**

12:20:38 A. **That's my Zip code, yes.**

12:20:40 Q. **And that, did it indicate that the total including tax after $0.65 in discounts and credits, was $13.44?**

12:20:52 A. **Yes, that's what it said.**

12:20:53 Q. **Was that charged to your credit card without any additional input?**

12:20:57 A. **Yes.**

12:20:57 Q. **Were all the elements of Claim 25 met by Alexa Shopping?**

12:21:01 A. **Yes.**

12:21:16 Q. **Just a few more questions, Dr. Polish. In this case, did you review the expert report of Amazon expert Dr. Johnson?**

12:21:27 A. **Yes.**

12:21:28 Q. **And did Dr. Johnson offer an opinion regarding possible non-infringing alternatives to the VoiceBox patent?**

12:21:36 A. **Yes, he did.**

12:21:37 Q. **And what impact did these non-infringing alternatives have on the number of turns that a user would have to use to perform the operations without practicing the patents?**

12:21:50 A. **In Dr. Johnson's view, the number of turns would increase.**

12:21:57 Q. **And would that increase -- let me ask this first. 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 order of, you know, of 2 or 3 number of turns, it may have been one-and-a-half turns, something like that.**

12:22:25 Q. **And do the VoiceBox patents reduce the number of 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 respect to how they interact with Alexa?**

12:22:40 A. **Yes, the whole point is to reduce the amount of effort the user has to go through.**

12:22:45 Q. **Would the patents in this case reduce the friction 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 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 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 evidence.)**

12:23:24

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

didn't hear that, but if you think there is, then you're going to have to tell me so I can deal with it. All right.

COURTROOM DEPUTY: All rise.

(A luncheon recess was taken.)

COURTROOM DEPUTY: All rise.

MR. YOON: I just have one brief procedural thing. Your Honor, the next witness after this one will be Mr. Peck, and Mr. Peck is the source code witness. Your Honor, the protective order, the parties agreed we wouldn't be making copies so for Exhibit 541(a) and 541(b), Mr. Smith will be using the Elmo but the witness won't be given a copy.

THE COURT: How are you going to give it to the jury when they deliberate?

MR. YOON: We'll submit it to the jury when they deliberate, there is no objection to being submitted, but the order limited the number of copies, but it will be submitted to jury, there is no dispute there, it's just two copies.

THE COURT: You guys agree?

MR. HADDEN: Fine, Your Honor.

(Jury entering the courtroom at 1:09 p.m.)

THE COURT: All right. Everyone, welcome back. Please proceed.

CROSS-EXAMINATION

13:09:52 **BY MR. HADDEN:**

13:09:53 Q. **Good afternoon, Dr. Polish.**

13:09:55 A. **Good afternoon.**

13:09:56 Q. **Now the Alexa NLU or Natural Language Understanding, uses machine learned neural networks, correct?**

13:10:08 A. **There are some neural networks components of the NLU.**

13:10:12 Q. **And those neural networks are domain models, those are machine learned DNN's and they're used to classify the user's intent, correct?**

13:10:24 A. **That's part of what the NLU does, it also does other things as I talked about.**

13:10:29 Q. **But you agree that the domain models are machine learned DNN's used to classify the user's intent, correct?**

13:10:40 A. **As I said, there are machine learned models within the NLU that are also rules and other things used to classify intent.**

13:10:51 Q. **Well, you prepared an expert report in this case that included your opinions about how Alexa worked and your opinions on infringement, correct?**

13:11:01 A. **Yes.**

13:11:02 Q. **Okay. And you have that in your binder in front of you. Let's look at your expert report. Mr. Patterson, can we look at Dr. Polish's opening report at paragraph 90. On page 31?**

13:11:24 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,**

those intents were generated by these deep neural networks that are run in the 1P pipeline; right? Do you not know how the intents were generated that you talked about all morning?

A. I know that the intents were generated within the NLU. I don't know -- what I'm not clear about is whether they were part of the 1P pipeline in particular.

Q. Okay. So you don't know other than that they came out of the NLU box, how they were actually generated?

A. Well, we talked about they were generated based upon the document we had from Mr. Vanee, that they were generated by executing a variety of different pipelines, including ones that contained rules including ones that contained finite data machines including one that contained models.

Q. You don't know whether these are generated by the 1P models that you talk about in your report?

A. I don't recall whether the -- whether what was in Mr. Vanee's diagram was entirely part of 1P or not entirely part of 1P.

Q. So in your report here, you claim that the information that's in paragraph 90, the information in a SIRF, is passed to all the domain models, roughly 37 to produce intents and token labels, do you see that?

A. Yes.

Q. And those 37 domain models, those are 37 of those

13:18:34 **machine learned deep neural networks, right?**

13:18:36 A. **Yes, that's my understanding.**

13:18:37 Q. **And each of those domain models produces multiple**

13:18:42 **results, including intent classification and name entity**

13:18:48 **recognition; right?**

13:18:50 A. **Yes.**

13:18:51 Q. **And then the Alexa NLU -- before I get there, yeah,**

13:19:02 **and the Alexa NLU sorts those results based on confidence**

13:19:10 **scores and retains the NBest to be passed along to**

13:19:18 **downstream processing.**

13:19:20 A. **Yes.**

13:19:22 Q. **Yes, and the output of the NLU retains multiple**

13:19:26 **interpretations of a customer's request right?**

13:19:29 A. **Yes, the different intents are different**

13:19:31 **interpretations.**

13:19:37 Q. **At least what you described in your report is what I**

13:19:41 **showed in the opening, right, so we have this NLU and it**

13:19:46 **receives the words from ASR, that's what's in that SIRF,**

13:19:53 **right?**

13:19:54 A. **The SIRF contains the recognized words.**

13:19:57 Q. **Right, that's the output of ASR, right?**

13:20:00 A. **The SIRF is the output of the ASR.**

13:20:03 Q. **And those words, that's what is provided to these**

13:20:07 **deep neural network domain models, and here I only show six,**

13:20:13 **but as you explained in your report, there is actually 37,**

13:20:17 **right?**

13:20:17 A. **Yes.**

13:20:18 Q. **And those guys receive whatever the words are, right,**

13:20:23 **and they process those neural networks, and each one comes**

13:20:29 **up with its own meaning, right?**

13:20:34 A. **So you used -- so first of all, the implication here**

13:20:38 **is that the entirety of the NLU is of these -- are these**

13:20:43 **DNN's, which it's not, and also you're calling them**

13:20:48 **meanings. They're different intents and associated slots.**

13:20:52 Q. **So let's just follow along with what you said in your**

13:20:55 **expert report. What you describe in your expert report is**

13:20:58 **exactly what I showed on the screen, which is the output of**

13:21:01 **ASR goes through each of the deep neural networks and they**

13:21:06 **all provide at least one, maybe multiple interpretations,**

13:21:10 **correct?**

13:21:12 A. **They each can provide one or more intents and slots.**

13:21:19 Q. **Well you called them interpretations in the expert**

13:21:22 **report, are they interpretations?**

13:21:24 A. **Yes, that's in the form of those intents.**

13:21:26 Q. **Right, so these are interpretations of what the words**

13:21:31 **meant that the user spoke; right?**

13:21:38 A. **I'm cagey about using the word meaning here because I**

13:21:42 **think it implies --**

13:21:43 Q. **Meaning today in the claim, and you talked about**

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

13:28:23 A. **Yes.**

13:28:25 Q. **Can we play the clip?**

13:28:26 **(Videotape deposition:)**

13:28:31 Q. **Is it correct that you did not review any of the Alexa source code that Amazon made available in this case?**

13:28:38 A. **That is correct.**

13:28:39 **MR. YOON: Objection, Your Honor.**

13:28:41 **THE COURT: Go ahead.**

13:28:42 **MR. YOON: I understood, Your Honor, under the procedure they're not supposed to be able to look at it.**

13:28:48 **THE COURT: We don't play it up there, that's not even impeachment. Why don't you ask him if he read it and looked at it, tell him what it says, but don't play it.**

13:28:59 **MR. HADDEN: Okay, Your Honor.**

13:29:00 **THE COURT: You did not ask him any questions I just looked at it --**

13:29:03 **BY MR. HADDEN:**

13:29:05 Q. **Do you see your testimony from your deposition on page 7, line 1?**

13:29:08 A. **Yes.**

13:29:09 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.

Q. **And you didn't tell or ask Mr. Peck to analyze certain features of the Alexa source code, is that right?**

A. **Yes, I did not tell Mr. Peck directly to do anything. I saw the questions that the lawyers were asking him, I contributed to that, and I saw what was going on through that interaction.**

Q. **And you could have reviewed the Alexa source code, it was all made available to you, correct?**

A. **Yes, I think so.**

Q. **But you didn't bother to look at any of that code that was made available to you?**

A. **I resist the term bother there, sir. I had somebody who was looking at the code very diligently and I interacted through the lawyers with him and had his report. I certainly took seriously my responsibility to confirm what I found with the code.**

Q. **But you had the opportunity to see exactly how what we saw on the screen work in the source code, and you chose not to look at it, correct?**

A. **I did not look at the source code. I did not review the source code on a source code computer.**

Q. **Can we put the demonstrative back up.**

Now, you agree that the output of NLU is a list of possible interpretations, right?

A. **The output-- my understanding is the output of these**

models is an NBest list of intents and slots.

Q. Right, that's what I have here is this list and meaning, right?

A. Okay.

Q. Now, if Alexa chose a context for all these DNN's processed the user's words, then my example from the opening of "do I need an umbrella," versus "I need an umbrella," wouldn't work, right?

A. I'm not sure I'm following.

Q. Sure. So if Alexa decided before figuring out the meaning of the user's words in all possible contexts, they decided beforehand that it was going to be focused on the shopping context, and we only ran the neural network for shopping, then if I said "Alexa, do I need an umbrella," I would get intent about buying an umbrella, and not a weather forecast, right?

A. No, I don't think that's -- I don't think that's right.

Q. Well, the weather forecast intent would come from the operation of the weather model, when it operates on the words, right, isn't that how it works?

A. So maybe we can re-examine this hypothetical. Give me the example that you're talking about.

Q. Sure. I say "Alexa, do I need an umbrella?"

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 claims, but the claims require identifying a context before you determine anything, right?**

13:35:30 A. **The claim talks about finding a context. And as I've talked about, that context can come from a variety of sources.**

13:35:42 Q. **Let's look at the claim to make it clear. Mr. Patterson, could we get PDX3-10? This is Claim 13 of the '681 patent from the counsel's presentation this morning. And before we get to the context, if we just blow up the preamble again. I think as you testified this morning, this is a claim to computer executable 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 computer code that's required by this claim was in Alexa, the best place to look for that would be in the Alexa source code, right?**

13:36:55 A. **I'm not sure I understand. My understanding is that there is servers that are running Alexa, and that those servers are embodied in executable code that's sitting on a**

server.

Q. Sure, but if we're going to understand whether or not Alexa has the specific computer instructions that are required by this claim, the place we should look is in the computer instructions in Alexa's source code; right?

A. The computer source code and the computer executable instructions are not the same thing.

Q. The computer source code is where it gets compiled to make the executable instructions the computer to run, right?

A. There are many steps between the source code and computer executable instructions. The point is, that what Alexa is doing, the Alexa server is running codes sitting on computers and that's what is -- what's being talked about here.

Q. Right, and if you're going to look at the code in a form that a human can read it, it would be in the source code, right?

A. Well, code in human readable form is source code. But this preamble is not talking about source code.

Q. You're telling me that the computer executable instructions in Alexa are not created from source code?

A. Well, of course they're created, they're created from source code. What this preamble is talking about is computer readable medium, not human, computer readable and computer executable instructions that are -- that are

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

same utterance, right?

A. Well, the short-term shared knowledge includes knowledge about the -- it includes other things as well.

Q. Sure, but the utterance in this element has to be the same utterance that we receive when we receive an utterance, right?

A. Well, it is accumulating short-term knowledge from the current utterance. The short-term knowledge includes knowledge accumulated from prior, that's what it means to be accumulating.

Q. Right, this is not that complicated. All I'm trying to confirm is when it says knowledge about the utterance, that has to be the same utterance that's referred to above where it says receive an utterance, right?

A. Again, it's not that complicated, short-term knowledge about the current conversation and it's accumulating some knowledge about the current utterance.

Q. Right, and then we go down to the identify context step. And we have to identify a context associated with the utterance again, right?

A. Yes.

Q. And that has to be the utterance that we received above; right?

A. Yes, it's the utterance.

Q. And if we go to the next element, we have to

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

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

13:45:05 you have to have a specific utterance and show that each of
13:45:09 these highlighted steps is performed with respect to that
13:45:13 utterance; isn't that what is required, isn't that the law?
13:45:17 A. I'm not a lawyer. I know that I showed that Alexa
13:45:22 was doing each of these things, and I showed documents
13:45:25 around it and I showed an example of Alexa in use.
13:45:30 Q. Let's be very clear. True fact or not, did you show
13:45:36 the specific utterance that is processed in the way required
13:45:39 by Claim 13?
13:45:52 A. It's possible that my demo example would fit this and
13:45:58 could be analyzed that way. That isn't how I presented it.
13:46:03 It's possible that my demo example shows receiving an
13:46:08 utterance and using -- and accumulating short-term
13:46:11 knowledge, identifying the context and establishing intended
13:46:14 meaning and generating a response, I didn't analyze it in
13:46:20 that way.
13:46:21 Q. But you didn't show any utterance that Alexa has ever
13:46:27 processed, that's processed as required by the steps of this
13:46:31 claim, right?
13:46:32 A. I showed a demonstration and I analyzed documents.
13:46:37 Q. You're not answering my question. Showing a
13:46:41 demonstration doesn't prove anything to the jury. What the
13:46:44 jury has to see is that you proved that each step of this
13:46:48 claim is performed on a specific utterance so that we can
13:46:51 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 example of what I showed on my demonstrative of that table
13:52:07 with the different meanings; right?
13:52:10 A. I'm not sure I know what you're referring to.
13:52:12 Q. Well, this is a portion -- right. Here we go. This,
13:52:16 what I showed here, what I have changed 6:45, meaning two,
13:52:21 meaning three, that is equivalent, if could go back, to what
13:52:25 you showed at the top of your slide where it says get set
13:52:33 volume intent, that would be one of the outputs from one of
13:52:36 the deep neural networks and it would include six, the
13:52:42 meaning, the meaning of six is volume level, right?
13:52:46 A. I don't know that this is the output of an DNN. This
13:52:51 is an output -- this is output from the NLU which as I said
13:52:55 does things besides the DNN, this could have been the result
13:53:01 of an DNN, but I don't know how it produced this.
13:53:05 Q. And the same is true for the music intent song name
13:53:11 equals six, so six went into the music DNN and it figured
13:53:16 out that that could be the name of a song, right?
13:53:18 A. Again, I don't know that it went through the DNN to
13:53:21 come up with this, it could have gone in through a rule or
13:53:25 gone in through an FST.
13:53:28 Q. And the result, though, is that the NLU has already
13:53:37 determined the meaning of the word six in each of these
13:53:41 possible contexts, right?
13:53:43 A. Those different -- those different meanings are
13:53:45 hypothetically in an NBest list.

Q. **They are meanings for the word six in each of the possible contexts, right?**

A. **I think that's what I just said.**

Q. **Now -- and at this point, we haven't used any short-term shared knowledge, right?**

A. **I don't know that that's -- that that's true. The NLU uses information from prior speech acts and from other pieces of information.**

Q. **Well let's look at that. Let's look at PTX-226, Mr. Patterson. Let's look at page 32. This is that multi-turn example you talked about, right?**

A. **Yes.**

Q. **And in fact, the user is saying six, that's what we saw on the top of your slide, PDX3-12, right, the result -- that six corresponds to the user saying six in this example, right?**

A. **They're both saying six.**

Q. **Now, if we look in the same document and we go to the next slide, 33, it explains that when the NLU processes that word six to come up with those different meanings, it says explicitly here, NLU models are currently not aware of dialogue context, right?**

A. **It says that, yes.**

Q. **This is saying in fact, when Alexa interprets the word six, it's not using any of that short-term shared**

knowledge about what the user had said before; right?

A. That's not how I interpret this, I interpret this as saying the NLU model, the statistical models, the DNN's, may not be aware of dialogue contexts but the rules and FST's are.

Q. You haven't seen any Alexa source code or documentation at least that you have shown the jury that says that rules are FST's are where dialogue context, have you?

A. I haven't shown the jury any source code.

Q. Or technical documents that said rules and FST's use dialogue contexts, right?

A. That's not true. This document has examples of rules and FST's running within the NLU along with models, and it's showing use of dialogue contexts in this very example. So from that I think it's quite reasonable --

Q. Let us walk through this example here to be a little clear.

A. Okay.

Q. Because what happens in this example, is we run six through the model with no dialogue context, as the Amazon document itself says. We get to the next slide, 34, please, Mr. Patterson. And this is what happens, right? The recognizer generates these intents and slots, these contexts 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 A. **It's a NBest list of hypotheses.**

14:00:58 Q. **Right, so these are exactly the hypotheses that are going to be output from the NLU, right?**

14:01:03 A. **This is showing what the NLU comes up with. I don't -- I think calling that establishing the meaning is misleading.**

14:01:08 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?**

14:01:12 A. **I'm sorry, we're talking about the '681?**

14:01:17 Q. **Claim 13.**

14:01:19 A. **Okay.**

14:01:20 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?**

14:01:49 A. **I don't remember -- I know I talked about it, I don't remember whether it was in the slide or not.**

14:01:55 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**

14:02:00

14:02:01

14:02:06

14:02:07

14:02:08

14:02:23

14:02:46

14:02:51

14:03:01

14:03:02

14:03:05

14:03:07

14:03:12

14:03:20

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

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

14:07:28 other than the ones that I presented.
14:07:29 Q. So you acknowledge that you did not show the jury
14:07:33 that this required element of the claim is actually done by
14:07:37 Alexa, right?
14:07:38 A. Alexa is in the NLU as we saw from the documents, is
14:07:46 -- is disambiguating words that have multiple meanings based
14:07:53 upon different pieces of information.
14:07:57 Q. Okay. But that is not what this claim requires. The
14:07:59 claim does not say disambiguate words using different pieces
14:08:05 he have information, that is not what this claim requires,
14:08:07 the claim requires these specific steps and you haven't
14:08:10 shown that Alexa does any of them, and specifically not this
14:08:19 one; right?
14:08:20 Let's go forward to the last element -- well,
14:08:38 let's go forward to the last element which is to generate
14:08:44 response to utterance, this would be PDX3-9, I think, can
14:08:52 you see that, Mr. Patterson. Sorry, that's not the right
14:09:09 number. We don't have a slide for this either. Let's go
14:09:15 back in to the claim itself, Mr. Patterson. Here we go.
14:09:24 So this claim element requires not just
14:09:28 generating a response, but it requires wherein the
14:09:33 conversational speech engine grammatically or syntactically
14:09:39 adapts the response based on the intended meaning, do you
14:09:43 see that?
14:09:43 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 grammatically or syntactically adapts a response, correct?
14:11:22 A. I don't recall specifically.
14:11:24 Q. You haven't seen any testimony from any Amazon
14:11:28 witness or engineer that says that Amazon grammatically or
14:11:33 syntactically adapts a response, correct?
14:11:35 A. They haven't so-they don't use those words, they say
14:11:38 that the response is in response to what the user has
14:11:42 requested.
14:11:43 Q. And the only evidence you provided to this jury today
14:11:47 is this diagram that doesn't say anything about
14:11:52 grammatically or syntactically adapting a response, correct?
14:11:57 A. I provided this and the fact that Amazon is
14:12:00 responding to a request from a user.
14:12:02 Q. Okay. Let's move to the '176 patent. If we look at
14:12:22 slide PDX3-32. You put up this slide in this element about
14:12:39 interpreting the recognized words or phrases wherein
14:12:45 interpreting the recognized words or phrases includes
14:12:49 establishing a context, do you see that?
14:12:51 A. Yes.
14:12:51 Q. And your example utterance was Alexa, I want to buy
14:12:57 an iPhone case; right?
14:13:08 A. That was one example I gave, yes.
14:13:10 Q. Right. When you say Alexa, I want to buy an iPhone
14:13:15 case, this context interpreter is not used at all?
14:13:20 A. This particular example here that was shown was

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

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

14:16:51 **ASR was already provided as input to the NLU.**

14:16:55 Q. **Let's be clear. You don't have any evidence that**

14:16:59 **re-ranker somehow operates because a user indicates that**

14:17:05 **Alexa made a mistake, right, you just made that up?**

14:17:09 A. **That the user said it? I don't think I said that**

14:17:13 **because the user said it made a mistake.**

14:17:15 Q. **You don't identify any actual predetermined event**

14:17:18 **that would cause Alexa's NLU to reinterpret the words or**

14:17:25 **phrases, correct?**

14:17:26 A. **It says that their system features and various kinds**

14:17:30 **of input including context that would cause it to re-rank.**

14:17:33 Q. **That wasn't the question. You haven't identified a**

14:17:37 **single predetermined event that would could cause Alexa to**

14:17:40 **somehow reinterpret the words or phrases, correct?**

14:17:43 A. **I haven't picked one out, no.**

14:18:07 Q. **Let's look at slide PDX3-42. This one that interpret**

14:18:14 **the natural language utterance based to the advertisement**

14:18:16 **and responsive to the existence of a pronoun in the natural**

14:18:20 **language utterance, determine whether the pronoun refers to**

14:18:22 **one or more of the product or service, and you understand**

14:18:27 **that for Alexa's NLU to interpret a pronoun, a speechlet has**

14:18:36 **to first provide a dialog act, right?**

14:18:42 A. **I'm sorry, say that again.**

14:18:43 Q. **Sure, for Alexa's NLU to actually determine the**

14:18:48 **meaning of a pronoun if it's going to do that, it has to**

first receive a dialog act from the speechlet, right?

A. Dialog act from the speechlet. I'm not sure I'm following what you mean in this context here.

Q. Sure. So you actually have no evidence other than what you thought you heard from Alexa, whether or not the NLU actually determines whether the pronoun refers to one or more of a product or service, correct?

A. So, I don't know whether it's -- I don't know what it's done with the pronoun, I don't know that -- what it thinks about the pronoun, what I know and what I have shown here, is that it is responsive to the existence of the pronoun in the natural language utterance. So the pronoun exists in one utterance, doesn't exist in another, and there are different responses.

Q. Right. But you don't know whether in either of these cases the pronoun is actually used by the NLU, is used to identify a product or service, right?

A. I don't know -- in this particular case, I didn't analyze any deeper than seeing that it clearly is doing exactly what's in the claim element.

Q. Well it's not, right, because this claim element requires that the system, right, Alexa, do a determination, which is whether the pronoun refers to a product or service. And we both know from your deposition and the evidence in this case, that when a user says buy it now, Alexa ignores

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

14:22:00 A. **It's clearly showing that when I use the pronoun, that I'm talking about the product that I'm trying to purchase.**

14:22:11 Q. **Well that's not true. As far as the NLU knows, this could be just like buy now or buy it now, the only difference in this example is one of those utterances got routed to something other than shopping; right?**

14:22:27 A. **I think the behavior of the system speaks for itself.**

14:22:32 Q. **Well, let's be clear, you haven't identified any Amazon engineer testimony that supports that the pronoun is being resolved in this example, right?**

14:22:45 A. **You said the pronoun is being resolved --**

14:22:48 Q. **Sure, you have no evidence from this slide from any Amazon engineer that this pronoun is being resolved in the way the claim requires; right?**

14:22:57 A. **I have no testimony from an Amazon engineer about it.**

14:23:03 Q. **And you have no document from Amazon --**

14:23:05 **THE COURT: I think he was --**

14:23:07 **You can finish your answer.**

14:23:11 **THE WITNESS: Thank you, Your Honor. I think that the claim doesn't say resolve, it's clear that the presence of the pronoun tells Alexa that we're talking about the thing that I'm buying it, and you can tell that by the fact that it's giving me the color stone, which is the color of the thing that is being purchased.**

BY MR. HADDEN:

Q. Let's be clear. You do not have any Amazon document that says that Amazon's NLU determines whether the pronoun it refers to anything; correct?

A. I am not presenting -- I am not presenting Amazon technical documents on this.

Q. You are not presenting Amazon source code that would show that the pronoun is determined to refer to anything; right?

A. I'm not presenting source code here.

Q. Let's look at the '703 patent. Now, the utterance that you used when you were explaining this example was Alexa, I want to buy an iPhone case, right?

A. Yes, I believe so.

Q. And when you say Alexa, I want to buy an iPhone case, you don't have any evidence to show that Alexa would obtain your shipping information; right?

A. Well it shipped it to my home, so --

Q. Let's be clear. Let's look at slide PDX3-55, please, Mr. Patterson. So if we look at the element obtain without further user input after the receipt of the user input shipping information, do you see that?

A. Yes.

Q. Right. So this stuff about without further user input means that the user can't say anything else after they

said in your example Alexa, I want to buy an iPhone case; right?

A. So, when I say buy it now --

Q. That's not the utterance that got you started. You got to have the same utterance, right, we start receive a user input, comprising a natural language utterance. And what you said was, I want to buy an iPhone case, right?

A. I did say that, that's right.

Q. So that's the user input. Now without any other user input, we have to obtain shipping information. Right? Alexa doesn't obtain your shipping information when you say I want to buy an iPhone case, right?

A. I believe when I say buy it now, it then goes through and --

Q. Sir you're not reading the claim. It says you have to obtain without further user input after the receipt of the user input. And the user input was I want to buy an iPhone case, right?

A. And the user --

Q. Let me finish the question, please, so it has to obtain the shipping information at that point without any further user input. So you haven't shown and it's a true fact that Alexa does not obtain the shipping information when you say Alexa I want to buy an iPhone case; right?

A. That may well be, when I say Alexa buy it now, at

that point it obtains my shipping information and my payment information and it process all the way through.

Q. Right and when --

THE COURT: Hold on, you're still cutting him off when he's still talking, you finish and then you start. Go ahead, sir.

MR. HADDEN: Sorry, Your Honor.

THE WITNESS: When I say Alexa buy it now, at that point, it obtains my shipping information, it obtains my payment information and it completes the purchase without any further user intervention.

BY MR. HADDEN:

Q. Sure. But if we look at what the claim requires, it requires that the shipping information has been obtained after you say in your example Alexa, I want to buy an iPhone case because it explicitly requires without further user input. If you were going to say buy it now, that would be further user input; right?

A. I think buy it now is the user -- is the initial user input.

Q. That's not what you told us this morning. This morning you told us, it was I want to buy an iPhone case. So let's say instead you said buy it now. If you say buy it now, the rest of the claim doesn't work because you don't identify the product based on buy it now; right? You need

14:28:59 **to pick one input and stick with it through the claim.**

14:29:16 A. **I think when I said I want to buy an iPhone case,**

14:29:21 **that was establishing various information, short-term**

14:29:24 **information, and I said buy it now, that was the utterance**

14:29:30 **that's being talked about here, and from that it completed**

14:29:34 **the purchase.**

14:29:38 Q. **Well, the problem is if you look at the identifying**

14:29:43 **stuff, right, identify without further user input after the**

14:29:48 **receipt of the user input, the product or service to be**

14:29:54 **purchased, right? Now the product you're trying to buy, the**

14:29:58 **iPhone case, it had already been identified before you said**

14:30:02 **buy it now. That's why in your presentation to this jury**

14:30:05 **this morning, you said instead, Alexa, I want to buy an**

14:30:09 **iPhone case, right?**

14:30:10 A. **I think you're misinterpreting here. The identify,**

14:30:15 **obtain and obtain steps, there is an identify, and then the**

14:30:20 **next step is obtain, and the next was obtain, those are all**

14:30:24 **internal. So the obtain, I have already provided payment**

14:30:27 **information to Amazon, so when it says obtain, I have**

14:30:31 **already provided that on the website. And when it says**

14:30:34 **obtain payment information, it's getting it from the**

14:30:39 **database. And when I say obtain -- obtain without further**

14:30:43 **user input shipping information, it's doing that from my**

14:30:49 **account, the long-term information that it has with Amazon.**

14:30:52 **So those identify, obtain, obtain, are internal.**

14:30:57 Q. **Understood. I'm going back to what your input was. Your input, Alexa I want to buy an iPhone case, is that where we start this claim, or do we start with the claim with the buy it now?**

14:31:11 A. **I think for the purposes of this analysis, it's Alexa buy it now.**

14:31:14 Q. **So if it's buy it now, nowhere did you show the identify without further user input receipt of the user input the product or service to be purchased. Right? That's not identified by buy it now, right?**

14:31:31 A. **That information is already part of the context of the conversation. I've already established it and the 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, please, Mr. Patterson. Now, this required that the system, Alexa, request, provide a request for user confirmation to use the payment information and the shipping information for a purchase transaction, correct? And you didn't show that at all; correct? You showed this slide. Nowhere did you show Alexa saying Dr. Polish, please confirm this payment information and the shipping information, right?**

14:32:41 A. **I think that's right, I think the receipt was sent -- 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 A. **So, the Speaker Voice ID is a -- its a component of**
14:34:20 **the Alexa system that you have to enroll, when you speak a**
14:34:25 **number of sentences, I think it was in my example, and it**
14:34:29 **learns your voice so that later on when you use the system,**
14:34:35 **it identifies you as you, and can customize the operation of**
14:34:39 **it for you.**
14:34:39 Q. **And is that long-term shared information that's used**
14:34:45 **in each communication with Alexa thereafter?**
14:34:48 A. **Yes, absolutely.**
14:34:49 Q. **Now, there was a discussion regarding -- if we could**
14:34:54 **have slide PDX3-55, on the screen. Do you recall there was**
14:35:10 **a discussion, if we could just focus on the right side for a**
14:35:13 **moment. There was a discussion regarding the I want to buy**
14:35:18 **an iPhone case and buy it now. Do you recall that?**
14:35:21 A. **Yes.**
14:35:21 Q. **And this is after you had already done the speaker**
14:35:24 **ID?**
14:35:24 A. **Yes.**
14:35:25 Q. **And so at the time you did this transaction because**
14:35:28 **of the speaker ID, what did Alexa know?**
14:35:31 A. **It knew it was me, and it had my payment and shipping**
14:35:35 **information.**
14:35:35 Q. **Okay. If we go to the bottom on there. The**
14:35:39 **statement buy it now, do you see that we had some discussion**
14:35:44 **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 understand it would refer to?**

14:35:51 A. **Well, it was referring to the previously established iPhone case that I had asked for.**

14:35:56 Q. **So when you said buy it now, that phrase, did Alexa 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 PTX-226, this is the natural language deep dive that has 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 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 given in the Amazon, actual Amazon technical document, shall**

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

14:38:15 **at a voice input device?**

14:38:17 A. **Yes.**

14:38:18 Q. **Now, let's go to see what happened when you say that,**

14:38:23 **Alexa says for what time? Do you see that?**

14:38:26 A. **Yes.**

14:38:26 Q. **And now the user says six. Do you see that?**

14:38:30 A. **Yes.**

14:38:31 Q. **Now, let's go to what is shown in the Alexa technical**

14:38:41 **document. If we can go to slide 36. At slide 36, the**

14:38:50 **statement six is made, do you see that?**

14:38:53 A. **Yes.**

14:38:53 Q. **And at the time the word six is made in this**

14:38:57 **conversation with Alexa, has short-term knowledge, shared**

14:39:02 **knowledge been accumulated about the current conversation?**

14:39:05 A. **Yes.**

14:39:06 Q. **And what is that?**

14:39:07 A. **Well, that we're working on an alarm and that we're**

14:39:13 **setting a notification.**

14:39:15 Q. **Right. And in fact, the previous intent was set**

14:39:19 **notification intent, that's to set the alarm?**

14:39:22 A. **Right, that's at least in part the context and it is**

14:39:25 **the short-term information.**

14:39:26 Q. **If we can highlight short-term shared knowledge about**

14:39:30 **the current conversation. And if we can highlight**

14:39:46 **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 A. **Then it changes the slot to be time equals 0600.**

14:41:11 Q. **Okay. And so looking at the next element, it says identify a context associated with the utterance, did Alexa do that?**

14:41:22

14:41:24 A. **Yes.**

14:41:24 Q. **Did it use both the short-term and long-term knowledge?**

14:41:27

14:41:27 A. **Yes.**

14:41:27 Q. **For example, with regards to the long-term knowledge, would that be the speaker ID?**

14:41:30

14:41:32 A. **Yes.**

14:41:33 Q. **And if you were changing the alarm, would the long-term knowledge include the previous alarm?**

14:41:36

14:41:39 A. **Yes.**

14:41:40 Q. **Now, if you look at the next element, establish an intended meaning for an utterance within the identified context, as shown in slide 38, does Alexa do that?**

14:41:44

14:41:46

14:41:53 A. **Yes.**

14:41:54 Q. **And if you go to the last element, generate a response to the utterance, if we can now go back to slide 32, after the user says morning, what does Alexa do?**

14:41:57

14:42:01

14:42:08 A. **Well, it has figured out the slot now and it comes back with alarm set for 6:00 a.m. tomorrow.**

14:42:13

14:42:16 Q. **So does Alexa practice that element in the example that you provided?**

14:42:20

14:42:20 A. **Yes.**

14:42:22 Q. **Now, there was some discussion about whether or not Alexa might ask the user to help it resolve ambiguity, do you recall that?**

14:42:35 A. **Yes.**

14:42:35 Q. **You can take that down, Mr. Smith. If we could go to slide 24.**

14:42:36 **Actually, go to slide 23, please. This is on PTX-226, slide 23, I believe, PTX, the actual exhibit. Next page. And next page. There it is.**

14:42:40 **Are there situations where for example in the utterance play Hunger Games that Alexa would find what they call true ambiguity?**

14:42:56 A. **Yes.**

14:43:09 Q. **In this situation here, the scores were close, .82 and point .81.**

14:43:13 A. **Right.**

14:43:17 Q. **Because the scores were so close, Alexa couldn't make a call as to what would be the blue arrow, is that fair to say?**

14:43:22 A. **That's right.**

14:43:24 Q. **If you go to the next page. When Alexa faces true ambiguity, does it ask the user to help it correct it?**

14:43:24 A. **Yes, this is an example of it, where it's asking the user for some information to help it out.**

14:43:30

14:43:32

14:43:32

14:43:36

14:43:39

14:43:39

14:43:40

14:43:45

14:43:49

14:43:52

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?

A. **Yes.**

Q. **Now, do you see the title of this slide is a two-step approach?**

A. **Yes.**

Q. **And what was Step 1?**

A. **So, Step 1 it says that recognizer generate the yes intent, no intent, or slots for the current turn utterance.**

Q. **And that's only the first step of the two-step approach, correct?**

A. **Yes.**

Q. **And if you go to the next slide. What occurs in Step 2?**

A. **In Step 2, the context interpreter merges the yes, no, multi-turn intent with previous intent and slots so it takes the previous intent that was somehow ambiguous and merges it with the results here to disambiguate.**

Q. **If we go to the next -- page 36. And that Step 2 is what we were discussing with regard to the clock setting example, correct?**

A. **Yes.**

Q. **Take that down.**

Just two more questions, Dr. Polish. Now, you answered multiple questions regarding source code during the cross-examination, do you recall that?

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 code, he was being communicated with by the lawyers. I saw the questions that were being provided to Mr. Peck and I added my own questions and consulted on what he was -- what he was tasked with looking for. And then I saw his results 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 identified in his report?**

14:49:05 A. **Yes.**

14:49:05 Q. **And you understand that Mr. Peck will be a witness in 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, 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 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 A. **Yes. My name is John Charles Peck, Jr.**

15:12:04 Q. **And generally what's your role in this case?**

15:12:08 A. **On this case I have been reviewing Amazon source code, Alexa source code.**

15:12:14 Q. **And what's your educational background, briefly?**

15:12:18 A. **Yes, so I have a bachelor of science in computer engineering in Carnegie Mellon University and I have a master of science from computer science from the University of California.**

15:12:33 Q. **Mr. Peck, what's your work experience be like since I graduated college?**

15:12:37 A. **I have been a practicing computer engineer for over twenty-five years, I have developed a lot of software and hardware products.**

15:12:46 Q. **And do you have any experience with voice related technology, like at issue in this case?**

15:12:57 A. **I do. I had a client several years ago who was interested in ASR technology. They wanted me to investigate and profile a software package called Sphinx, it performs ASR.**

15:13:13 Q. **Do you have any experience with big computer systems with lots of servers?**

15:13:17 A. **I do. Part of my prior employment at Foxfire Technologies Corporation, and Transarc Corporation which made distributed file systems, and also with a company**

called Adion Systems, a client that interacted with servers on a network.

Q. And during your professional career, have you ever been awarded patents?

A. Yes, I have. I am the named inventor on ten U.S. patents issued.

Q. And in what kind of expertise are you bringing to this case?

A. In this case I performed a source code review of the Amazon Alexa source code.

Q. And have you done source code review like this in other litigation matters?

A. Yes, I have. I have performed source code review on several other litigation related matters in the past.

MR. SMITH: Your Honor, at this point, VoiceBox offers Mr. Peck as an expert witness in the field of computer source code analysis.

MR. HADDEN: No objection.

THE COURT: All right. He will be so recognized.

MR. SMITH: Thank you, Your Honor.

BY MR. SMITH:

Q. Mr. Peck, first of all, what is source code?

A. So source code are a set of commands that are given 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?**

A. **Yes, I was.**

Q. **And were you asked to compare the source code to the patent claims like -- or to the patent claims in this case?**

A. **No, I was not.**

Q. **What was your role as you understood it relative to Dr. Polish?**

A. **I understood that Dr. Polish would be performing that comparison between the source code and the claims.**

Q. **And aside from looking at source code, did you look at any other types of evidence in the case?**

A. **Yes. I attended depositions from Amazon's witnesses. I read several deposition transcripts. And I reviewed Amazon technical documents.**

Q. **And those depositions, what sort of employees were you going to depositions for?**

A. **These were Amazon engineers who worked on the Amazon Alexa system. There were several witnesses who worked on different parts of the Amazon Alexa system, including the NLU, the ASR, and something called ranking arbitration as well as the shopping.**

Q. **Based on your analysis of the source code, would you say that the diagram that Amazon has been showing that they say indicates Alexa, was that accurate based on the source code analysis that you did?**

A. **No, I don't think that's entirely accurate.**

Q. **Okay. Well, for one thing, did you take a look at a source code for what we have been referring to as the ASR in this case?**

A. **Yes, I did. I reviewed source code for two different versions of Amazon Alexa ASR, there was an older version and a newer version that was put to use later.**

Q. **And did you review any--strike that.**

Did you find any source code for an NLU?

A. **Yes, I did.**

Q. **And did you review that code for the NLU, too?**

A. **Yes, I reviewed that source code for the NLU.**

Q. **And were you able to identify source code for reusing information from a prior utterance?**

A. **Yes, I did identify a source code that uses information from a prior utterance.**

Q. **Now, Mr. Peck, I'm going to show you a document here on the screen here. My understanding is it's sensitive information from Amazon. So we don't have extra copies. But I'm just going to put it up on the Elmo for you to take a look.**

And first of all, could you take a look at the bottom here and just identify the exhibit there, exhibit number there in the middle?

A. **Yes. So this source code is stamped with a Bates 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

something that is public class, contextual intent carrier IPMPL. Do you see that?

A. Yes.

Q. What does that refer to?

A. This is the name of this class.

Q. And what does this class do generally.

A. So this is class is deciding which intents are to be carried forward into the current set of intents being considered.

Q. Why don't we go down here do you see line 71 where it says intent carryover engine, rules engine, do you see that?

A. I see that.

Q. And based on your analysis, did you determine that Alexa as an entire engine devoted to rules?

A. Yes, in this particular model, there is a particular engine put in use that uses rules to perform its function.

Q. Is this file that we just looked at Exhibit 541(b) another example of Alexa using manual rules in 2020 or later?

A. Yes, it is.

MR. SMITH: Thank you, Mr. Peck. Pass the witness.

THE COURT: Cross.

MR. HADDEN: You can leave that up.

CROSS-EXAMINATION

15:29:58 **BY MR. HADDEN:**

15:30:02 Q. **Good afternoon, Mr. Peck.**

15:30:04 A. **Good afternoon.**

15:30:06 Q. **Now, you're not an expert in Natural Language Understanding, are you?**

15:30:12 A. **I'm not an expert on Natural Language Understanding, I'm a practicing engineer of over 25 years with experience in software and hardware development.**

15:30:22 Q. **Thank you, sir.**

15:30:23 **And you're not offering any opinion to the jury here today that Amazon infringes any of these patents, are you?**

15:30:31 A. **No, I am not.**

15:30:33 Q. **And you're not offering any opinion to the jury today that these patents are valid, are you?**

15:30:38 A. **No, I am not.**

15:30:40 Q. **Thank you.**

15:30:41 **Your assignment, if I understand it correctly, was to answer questions from VB Assets's lawyers, is that right?**

15:30:53 A. **Yes, that's correct.**

15:30:54 Q. **And Dr. Polish didn't ask you any questions, did he?**

15:31:00 A. **I received communication from counsel about what parts to look at. Dr. Polish testified today about that process.**

Q. **Okay. But when you were doing your analysis, you weren't talking to Dr. Polish, correct?**

A. **No, I didn't talk to Dr. Polish directly. My instructions came from counsel.**

Q. **Can we look at JTX-001, please, Mr. Patterson. Claim 13.**

A. **I have that in front of me.**

Q. **Now is it fair to say that between you and Dr. Polish, you're the one who knows most about how Alexa source code operates?**

A. **I would say I reviewed the source code for Amazon Alexa.**

Q. **If we look at this Claim 13, did VB Assets' counsel, did they ask you to go find out if there was Alexa source code that accumulated short-term shared knowledge about the current conversation?**

A. **So I recall my report addressing questions that counsel presented me, those particular questions are in my report. I don't recall this point being a particular question being asked.**

Q. **Okay. So even though this claim requires computer code that accumulates short-term shared knowledge, VB Assets's lawyers didn't ask you to try to find that in Alexa; correct?**

A. **No, the questions that I was asked, I don't think**

relate to the claim language here. They were another terms that were questions.

Q. If we look at the next element C, which requires accumulating long-term shared knowledge, is it also the case that VB Assets's lawyers didn't ask you to find out whether Alexa has source code that accumulates long-term shared knowledge?

A. I don't remember a question with those exact terms in it.

Q. Okay. If we go to the next element, identify a context associated with the utterance wherein the conversational speech engine identifies the context associated with the utterance from the short-term shared knowledge and the long-term shared knowledge, is it also the case that VB Assets's lawyers didn't ask you to find out whether or not the Alexa source code does that?

A. Yeah, I don't remember a question posed in those terms.

Q. Okay. And the same with element E, establish an intended said meaning for the utterance, and then continues, do you see the text there, sir, is it true that VB Assets's lawyers didn't ask you to try to find out whether the Alexa source code does this?

A. That's correct, I don't recall such a question being 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,**

Mr. Patterson. This is JTX-003, Mr. Peck. Look at Claim 40. Claim 40 has this piece at the end, an adaptive misrecognition engine configured to determine that the conversational language incorrectly interpreted the words or phrases, and it continues, do you see that language, Mr. Peck?

A. I see that.

Q. Did VB Assets's lawyers ask you to try to figure out whether or not the Alexa source code performs that step?

A. No, I don't recall a question like that being posed to me.

Q. Let's look at JTX-005, sir, this is the '097 patent. If we could go to Claim 23. This claim it requires the last element, requires interpreting the natural language utterance based on the advertisement and responsive to the existence of a pronoun in the natural language utterance, determine whether the pronoun refers to one or more of the product or service, or a provider of the product or service?

Did VB Assets's lawyers ask you to try to find out whether or not the Alexa source code did that.

A. No, I don't recall a question put in those terms to me.

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

language understanding.

Q. Dr. DiCristo, during the time at VoiceBox, you did lead a large team of engineers and scientists; correct?

A. I did lead a large team of people, which included engineers, scientists, program managers, TPM, and so on, technical program managers and so on.

Q. What was VoiceBox Cybermind?

A. Cybermind was a product that we developed and that leveraged existing technologies such as speech recognition and Natural Language Understanding, and allowed people to -- to communicate with the machine and turn some command into actions.

Q. When you prepared your LinkedIn profile, you believed that the Cybermind was an Amazon-like system; correct?

A. Well, I think the "like" -- you know, when we say like, of course you can take into different interpretation here. It's open for interpretation.

But on the surface, they have like appearance. Of course, their implementation could be quite different.

Q. But you described Cybermind as the Amazon Echo-like system; correct?

A. Again, I think like is very -- it's Echo-like. Right. Like, is there -- there is different flavors of like.

And as -- as a person, as a customer without any

technical background, you could consider that they look alike.

Q. And they operated alike, correct?

A. That would -- that's actually very vague. They are -- I think they are -- they are very different on the -- on the way they are implemented in terms of code is different. The system that they use to recognize speech is different. There a lot of differences that we could go into.

Q. But the Cybermind, the first version of it came out in 2006, correct?

A. So the Cybermind product was actually never productized, but it was demoed. And I definitely -- I definitely think I remember even from the video that in 2006, we had something that looked like that.

Q. And the -- to your knowledge, the Amazon Echo product was publicly announced in the 2014 time frame; correct?

A. I don't recall when exactly the Amazon system came out, so I think it's hard for me to say. 2014 or 15 or around that time frame.

Q. In 2013, IEEE ranked VoiceBox Technologies number 13 in patent power for the computer software industry. Do you see that?

A. Yes, I do.

Q. And you put that statement in your LinkedIn profile; 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 frame, the IEEE had ranked VoiceBox Technologies as having the 13th most significant patent portfolio in the computer software industry; correct?**

15:44:00 A. **So I remember actually seeing a document that was actually provided by Mike Kennewick. That was a link where the IEEE had ranked different companies of something they call patent power. And that seems like a good thing to put on your LinkedIn profile.**

15:44:19 Q. **So sir, you -- you began working after VoiceBox for 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, -- January 1st. I'm sorry. I don't know why I said July 1st. January 1st, 2022.**

15:44:51 Q. **Okay. Dr. DiCristo, looking at Exhibit 48, do you see the top e-mail is from Rich Kennewick, dated April 30th, 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 steps-VoiceBox/Intel." Do you see that?**

15:45:14 A. **Yes, I see that.**

15:45:15 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

high level, the context is general information of prior event, basically, at the time of the request.

Q. And those prior events could include prior conversations, correct?

A. Again, it's prior information that was available before that -- that I know for -- that -- that conversation.

Q. And so prior information could include information that came from prior conversations, correct?

A. I'm not sure I understand the question, because I guess, like I said, it depends on how we define context. As I mentioned, I define context as prior information that was either provided by the system or by a person. And an agent could leverage that information if they were designed to do so, or if that was relevant.

Q. And looking at Item 10 in the Intel documentation, it says: "Learning from interactions-wants system to learn from previous conversation."

Do you see that?

A. I can see that.

Q. And one of the capabilities that the VoiceBox system had and its agents had was the capability of learning from previous conversations, correct?

A. So I guess it's hard for me to answer, because it's -- like you see, learn is in quote marks, so learn has a lot of vague notion of it. But I can say that actually, there

15:49:39 were no artificial intelligence in that -- in that -- at
15:49:44 that time.
15:49:44 Q. But the operation of the VoiceBox system could adapt
15:49:48 over time based on prior conversations, correct?
15:49:52 A. I don't recall that capability.
15:49:57 Q. And Dr. DiCristo, why don't you just take a moment to
15:50:01 look at Exhibits 55, 56, and 57. And when you've had a
15:50:06 chance to look at those text messages, I'll ask you a
15:50:09 question. Okay?
15:50:11 A. Okay.
15:50:12 Q. Dr. DiCristo, these are texts that you exchanged with
15:50:17 Kelley Price in the January 2006 to January -- sorry --
15:50:23 January 6, 2017, to January 9, 2017, time frame; correct?
15:50:30 A. Yes, this is correct.
15:50:35 Q. And these are screen shots of the texts that you
15:50:42 exchanged, correct?
15:50:43 A. This is correct.
15:50:45 Q. And sir, with regards to this, Ms. Price was a
15:50:49 VoiceBox employee; correct?
15:50:50 A. Kelley Price was a -- dealing with human resources,
15:50:55 as I previously stated.
15:50:57 Q. And so, she was the human resources at VoiceBox;
15:51:02 correct?
15:51:03 A. She was a human resources at VoiceBox. She was a
15:51:07 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 **exhibit PTX-42, 48, 55, 56, 57, and 58.**

15:54:28 **MR. HADDEN: No objection.**

15:54:31 **THE COURT: All right. Thanks. They are**

15:54:32 **admitted.**

15:54:33 **(PTX Exhibits No. 42, 48 55, 56, 57 and 58 were**

15:54:34 **admitted into evidence.)**

15:54:34 **MR. YOON: And now, Your Honor, we would like to**

15:54:36 **play to deposition testimony of Rino Caruccio, who is the**

15:54:41 **vice-president of financial analysis for devices at Amazon.**

15:54:46 **THE COURT: All right.**

15:54:48 **(Videotape deposition of Rino Caruccio:)**

15:54:50 Q. **Mr. Caruccio, Exhibit 173 is an e-mail from someone**

15:54:55 **named Andy Vu dated September 8, 2021, with the subject**

15:55:02 **line, Re: OP1 outputs review, ten-year P&L, LTVLF and**

15:55:11 **IBVLF. Do you see that?**

15:55:11 A. **Okay. I have it, yes. I see the e-mail.**

15:55:14 Q. **Mr. Caruccio, are you familiar with an OP1, what an**

15:55:21 **OP1 is at Amazon?**

15:55:23 A. **Yes, I am.**

15:55:24 Q. **What is that?**

15:55:25 A. **OP1 is a mechanism within the Amazon company usually**

15:55:30 **held in September/October time frame whereby businesses,**

15:55:40 **products, review their strategy including investment needed,**

15:55:45 **financial projections, financial forecast for the next 12 to**

15:55:51 **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 a theme in terms of determining where to invest and where to
15:58:19 prioritize.
15:58:25 (End of videotape.)
15:58:28 MR. YOON: Your Honor, at this time, we would
15:58:30 seek to move into evidence exhibit PTX-173.
15:58:34 MR. HADDEN: No objection.
15:58:35 THE COURT: Thank you. It is admitted.
15:58:35 (PTX Exhibit No. 173 was admitted into
15:58:39 evidence.)
15:58:39 MR. YOON: At this time, Your Honor, the
15:58:40 plaintiff will now present the video deposition testimony of
15:58:45 Vinod Krishnan, who is the director of engineering and
15:58:49 science for Alexa Shopping at Amazon.
15:58:54 (Videotape deposition of Vinod Krishnan:)
15:59:00 Q. All right. In your folder, for the record,
15:59:03 Exhibit 78 is a document about with Bates numbers ending in
15:59:09 58040. Does the Alexa Shopping team use Wikis to document
15:59:15 the Alexa Shopping technology?
15:59:21 A. As a team, Alexa shop be uses many -- many
15:59:27 technologies to capture the thought process. Wikis could be
15:59:41 one of those.
15:59:42 Q. And so Exhibit 78, we're looking at a Wiki for a
15:59:46 document in the technology of Alexa Shopping, right?
15:59:50 A. It looks like this Wiki, or this document, captures a
15:59:55 version of -- of Alexa Shopping.

16:00:04 Q. **And for the record, Exhibit 80 is an e-mail and attachments, the Bates number on the first page ends in 342249.**

16:00:17 **Mr. Krishnan, do you recognize Exhibit 80 to be an e-mail?**

16:00:23 A. **Exhibit 80, the first part of Exhibit 80 is an 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, 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 2018?**

16:01:02 A. **Yes.**

16:01:04 Q. **And if you look at -- so Michelle Solano sent the e-mail and she says please find attached the docs for today's 2:00 p.m. PT 2019 NA consumer OP1: Alexa Shopping, 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 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 an ending Bates number of 85508. And do you recognize this to be a Wiki page from the Alexa Shopping Wiki that is entitled "Dialog Platform"?**

A. **Yes.**

Q. **And Exhibit 87 is -- Exhibit 87 is the -- this is an e-mail attaching the 2022 OP1; correct?**

A. **That's correct.**

Q. **An e-mail you received?**

A. **Yes.**

Q. **And just like the other OP1's, the 2022 OP1 was presented to Amazon senior leadership?**

A. **It was presented to Amazon consumer and Alexa senior leadership.**

(End of video.)

MR. YOON: Your Honor, at this time we would like to move into evidence exhibit PTX-78, PTX-80, PTX-84, and PTX-87.

MR. HADDEN: No objection.

THE COURT: Is it clear what the corresponding exhibit is with the script.

MR. YOON: I believe so, Your Honor.

THE COURT: I have been trying to follow it, I just want to make sure that some day we don't have a question as to what exhibit was talked about versus what exhibit you entered.

MR. YOON: Yes, the exhibits that were talked about in the transcript are the videos, are the ones that 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**

didn't have time to secure patents on our own technology, we wanted to make sure that we had some assets that we could counter claim or use in discussions or negotiations with others. And we bought a small patent portfolio with Comverse.

Q. Exhibit 158 is a document entitled objections and responses of defendants to plaintiff VB Assets LLC's first set of interrogatories, numbers 1 through 14?

A. I see that.

Q. And have you seen Amazon's interrogatory responses before today?

A. I may have seen one subsection of one of these, but I don't recall this one.

Q. Why don't we go to page 21 of Exhibit 158.

A. Okay.

Q. And the bottom of page 21, do you see there is a sentence that says "Amazon became aware of U.S. Patent Nos. '8,886,536; 8,073,681; and 9,015,049 in 2015." Do you see that?

A. Yes.

Q. And do you understand that the three patents I just quoted from the interrogatory response have been referred to in this case as the '536, the '681, and the '049 patents?

A. I do.

Q. Do you know how Amazon became aware of the those

three patents in 2015?

A. My understanding is those were cited in prosecution of our own patent applications.

Q. And does -- and do you have any other facts about the context for how those patents came to be cited in Amazon patents?

A. A patent examiner would have thought they needed to be cited and cited them. Often times it's got the right words in a search result. Sometimes it's something they are aware of. I don't know why or how that came about.

Q. Do you know if the citation was from the patent examiner or if it was from the patent prosecution counsel of Amazon?

A. I don't. I don't know.

(End of videotape deposition.)

MR. YOON: Your Honor, at this time we would like to move into evidence PTX-158.

THE COURT: So you got a forty-page document of their interrogatory responses that has a whole bunch of stuff in there that's not related to what he just talked about. Why are we doing this?

MR. YOON: I apologize. That was what was used at deposition. We can excerpt it so it's just one or two pages.

THE COURT: I'm not admitting this. You can

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