EXHIBIT A

THIS EXHIBIT HAS BEEN REDACTED IN ITS ENTIRETY

EXHIBIT B

IN THE UNITED STATES DISTRICT COURT

FOR THE DISTRICT OF DELAWARE

PERSONALIZED USER)
MODEL, LLP,)
Plaintiff,)
vs.) CA number 09-525 (LPS)
GOOGLE, INC.,)
Defendant.)

\- - - - -

VIDEOTAPED DEPOSITION OF JAIME CARBONELL

WASHINGTON, D.C.

NOVEMBER 27, 2012

The videotaped deposition of JAIME CARBONELL was convened on Tuesday, November 27, 2012, commencing at 10:05, at the law offices of SNR Denton, located at 1301 K Street, Northwest, in Washington, D.C., before Paula G. Satkin, Registered Professional Reporter and Notary Public.

\- - - - -

Job No. CS1565706

illegal and so forth. A disjunctive normal form allows you to combine one or more of these different target expressions into a single expression and to learn it from incoming data streams without having to first categorize it as to which type of potential fraud it might be. It's a theoretical method with practical applications.

Q. And your -- the machine learning artificial intelligence approach, did that involve estimation of probabilities?

A. Estimation --

MS. BENNETT: Objection. Form.

THE WITNESS: Estimation of parameters, which can be interpreted as probabilities.

BY MR. PERLSON:

Q. Okay. So it would have involved both estimating parameters and estimating probabilities; is that correct?

A. So the parameters are part of the learn model, the learn function, and the probabilities are probabilities of the outcome.

Q. And both -- and both of those aspects would have been part of what was

disclosed in your book; right?
A. Yes. It was disclosed subsequent to the book in more detail than in the original book. The original book was the start of the field.
Q. Right. It was proposed in -- okay. But let's then use the last version of the book. The last version of the book would have disclosed that; right?
A. It's -- yes. It -- I need to correct the presumption in your -- in your question.
Q. Sure.
A. It wasn't three versions of the same book. It was three volumes, three different books. And they taught different aspects much machine learning.
And aspect of estimating parameters was disclosed there, but it was developed in much greater detail in subsequent work post-1990.
Q. Okay. By you or others?
A. Yes and yes.
Q. Okay. And was -- does the Machine Learning and Artificial Intelligence Approach

discuss creating models using machine learning?

MS. BENNETT: Objection. Form.

THE WITNESS: Yes, it does.

BY MR. PERLSON:

Q. And what sort of -- does -- does Machine Learning and Artificial Intelligence Approach have -- discuss any practical applications of machine learning?

A. Yes. Several examples are given. Applications in disease diagnosis, in agriculture, in medical diagnosis. I guess that's similar to disease diagnosis. In planning such as robotic planning and so on.

Q. And was -- you would agree that in -- that throughout the 1990s, machine learning was applied to a variety of practical applications; right?

MS. BENNETT: Objection. Form.

THE WITNESS: Yes, sir.

BY MR. PERLSON:

Q. And that would have included search engines; correct?

A. No. I do not believe so. The applications to search engines started towards the end of the 1990s, not in the greater part of that decade.

Q. Fair enough. But machine learning was used in the context of search engines in -- let's say at least by 1998; correct?

MS. BENNETT: Objection. Form.

THE WITNESS: It was starting to be used.

BY MR. PERLSON:

Q. And -- and machine learning also would have been used in the context of personalization by 1998; correct?

A. No --

MS. BENNETT: Objection. Form.

THE WITNESS: Incorrect.

BY MR. PERLSON:

Q. Your -- you're telling this Delaware jury that in no circumstance was personalization used in machine learning by 1998?

A. No --

MS. BENNETT: Objection. Form.

THE WITNESS: I'm stating that the field did not focus on personalization in machine learning at that time. Personalization requires large amounts of information, data collected about individuals, about each specific individual in order to build a model that is tailored or customized or instantiated for that particular individual.

Prior to 2000, such information was almost never available. Search engines did not go around capturing the behavior of users at that time. Electronic commerce, Mechanical Turk for crowd sourcing, LinkedIn and Facebook didn't exist. So all the sources of electronic data in sufficient quantities to personalize was really not available in that decade.

BY MR. PERLSON:

Q. So in 1999 it really was not possible to personalize search engines; is that your testimony?

A. My testimony is that it was not at that time practiced.

Q. And -- but was it possible or not possible?

MS. BENNETT: Objection. Form.

THE WITNESS: In order to make it possible, it would have required a collection of all of this data. It would have required a method that was similar to the patents in suit, and these were not available at that time. These are preconditions.

BY MR. PERLSON:

Q. Okay. And do the patents in suit provide a way that you would be able to collect sufficient information in order to actually provide personalization of search results?

MS. BENNETT: Objection. Form.

THE WITNESS: The patents in suit teach how to collect the information unobstrusively -- in other words, "transparently" I think is the word that is used in the claim language -- based on user click-through data. That is one of the necessary ingredients in order to collect information in large quantity about multiple users so they could be able to personalize user models about each one. It also required other ingredients.

BY MR. PERLSON:

Q. What other ingredients?

A. It required a learning machine with a suite of parameters and a method of estimating values for these parameters based on the personalized user data that is collected, and then it required a -- that learning machine to be customized; in other words, begins with a generic learning machine set of parameters and there's a mathematical update function that looks at the incoming data, customizes the learning machine, and so now you have customized version of that learning machine per each user that has provided sufficient data. Those are all required ingredients.

It also requires analysis of the incoming information in the form of documents or electronic files in order to be able to use the learning machine to make a prediction, posterior probability, which is a degree of -- numerical degree of belief on whether that document would be of interest to the particular user for whom the learning machine was estimated -- the parameters of the learning machines were estimated.

Q. And -- so you mentioned documents, analyzing documents. Now, would an example of that would be a web page? Is that -- is that accurate?

A. If your question is whether a web page is an example of a document, the answer is yes.

Q. Okay. And in order to analyze the web page, you'd actually have to look at the page; right?

MS. BENNETT: Objection. Form.

THE WITNESS: Yes.

BY MR. PERLSON:

Q. You can't just look at the URL that's the pointer to the page. You would agree with that; right?

MS. BENNETT: Objection. Form.

THE WITNESS: That's essentially correct.

BY MR. PERLSON:

Q. That's essentially correct?

A. The URL provides a small amount of information pertinent to the page.

Q. Well, do you think just looking at the URL of a web page is sufficient to analyze documents in -- in the context of this patent?

A. If the document is a document to which the URL points -- is that what you mean by analyzing?

Q. Yeah.

A. No, I don't think it's sufficient.

Q. Now let's talk a little bit about the probability that you mentioned. You referred to estimating a probability of interest in a document; is that right?

MS. BENNETT: Objection. Form.

THE WITNESS: Yes.

BY MR. PERLSON:

Q. What is that?

A. Okay. What is what?

Q. What does that mean?

A. Interest in the document or probability?

Q. Estimating a probability?

MS. BENNETT: Objection. Form.

THE WITNESS: Estimating a probability means using available information usually in a Bayesian setting, which is the majority of settings under which -- Bayesian. I can spell it out for you later. Bayesian is a type of probability in statistics. It's the dominant type that is used here. So in order to answer your question, I need to make a couple of assumptions, and one assumption is that I am using probability in the same way that Dr. Jordan used it, your own expert, and so forth, which is the Bayesian way.

In that case you would combine the available data from the past. In other words, judgments made by that user. If you're estimating the probability of interest of a document for a particular user, judgments made by that user for previous documents. The similarity between those previous documents and new documents. This is why you need the content analysis in order to establish that similarity. And you would combine it in a mathematical form that would take the parameters, the values of the parameters, I should say, and from it compute a number, a numerical degree that would estimate the extent to which the new document is of interest to the user for which that model was made.

That number, if you want to be -- use strict probabilities, would then be renormalized to the 0-1 interval to get -- probability cannot be less than 0, it cannot greater than 1, or it can be left un-normalized, in which case it normally would not be called a probability in the technical sense but it would be called a probability in the popular sense, like a percentage.

Q. Okay. So you would agree, though, that the probability is a number between 0 and 1; right?

MS. BENNETT: Objection. Form.

THE WITNESS: The Court's construction is broader than that.

BY MR. PERLSON:

Q. Okay. Well, let me just first ask you, in the context of Bayesian statistics, probability has to be between 0 and 1; you would agree with that?

MS. BENNETT: Objection. Form.

THE WITNESS: In the technical definition, yes. In the popular definition, it doesn't -- is not confined to that.

BY MR. PERLSON:

Q. Okay. But in -- what do you mean "the popular definition"?

A. We use probabilities in the language more loosely than in the technical sense, and the Court's construction requires only it be a numerical -- measure the degree of belief.

Q. And in -- in the patents in this

A. Which one do you mean?
Q. Just the general concept.
A. No. They did not invent.
Q. I mean, transparently monitoring user interactions was known before the patents in this case generally?
MS. BENNETT: Objection. Form.
THE WITNESS: Generally? It was known at least to some people. Yes.
BY MR. PERLSON:
Q. It would have been known to someone of skill in the art; correct?
MS. BENNETT: Objection. Form.
THE WITNESS: Just transparently monitoring in general, it would have been, yes.
BY MR. PERLSON:
Q. And -- and certainly documents were -- documents that had been -- well, let me back up.
Monitoring documents that a user had interacted with would have also been known in the art before the patents; correct?
MS. BENNETT: Objection. Form.
THE WITNESS: Keeping track of which documents the user would have interacted

with was known prior to the patents.

BY MR. PERLSON:

Q. And using -- in connection with the transparently monitoring phrase, would you agree that if a user was required to indicate that -- well, scratch that.

You would agree that estimating parameters of learning machines at least generally were known to one skilled in the art before the patents?

MS. BENNETT: Objection. Form.

THE WITNESS: Not in the way that it was described in the patents, in the patent claims.

BY MR. PERLSON:

Q. But I'm just asking you generally. I mean, the estimating parameters of learning machines was known just generally to one of skill in the art; right?

MS. BENNETT: Objection. Form.

THE WITNESS: You take it out of the context of the patent and out of the context of search engines, out of the context of personalization, the answer is yes.

BY MR. PERLSON:

Q. And, in fact, estimating parameters of a learning machine generally is just kind of a basic concept of machine learning; isn't it?
MS. BENNETT: Objection. Form.
THE WITNESS: Yes.
BY MR. PERLSON:
Q. I mean, that would have been disclosed in your work in the machine learning --
A. Yes. Estimating parameters of a learning machine is part of the process of machine learning and has been since the '80s.
Q. Okay. And the same is true with estimating probabilities; right?
MS. BENNETT: Objection. Form.
THE WITNESS: Estimating probabilities as the output of a learning machine wasn't as widely practiced. It was narrower, but still -- it still was part of the prior existing art.
BY MR. PERLSON:
Q. You're familiar with the -- the WebWatcher prior art?
A. Yes.

Q. And the WebWatcher prior art used machine learning to determine a -- whether a user would be interested in a document; correct?
MS. BENNETT: Objection. Form.
THE WITNESS: No.
BY MR. PERLSON:
Q. Why not?
A. For actually both parts of your question is no.
Q. Okay.
A. It was not targeted at a specific user, and it was not targeted at a document.
Q. Well, but it was used in connection with providing documents of interest to users; right?
MS. BENNETT: Objection. Form.
THE WITNESS: It was used in connection with providing suggestions of hyperlinks that the user may want to follow in the user's navigation.
BY MR. PERLSON:
Q. Okay. And it used machine learning as part of that process; correct?
A. It used -- it used machine learning as part of that process. Right. The

models were not -- I actually used it on a trial basis. The models were not specific to a user. It was a generic model for users. And the -- what it did was, among the different links on the page, it would recommend you might want to follow this one or this one but not these two others. It was right about half the time.

Q. Okay. And whether it's right or not, I mean, there's no requirement in the patents that -- that in this case, that the results are correct or incorrect in determining whether -- users' interest; right?

MS. BENNETT: Objection. Form.

THE WITNESS: I don't know the legal definition, but usually for something to be effective, you would assume it would have to be better than a random chance.

BY MR. PERLSON:

Q. Well, but let's say -- let's say that I practice every single element of this claim in a system and then I go back and I talk to the -- let's say I've implemented this -- every element of one of the asserted independent claims in this case, and you've -- and delivered results to the user and I talked to the user

Q. I'm talking about the Pazzani one?

A. Naive Bayesian predates Pazzani and it's the -- as the name suggests, the simplest of all the Bayesian methods. There are more complex ones of which, for example, Dr. Jordan knows well. He created one of them. This is not the totality of the Bayesian. This is just one instance among the multiple Bayesian methods.

Q. Understood. Do the patents talk about using Bayesian?

MS. BENNETT: Objection. Form.

THE WITNESS: I have to go back and reread to see which methods, the patent suggests several methods -- they probably mention Bayesian methods as those were common at that time.

BY MR. PERLSON:

Q. And then it goes on to discuss the last -- if you look at the second to last line it says -- how do you pronounce that? Joachims?

A. Joachims.

Q. Joachims, that's the WebWatcher?

A. Yes.

Q. So Joachims goes on to say,

"Joachims introduced probabilistic TFIDF that takes into account document representation" -- what is that little symbol there?

A. That's a theta.

Q. "Theta and defines probability for class C for given doc that contains words W," and then it lists this long formula here?

A. Yes.

Q. What is that formula?

A. That is a probabilistic version of the basic TFIDF version, which is -- which will give you numbers in the zero to one range and will continue.

Q. Okay. And then it mentions the WebWatcher also used naive simple Bayesian classifier on frequency vectors, the same as we used in Personal WebWatcher. Do you see that?

A. Yes. The naive simple Bayesian classifier is more straightforward than this particular formula.

Q. Okay. And so -- but that's the learning algorithm that is being used in the Personal WebWatcher; is that right?

MS. BENNETT: Objection. Form.

THE WITNESS: I believe Personal

WebWatcher -- excuse me, that WebWatcher tried both of those algorithms and I believe that Personal WebWatcher tried naive Bayes and tried k-nearest neighbors, just letter K.

BY MR. PERLSON:

Q. Then if you look on the last sentence of that paragraph states, "the current version of PWW uses a naive simple Bayesian classifier for vectors to generate"?

A. Sorry. I lost the paragraph.

Q. Sure. See where it says structure of the Personal WebWatcher on page 7?

A. Yes. Yes. I see that.

Q. The sentence immediately before that.

A. I got it. I read it.

Q. It says, "the current version of Personal WebWatcher uses a naive simple Bayesian classifier on frequency vectors to generate a model of user interest that is used for advising hyperlinks." Do you see that?

A. That's right. That is not the full formula given above.

Q. Okay. But that's what it says right here?

A. Yes. That's what she used.

Q. Okay. That's something that is -- so in generating this model of user interest there would be estimating the parameters; correct?

MS. BENNETT: Objection to form.

THE WITNESS: It would be estimating the parameters of the naive Bayesian algorithm. Those parameters would be coefficients or weights on the various terms that would be included in the naive Bayesian formulation.

BY MR. PERLSON:

Q. And the Personal WebWatcher discloses using this naive Bayes machine learning to make probabilities; right?

MS. BENNETT: Objection. Form.

THE WITNESS: That's not quite right. First of all, just to clarify. Let's go back to that paragraph that you were quoting from. The beginning of it, short paragraph. "We decided to test different learning algorithms on PWW data since there it's not clear which algorithm as many appropriate." Later on in the paper they describe the K next neighbor as the other one they applied. So naive Bayes was one of the two -- still using naive Bayes to make a binary prediction. If you look at the top of page 11 it was trying to predict -- bottom of page 10 as well, it was trying to predict positive or negative.

BY MR. PERLSON:

Q. I'm sorry, where are you pointing to?

A. I'm pointing to the bottom of page 10 and the top of page 11.

Q. So where does it say it's a binary prediction?

A. The second paragraph in section 4.2 and the simple formula that follows it.

Q. Okay. I thought you had just directed me to -- maybe I'm looking at the wrong page.

A. Page 10 and 11 of the article, not the complicated number there.

Q. Okay. So in your view deciding whether or not the user's interested in the document as a binary prediction is not a probability?

MS. BENNETT: Objection. Form.

THE WITNESS: That's correct.

BY MR. PERLSON:

Q. In estimating parameters could a parameter be a binary parameter such as zero to one?

MS. BENNETT: Objection. Form.

THE WITNESS: We're confusing issues here. One of them is the output of a system and another one is internal to the model. Parameters are internal to the model. The prediction is the output of the model.

BY MR. PERLSON:

Q. Understood.

In building the model could a variable would be sports and could -- would estimating a parameter, could that be merely determining a binary decision of whether I'm interested in sports or not?

MS. BENNETT: Objection. Form.

BY MR. PERLSON:

Q. In assigning a one to yes I'm interested or zero to not interested?

MS. BENNETT: Objection. Form.

THE WITNESS: The parameters of a model can be continuous, can be discrete, and

EXHIBIT C

THIS EXHIBIT HAS BEEN REDACTED IN ITS ENTIRETY