IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

PERSONALIZED USER MODEL LLP	:	
	:	
Plaintiff,	:	
	:	
v.	:	Civ. No. 09-525-LPS
	:	
GOOGLE INC.	:	
	:	UNSEALED ON
	:	JANUARY 26, 2012
Defendant.	:	

OPINION

Karen Jacobs Louden, Jeremy A. Tigan, Wilmington, DE, MORRIS, NICHOLS, ARSHT & TUNNELL, LLP.

Marc S. Friedman, New York, NY; Mark C. Nelson, Dallas, TX; Jimmy M. Shin, Jennifer D. Bennett, Palo Alto, CA, SNR DENTON US LLP.

Attorneys for Plaintiff.

Richard L. Horwitz, David E. Moore, Wilmington, DE, POTTER ANDERSON & CORROON, LLP.

Charles K. Verhoeven, David A. Perlson, Antonio R. Sistos, San Francisco, CA, QUINN EMANUEL URQUHART & SULLIVAN, LLP.

Attorneys for Defendants.

January 25, 2012
Wilmington, DE

STARK, U.S. District Judge:

Personalized User Model, L.L.P. ("PUM") filed the instant lawsuit on July 16, 2009, alleging that Defendant Google, Inc. infringes two of PUM's patents: U.S. Patent No. 6,981,040[1] ("the '040 patent"); and U.S. Patent No. 7,685,276[2] ("the '276 patent"). (D.I. 1) The parties filed a Joint Claim Construction Chart ("JCCC") laying out the claim terms in dispute. (D.I. 115) Briefing was completed on December 13, 2010, and the Court held a *Markman* hearing on January 11, 2011. (*See* D.I. 170 (Hr'g Tr., January 11, 2010, hereinafter "Tr.")) At the Court's direction, the parties submitted post-hearing letter briefs. (D.I. 160; D.I. 161; D.I. 163) The Court now proceeds to provide constructions of the disputed terms.

I. BACKGROUND

A. Overview of the Technology

Because of the magnitude of material available on the internet, users need some type of filtration system to help sort through the overwhelming amount of irrelevant and useless information.[3] The technology involved in this lawsuit relates generally to this field of art: services and methods that provide a more personalized experience for internet users based on an individual user's interests. More specifically, the technology filters information available on the internet to target more relevant internet search results, advertisements, news, and other

---

[1]The '040 patent is found at D.I. 117 Ex. A.

[2]The '276 patent is found at D.I. 117 Ex. B.

[3]For example, as PUM explains, search results for the word "jaguar" would return information related to the animal jaguar, the car Jaguar, and the Jacksonville Jaguars football team. *See* PUM Technology Tutorial (D.I. 153). In PUM's words, the problem is related to "information overload," or the "proverbial needle in a haystack."

information for an individual internet user. (D.I. 119 at 2; D.I. 116 at 1-2)

These personalized computer network services function by creating a user profile based in part on a user's interactions with data on the internet. The technology monitors a user's interactions with data and creates a "user profile," which is based on the user's "interests." The technology is dynamic, in that it constantly updates the user profile based on new information obtained through further monitoring of the user's interactions with network information.

Prior art in the field was divided into two basic categories: information filtering techniques and collaborative filtering techniques. ('040 patent, col. 1 lines 29-30) Information filtering techniques represented a user as a profile, often characterized by a set of informative words or categories in a set of user-associated documents. These techniques would then parse new documents for content to determine how often the informative words or categories appeared in the document. (*Id.* at col. 1 lines 57-60) The existing information filtering techniques suffered from three primary drawbacks. First, since the filtering process proceeded in two steps – collecting documents followed by filtering the documents that were collected – the personalized filtration was only applicable to the second step of the process. (*Id.* at col. 1 line 61-65) The underlying assumption of the information filtering process is that each and every network document is presented with the user's personal filter, which is impractical given the extent of the amount of information on the internet. (*Id.* at col 1 line 67) Second, the existing information filtering techniques had difficulty handling dynamic documents – those documents that constantly change, such as news sites or sports sites. (*Id.* at col. 2 lines 3-5) A final problem was that user profiles were relatively limited. They were based only on sets of informative words and often based on only one mode of user-interaction. (*Id.* at col. 2 lines 6-10) As a result, updates

to the user profile, if they occurred at all, were minimal.

Collaborative filtering, on the other hand, is based on a database of an individual's user opinions that is then compared to databases of opinions of similar users. (*Id.* at col. 2 lines 16-17) In order to function properly and make predictions about the relevancy of a particular document to a particular user, collaborative filtering techniques must mine a very large set of offline data from a large number of users to generate associational rules and patterns. (*Id.* at col. 2 lines 20-24) The major drawbacks of existing collaborative filtering techniques were that they did not take into account the content of a document, and they could not "incorporate user context." (*Id.* at col. 3 lines 9-11) Additionally, since collaborative filtering techniques are based on user opinions, they do not have the capacity to incorporate documents that have "never been rated" or that are "rated by only a few users." (*Id.* at col. 3 lines11-15)

Several prior art references combined advantages of the information filtering and collaborative filtering techniques. For instance, U.S. Pat. No. 5,867,799, issued to Lang, incorporates facets of both methods. (*Id.* at col. 3 lines 24-26) However, according to the '040 patent, the Lang patent, including a later extension of the Lang patent (U.S. Pat. No. 5,983,214), does not "maintain an adaptive content-based model of a user that changes based on user behavior, allow for real-time updating of the model, operate during the collection stage of information retrieval . . . make recommendations for items or documents that have never been evaluated, or model a user based on different modes of interaction." (*Id.* at col. 3 lines 39-45)

B. **Patents-in-Suit**

The '040 patent, entitled "Automatic, Personalized Online Information And Product Services," was granted by the U.S. Patent and Trademark Office ("PTO") on December 27, 2005.

The '040 patent, which lists as inventors Yochai Konig, Roy Twersky, and Michael Berthold, contains a total of 62 claims. PUM asserts that Google infringes claims 1, 11, 21, 22, and 34. (D.I. 119 at 1) Of these asserted claims, only claim 1 is independent. Claims 11, 21, and 22 depend from claim 1, while claim 34 depends from independent claim 32, which is not asserted in the instant lawsuit. (*Id.*)

The '276 patent is a continuation patent; it shares the same title and inventors as the '040 patent and essentially the same specification. The key difference between the two patents is that the '276 patent concerns when the user is performing a "search query." ('276 patent, col. 31 line 53) PUM asserts the following claims of the '276 patent: claims 1, 3, 5-7, 14, and 21-24. Claims 1 and 23 are independent claims. Claim 24 depends from claim 23, while the remaining asserted dependent claims depend from claim 1. ('276 patent, col. 31-32)

C. **Disputed Terms**

While most of the disputed claims arise in both patents, the parties largely focus on the '040 patent. Examples of the disputed claim terms as they appear in the '040 patent, with emphasis added to show the disputed terms, appear below:[4]

[4]Several claims in the '040 patent are made up of different limitations that comprise the claim. When the claim's limitations are denominated by an alphabetic letter, the Court will refer to the specific limitations by both their claim number as well as their alphabetic signifier. For example, claim 1 of the '040 patent is comprised of six limitations, from "a" to "f." Hence, the Court will refer to limitation (f) of claim 1 as limitation 1(f).

1. A computer-implemented method for providing automatic, personalized information services to ***a user u***, the method comprising:

a) transparently monitoring user interactions with data while ***the user*** is engaged in normal use of a computer;

b) updating ***user-specific data files***, wherein ***the user-specific data files*** comprise the ***monitored user interactions with the data*** and a ***set of documents associated with the user***;

c) ***estimating parameters of a learning machine***, wherein the parameters define ***a User Model specific to the user*** and wherein ***the parameters*** are estimated in part from ***the user-specific data files***;

d) analyzing ***a document d*** to identify properties of ***the document***;

e) ***estimating a probability P(u/d) that an unseen document d is of interest to the user u***, wherein ***the probability P(u/d)*** is estimated by applying the identified properties of ***the document*** to ***the learning machine*** having ***the parameters*** defined by ***the User Model***; and

f) using ***the estimated probability*** to provide automatic, personalized information services to ***the user***.

11. The method of claim **1** further comprising ***estimating a posterior probability P(u/d,q) that the document d is of interest to the user u, given a query q submitted by the user***.

21. The method of claim **1** further comprising sending to a third party web server ***user interest information derived from the User Model***, whereby the third party web server may customize its interaction with ***the user***.

22. The method of claim **1** wherein the monitored ***user***

interactions include a sequence of interaction times.

32. A program storage device accessible by a central computer, tangibly embodying a program of instructions executable by the central computer to perform method steps for providing automatic, personalized information services to ***a user u***, the method steps comprising:

a) transparently monitoring user interactions with data while ***the user*** is engaged in normal use of a client computer in communication with the central computer;

b) updating ***user-specific data files***, wherein ***the user-specific data files*** comprise the ***monitored user interactions with the data*** and a ***set of documents associated with the user***;

c) ***estimating parameters of a learning machine***, wherein ***the parameters*** define ***a User Model specific to the user*** and wherein ***the parameters*** are estimated in part from ***the user-specific data files***;

d) analyzing ***a document d*** to identify properties of ***the document***;

e) ***estimating a probability P(u/d) that an unseen document d is of interest to the user u***, wherein ***the probability P(u/d)*** is estimated by applying the identified properties of ***the document*** to ***the learning machine*** having ***the parameters*** defined by ***the User Model***; and

f) using ***the estimated probability*** to provide automatic, personalized information services to ***the user***.

34. The program storage device of claim 32 wherein analyzing ***the document d*** provides for the analysis of documents having multiple distinct media types.

II. LEGAL STANDARDS

"It is a bedrock principle of patent law that the claims of a patent define the invention to which the patentee is entitled the right to exclude." *Phillips v. AWH Corp.*, 415 F.3d 1303, 1312 (Fed. Cir. 2005) (internal quotation marks omitted). Construing the claims of a patent presents a question of law. *See Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 977-78 (Fed. Cir. 1995), *aff'd*, 517 U.S. 370, 388-90 (1996). "[T]here is no magic formula or catechism for conducting claim construction." *Phillips*, 415 F.3d at 1324. Instead, the court is free to attach the appropriate weight to appropriate sources "in light of the statutes and policies that inform patent law." *Id.*

"[T]he words of a claim are generally given their ordinary and customary meaning . . . [which is] the meaning that the term would have to a person of ordinary skill in the art in question at the time of the invention, i.e., as of the effective filing date of the patent application." *Id.* at 1312-13 (internal citations and quotation marks omitted). "[T]he ordinary meaning of a claim term is its meaning to the ordinary artisan after reading the entire patent." *Id.* at 1321 (internal quotation marks omitted). The patent specification "is always highly relevant to the claim construction analysis. Usually, it is dispositive; it is the single best guide to the meaning of a disputed term." *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996).

While "the claims themselves provide substantial guidance as to the meaning of particular claim terms," the context of the surrounding words of the claim also must be considered. *Phillips*, 415 F.3d at 1314. Furthermore, "[o]ther claims of the patent in question, both asserted and unasserted, can also be valuable sources of enlightenment . . . [b]ecause claim terms are normally used consistently throughout the patent . . . ." *Id.* (internal citation omitted).

It is likewise true that "[d]ifferences among claims can also be a useful guide . . . . For example, the presence of a dependent claim that adds a particular limitation gives rise to a presumption that the limitation in question is not present in the independent claim." *Id.* at 1314-15 (internal citation omitted). This "presumption is especially strong when the limitation in dispute is the only meaningful difference between an independent and dependent claim, and one party is urging that the limitation in the dependent claim should be read into the independent claim." *SunRace Roots Enter. Co., Ltd. v. SRAM Corp.*, 336 F.3d 1298, 1303 (Fed. Cir. 2003).

It is also possible that "the specification may reveal a special definition given to a claim term by the patentee that differs from the meaning it would otherwise possess. In such cases, the inventor's lexicography governs." *Phillips*, 415 F.3d at 1316. It bears emphasis that "[e]ven when the specification describes only a single embodiment, the claims of the patent will not be read restrictively unless the patentee has demonstrated a clear intention to limit the claim scope using words or expressions of manifest exclusion or restriction." *Liebel-Flarsheim Co. v. Medrad, Inc.*, 358 F.3d 898, 906 (Fed. Cir. 2004) (internal quotation marks omitted), *aff'd*, 481 F.3d 1371 (Fed. Cir. 2007).

In addition to the specification, a court "should also consider the patent's prosecution history, if it is in evidence." *Markman*, 52 F.3d at 980. The prosecution history, which is "intrinsic evidence," "consists of the complete record of the proceedings before the PTO [Patent and Trademark Office] and includes the prior art cited during the examination of the patent." *Phillips*, 415 F.3d at 1317. "[T]he prosecution history can often inform the meaning of the claim language by demonstrating how the inventor understood the invention and whether the inventor limited the invention in the course of prosecution, making the claim scope narrower than it would

otherwise be." *Id.*

A court also may rely on "extrinsic evidence," which "consists of all evidence external to the patent and prosecution history, including expert and inventor testimony, dictionaries, and learned treatises." *Markman*, 52 F.3d at 980. For instance, technical dictionaries can assist the court in determining the meaning of a term to those of skill in the relevant art because such dictionaries "endeavor to collect the accepted meanings of terms used in various fields of science and technology." *Phillips*, 415 F.3d at 1318. In addition, expert testimony can be useful "to ensure that the court's understanding of the technical aspects of the patent is consistent with that of a person of ordinary skill in the art, or to establish that a particular term in the patent or the prior art has a particular meaning in the pertinent field." *Id.* Nonetheless, courts must not lose sight of the fact that "expert reports and testimony [are] generated at the time of and for the purpose of litigation and thus can suffer from bias that is not present in intrinsic evidence." *Id.* Overall, while extrinsic evidence "may be useful" to the court, it is "less reliable" than intrinsic evidence, and its consideration "is unlikely to result in a reliable interpretation of patent claim scope unless considered in the context of the intrinsic evidence." *Id.* at 1318-19.

Finally, "[t]he construction that stays true to the claim language and most naturally aligns with the patent's description of the invention will be, in the end, the correct construction." *Renishaw PLC v. Marposs Societa' per Azioni*, 158 F.3d 1243, 1250 (Fed. Cir. 1998). It follows that "a claim interpretation that would exclude the inventor's device is rarely the correct interpretation." *Osram GmbH v. Int'l Trade Comm'n*, 505 F.3d 1351, 1358 (Fed. Cir. 2007). Thus, if possible, claims should be construed to uphold validity. *See In re Yamamoto*, 740 F.2d 1569, 1571 (Fed. Cir. 1984).

III. CONSTRUCTION OF DISPUTED TERMS

The patents-in-suit, and the underlying technology, involve intricate and complicated issues of computer science and machine learning. Unsurprisingly, the disputed terms do not lend themselves to easy categorization. In the parties' initial briefing, PUM and Google presented the terms in dispute in substantially different orders. (*Compare* D.I. 119 *with* D.I. 116) Consequently, the Court found it difficult to identify the key issues in dispute. At the claim construction hearing, the Court directed the parties to meet and confer to provide the Court with a joint proposal as to how the Court could most effectively address the issues and organize its Opinion. Unfortunately, the parties' submissions provided neither clarification nor simplification.[5] The Court has decided to approach the disputed terms essentially in the order in which they arise in the patents.

A. '040 Patent

1. Preamble to Claim 1

Claim 1 of the '040 patent begins, "A computer-implemented method for providing automatic, personalized information services to a user u . . ." ('040 patent, col. 32 lines 24-25) "User u" or "user" is one of the terms in dispute.[6]

a. "user u"

The term "user" or "user u" is used throughout the claims and the specification. Indeed,

---

[5]The parties "agree" on seven broad categories within which to group the terms in dispute. (D.I. 160 at 6; D.I. 161 at 2-5) Beyond this, however, they cannot agree on other points, such as the order in which terms should be addressed within each of these categories.

[6]User is also one of several terms about which Google contends there is no antecedent basis. (D.I. 160 Ex. A at 7) The Court will address the antecedent basis issue later in this Opinion.

the very first sentence of the patent's specification provides, "This invention relates generally to methods for personalizing a user's interaction with information in a computer network." ('040 patent, col. 1 lines 13-14) At first glance, the parties' proposed constructions seem similar: Google proposes that the term be construed to mean "person operating a computer," while PUM proposes "person operating a computer as represented by a tag or identifier." The parties, thus, do not dispute that user at least covers the actual person operating the computer. The issue is whether the term is limited to the actual person.

Google argues that the entire patent operates on the presumption that a user is a person: "As a matter of common sense, personalization services are provided to persons." (D.I. 16 at 18) Particularly persuasive in Google's view are references in the specification to a user's gender. The written description, for example, provides that, "Products that are of high interest to the user are suggested to ***him or her*** for purchase." ('040 patent, col. 31 lines 58-59 (emphasis added); *see also id.* at col. 23 lines 55-57) ("If the user feels that the User Model is not an adequate representation of ***him or her***, the user may submit user modification requests.") (emphasis added) Google argues that PUM, in its construction, has "dropped in out of nowhere" the language that suggests that a user is "represented by a tag or identifier." (D.I. 116 at 18)

PUM, on the other hand, contends that Google is engaged in nonsensical wordplay. According to PUM, "[c]omputers identify users by their electronic tags or identifiers," such that "tag or identifier is a representation of the user." (D.I. 119 at 11) Google ignores "basic principles of computer science" by attempting to limit a user to the "actual physical person operating the computer." (*Id.*) PUM contends that the specification expressly defines user and the user u, as it states: "The following notation is used in describing the present invention. The

user ***and his or her associated representation*** are denoted with u . . . ." ('040 patent, col. 9 lines10-12) (emphasis added) PUM acknowledges that in lay parlance, user is understood as the physical person operating the computer, but it asserts that one with ordinary skill in the art of computer science and machine learning would understand that the term also includes that person's tag or identifier. (D.I. 119 at 11-12)

The claim language – and the parties' constructions – confirm that, at the least, user must include a person. For example, claim 5 states that the patented method "transparently monitor[s] user interactions" during "multiple distinct modes of user interaction with network data," including, *inter alia*, "network searching" mode, "email reading" mode, "email writing" mode, and "document writing" mode. ('040 patent, col 32 lines 55-64; *see also id.*, col. 4. lines 58-61) It is people, of course, who undertake these activities, including reading and writing emails.

It is also true that the technology in the patent does not employ facial recognition technology or fingerprinting to recognize a particular person that is operating the computer; also obvious is that the person operating the system is not actually inside the computer. (Tr. at 32) Hence, what the technology directly interacts with is not a person, but instead an electronic identifier that is associated with the person. The personalization services are provided directly to a machine, through which a person receives the services.

The specification supports this view. It does so, first, in noting – as PUM emphasizes – both a user ***and*** his or her associated representation. ('040 patent, col. 9 lines 10-12; D.I. 119 at 11) Even if not a definition, this language is strong evidence that the patent contemplates that user u would include the person's associated representation. Second, the patent's specification speaks of "clustering" users, which is not something done (in the context of the art) with actual

persons. Thus, when the specification indicates that, "Distances between users are calculated to determine similar users, who are clustered into clusters of similar users," the patent is not referring to physical distances. ('040 patent, col. 5 lines 27-28; *see also id.* at col. 9 lines 21-25; D.I. 136 at 16)

The claim language and the specification refer to the user u as being both the actual physical person and his or her representation. Accordingly, the Court construes "user" and "user u" to mean "a person operating a computer or the associated representation of the user."[7]

2. Limitation 1(b)

The next disputed term that arises in claim 1 of the '040 patent appears in limitation 1(b), which recites, "updating user-specific data files, wherein the user-specific data files comprise the monitored user interactions with the data and a set of documents associated with the user." ('040 patent, col. 32 lines 29-32) Two terms are in dispute: (i) user-specific data files; and (ii) monitored user interactions with data.[8] (D.I. 160)

a. "user-specific data files"[9]

The parties dispute whether the data files must be "files." (D.I. 131 at 11) Google contends that they must be. (D.I. 116 at 14) In PUM's view, by contrast, the specification and

---

[7]The Court has not adopted PUM's "tag or identifier" language, as this language – unlike "associated representation" – does not appear in the patent.

[8]The term "set of documents associated with the user" had also been in dispute, but the parties agreed after the hearing that "set" should be construed as "group or collection" and the phrase "set of documents associated with the user" need not be separately construed by the Court. (D.I. 160 at 1) The Court accepts the parties' joint proposal.

[9]The parties' dispute as to the meaning of the "user-specific" portion of this and other terms will be addressed in connection with the term "User Model specific to the user," which arises in limitation 1(c).

the claim teach that user-specific data files means the monitored user interactions with data and a set of documents[10] associated with the user. (D.I. 119 at 13) PUM suggests that the word files is not intended to be a limitation to any specific kind of structure. (D.I. 136 at 15)

PUM is correct that the claims and the specification repeatedly refer to the user-specific data files as somehow involving or relating to monitored user interactions and documents associated with the user. The claim itself notes that the monitored user interactions with data and the set of documents "comprise" the user-specific data files. ('040 patent, col. 32 lines 29-33) Generally, "comprise" is used in patent claims to mean "include." *See, e.g., Amgen Inc. v. Hoechst Marion Roussel, Inc.*, 314 F.3d 1313, 1344-45 (Fed. Cir. 2003) ("'Comprising' . . . means that the named elements are essential, but other elements may be added and still form a construct within the scope of the claim. The word 'include' means the same thing."). Consistent with PUM's reasoning, the specification at one point states that the user-specific data files "include a set of documents and products associated with the user as well as monitored user interactions with data." ('040 patent, col. 8 line 67-col 9 lines 1-2) Thus, the user-specific data files "include" the monitored user interactions and a set of documents.[11]

Google's proposed construction would arguably impose a specific kind of structural limitation onto the claim, which would be improper, as the claim language itself does not impose such a limitation, and the specification notes that "any suitable data structure may be used."

---

[10]The parties' dispute as to the meaning of "document" will be addressed in connection with limitation 1(d).

[11]To the extent that PUM seeks to limit data files only to monitored user interactions with data and a set of documents associated with the user, the Court rejects that notion. As the case law provides, and as Google discussed at the hearing, comprising is "open-ended." *See Genentech, Inc. v. Chiron Corp.*, 112 F.3d 495, 501 (Fed. Cir. 1997).

('040 patent, col. 22 lines 32-33)

The Court, thus, agrees with PUM. The Court will construe "user-specific data files" to mean "the monitored user interactions with data and a set of documents associated with the user."

b. "monitored user interactions with the data"

The dispute about this term is not entirely clear. (*See, e.g.*, Tr. at 37) (PUM's counsel: "We don't quite understand what defendant is going to contend [monitored user interactions with data] are."). Google argues that the monitored user interactions with data must be "obtained from the monitoring step of 1(a)." (D.I. 131 at 12) Google contends that PUM's proposed construction introduces "the collected information" phrase with no antecedent basis and appears designed to avoid the clear reference back to the previous step. (*Id.*) PUM responds that the specification contemplates that the monitored user interactions with data are "the collected information about the user's interactions with data." (D.I. 119 at 12) In PUM's view, Google's construction simply reorders the words rather than provides a definition that would be helpful to the jury. (*Id.* at 14; *see also generally Am. Patent Dev. Corp v. Movielink, LLC*, 604 F. Supp. 2d 704, 716 (D. Del. 2009) (rejecting construction that is "merely a verbose paraphrasing of claim language that otherwise offers little to assist one of skill in the art in understanding the claims"))

The specification makes clear that the User Model is updated by monitoring the user's interactions with data. ('040 patent, col. 8 lines 60-66) The User Model is a "dynamic entity," that is "constantly and dynamically updated," where the updating happens in "real-time" based on "all user actions." ('040 patent, col. 21 line 64-65; *id.* at col. 3 lines 39-41) The monitored user interactions come from "[m]ultiple distinct modes of interaction," including, "network

searching, network navigation, network browsing, email reading, email writing, document writing, viewing pushed information, finding expert advice, product information searching, and product purchasing." ('040 patent, col. 22 lines 1-4) The user interactions include positive and negative patterns, such as a user following a search query, saving a document in favorites or bookmarks, ignoring search results or pushed news, etc. (*Id.* at col. 22 lines 15-25) Monitored user interactions with data ultimately result in "collected information" that is used to refine the parameters of the User Model. ('040 patent, col. 21 lines 63-64) PUM's construction more accurately captures the process involved in monitoring the user's interactions and aligns with how the monitored user interactions are used to update the User Model.

Therefore, the Court will construe "monitored user interactions with the data" to mean "the collected information about the user's interactions with data."

3. Limitation 1(c)

Limitation 1(c) recites "estimating parameters of a learning machine, wherein the parameters define a User Model specific to the user and wherein the parameters are estimated in part from the user-specific data files." ('040 patent, col. 32 lines 33-36) Both parties agree that the "main and overarching disputes" relate to the construction of the "user-specific" or "specific to the user" terms contained in this limitation. (D.I. 131 at 1; D.I. 136 at 2) These disputes, in fact, take up approximately one-third of the parties' briefing and consumed the bulk of the time at the hearing. The parties agree that, at the least, the following terms need to be construed by the Court: (i) estimating parameters; (ii) learning machine; (iii) User Model specific to the user;

and (iv) user-specific learning machine.[12] (D.I. 160; D.I. 161)

a. "estimating parameters"[13]

This term presents two disputes: (1) what are the parameters themselves; and (2) must every parameter be used to estimate the probability of a user's interest in a document? (D.I. 116 at 13-14; D.I. 119 at 16-17)

(i) "parameters"

Both parties use "values and weights" and "variables" in their proposed definitions of parameters. (D.I. 119 at 16) Google characterizes the dispute as whether the "values and weights" are the parameters or, in the alternative, whether the "variables," which are assigned values and weights, are the parameters. (D.I. 116 at 13-14) In Google's view, the parameters of a learning machine are the variables themselves, and the "estimating" limitation of this term involves estimating a value or weight for each of the parameters. (D.I. 131 at 10) Importantly for Google, it is the variables themselves that are used to estimate probabilities. (D.I. 116 at 13-14)

PUM contends, however, that the parameters are values or weights, as opposed to variables that are given a weight or value. (D.I. 136 at 3) PUM argues that in order for

---

[12]While the parties agreed that the Court should have a "Learning Machine Terms" grouping, PUM lists five disputes under this category, while Google lists only four. Nor could the parties agree on the order in which the Court should address these terms. (D.I. 160; D.I. 161; D.I. 163)

[13]The term "estimating" is also a disputed term in a different context. That dispute involves whether "estimating" should mean "calculating" or "approximating or roughly calculating" and arises in connection with the "estimating a probability" language found in limitation 1(e). (D.I. 116 at 16-17; D.I. 119 at 23-25) The Court defers discussion of that dispute until later in this Opinion.

parameters to be "tunable," or to be "estimated," the parameters must be capable of being numerically approximated or roughly calculated; the parameters, then, must encompass a mathematical component. (*Id.*) PUM further argues that the initialization mode of the invention proves that the parameters are values or weights. According to the specification, during the initialization phase of the user model, documents are analyzed to determine the initial parameters. ('040 patent, col. 17 lines 48-54) Later, the parameters are updated based on the user's interactions with data. (*Id.* at col. 21 lines 63 to col. 22 line 7) In PUM's view, Google's definition would require changing the parameters to something else, "modifying the actual variables in some way or creating new ones," which would fail to capture the tunability of the parameters. (D.I. 163 at 5) PUM fears that Google's construction might result in the parameters being "words and things like that." (Tr. at 117)

As PUM suggests, the claim language does indicate a mathematical element to the parameters. The parameters are estimated, according to limitation 1(c). ('040 patent, col. 32 line 33) It is difficult to understand how changing the variables themselves would be consistent with estimating parameters. The specification also supports the conclusion that parameters must be numbers. Parameters are "stored," "tuned," "estimated," and "continually updated." ('040 patent, col. 8 lines 45-50)

In the Court's view, Google's construction would result in the parameters being the variables themselves, despite Google's protestations to the contrary. Under Google's construction, as PUM points out, "each user would have a separate model made up of hundreds of thousands of words and all these other things." (Tr. at 117) The Court will adopt PUM's proposal and construe "parameters" to mean "values or weights."

(ii) "estimating parameters of a learning machine"

Google's proposed construction of this term is "estimating a value or weight of each of the variables that are used by the [user-specific] learning machine ***to calculate a probability***." (D.I. 131 at 9) (emphasis added) PUM's proposal is essentially the same except it omits the final "to calculate a probability" phase. The issue, then, is whether to include this phrase.

According to PUM, there is nothing about the claim term "estimating parameters of a learning machine" that involves calculating probabilities. (D.I. 119 at 16) Google argues that PUM cannot cite a single example of a parameter that would not be used to estimate probabilities; nor does PUM indicate what purpose a parameter would have other than estimating a probability. (D.I. 131 at 9)

The Court agrees with PUM that "to calculate a probability" is not a component of this term. Google's proposed construction conflates what parameters are with how they are used. Nor does limitation 1(c) require that every parameter be used to estimate a probability.

The Court will construe the term "estimating parameters of a learning machine" to mean "estimating values or weights of the variables of a learning machine."

b. "learning machine"

The term learning machine appears in claims 1 and 32 of the '040 patent and claims 1, 5, and 23 of the '276 patent. (D.I. 115 at 5; D.I. 116 at 5; D.I. 119 at 18) Google argues that a learning machine should be construed as a "program that contains parameters used to calculate a probability, and where the predictive ability of the program improves over time with the addition of new data." (D.I. 116 at 5) PUM argues that learning machine should be construed as "a model and/or mathematical function that is used to make a prediction or intelligent decision that

attempts to improve performance in part by altering the values/weights given to its variables depending upon past observations or experiences." (D.I. 116 at 5) The parties agree that a learning machine makes a prediction and that the predictive ability should theoretically improve over time.[14] (D.I. 119 at 18) The disputes are: (1) whether the learning machine must be a program or whether it is more properly understood as a mathematical function or model; and (2) whether the learning machine operates on "new data" or "past observations or experiences."[15] (D.I. 119 at 18; D.I. 116 at 5-8; D.I. 131 at 6-8; D.I. 136 at 9-11)

(i) mathematical function dispute

In Google's view, the learning machine is not a mathematical function or a model; instead, a learning machine is appropriately labeled as a program. (D.I. 116 at 8) Google contends that the language "mathematical function" or "model" used in PUM's proposed construction more properly describes the "user model," which is a separate term in the '040 patent. (D.I. 116 at 8) Google argues that there are "three limitations" in the claims: a learning machine, a user model, and parameters. (D.I. 138 at 9) PUM is improperly attempting to conflate the user model with the learning machine. (D.I. 116 at 8) Google also contends that the section of the specification upon which PUM relies to suggest that a learning machine may be a model demonstrates the opposite: "That the 'User Model' may be an 'implementation' of a

---

[14]Google initially proposed a definition of this term that implied that the predictive ability of the learning machine ***necessarily*** improved ***each*** time. At the hearing, however, Google conceded that the learning machine "attempted" to improve and agreed to a modification of its construction to reflect this concession. (Tr. at 75; *id.* at 127; *see also* D.I. 160 at 3)

[15]In the briefing, the parties disputed whether the learning machine made an "intelligent decision." (D.I. 131 at 6) PUM, however, later agreed to eliminate this language in its proposed construction. (Tr. at 20)

learning machine does not mean that a model is the ***same thing*** as the learning machine. The claims plainly recite 'User ***Model*** and 'learning machine' as separate limitations . . . ." (D.I. 131 at 9; *see also* Tr. at 77) (emphasis in original)

PUM argues that learning machines are not limited to "programs." (D.I. 119 at 19) PUM notes that the specification contains many references to learning machines that are not necessarily embodied in a program. (D.I. 119 at 18) PUM also points to technical literature that purportedly demonstrates that a learning machine may also be a model or mathematical function. (*Id.*)

The Court agrees with PUM that learning machine as used in the patents-in-suit is more accurately described as a model or a mathematical function. As the specification notes, "any model that is a learning machine" is included in the scope of the present invention. ('040 patent, col. 8 lines 52-53) That phrase is most sensibly understood to mean a learning machine can be categorized as a "model." Also persuasive is the fact that the User Model is specifically described ***as*** a learning machine. (040 patent, col. 8 lines 43-45) Google's argument that the User Model is not the same thing as the learning machine is inapposite here. PUM does not argue that the two terms are the same item, nor could it. (D.I. 118 Ex. C at PUM 0067743)

Furthermore, Personal Web, the preferred embodiment of the invention, defines and updates the User Model. ('040 patent, col. 8 line 63) The process, as the patent makes clear, is continuous and recursive. ('040 patent, col. 4 lines 63-64; *see also* D.I. 131 at 8) In that case, the User Model is specifically described as a "function" based on an "underlying mathematical framework" that includes "algorithms." ('040 patent, col. 8 lines 33-37) If the User Model is a kind of learning machine, which the specification indicates is the case, and if the User Model is a function based on an underlying mathematical framework, then the learning machine can

reasonably be said to be a mathematical function. Program, on the other hand, never appears in the specification with respect to the learning machine.

The Court declines to accept Google's "program" limitation for the term learning machine. The Court will include in its construction of "learning machine" that it be a "mathematical function and/or model."

(ii) "observations and experience" dispute[16]

The parties dispute how to characterize the information upon which the learning machine acts. (D.I. 131 at 8) PUM takes issue with Google's use of the term "new data," insisting instead that the learning machine improves based on "past observations or experience." (D.I. 119 at 20) Subsequently, Google agreed to modify its definition by removing "new data" and replacing it with "monitored user interactions." (D.I. 131 at 8) In PUM's view, however, Google's concession still leaves out other potential sources of information, such as a set of documents associated with the user and "world knowledge." (Tr. at 24) Google, for its part, sees a danger that PUM's use of the word "past" might allow a "one-time input" rather than an ongoing process. (D.I. 116 at 8) Both parties criticize the other's definition as being too vague or confusing. (D.I. 136 at 10; D.I. 116 at 7)

In the Court's view, neither party's proposal is satisfactory. Given the danger of

---

[16]The parties also disagree about how to phrase the tasks that the learning machine performs. Like it did previously with the "estimating" term discussed above, Google attempts to import a "to calculate a probability" limitation into this term. (D.I. 136 at 9) For the same reasons the Court articulated previously, the Court does not adopt Google's limitation. The specification repeatedly emphasizes that the invention estimates probabilities. Indeed, Google's own briefing is instructive on this point, as Google's discussion of this issue highlights how often the word "estimates" is used and how infrequently "calculates" is used in the patent. (D.I. 116 at 6; *see also* D.I. 131 at 6 (Google brief with heading, "A Learning Machine Estimates Probabilities"))

disguising the ongoing nature of the learning process that PUM's phrase ("past observations and experience") entails, the Court declines to adopt PUM's proposal. The patent's specification consistently notes that "user-specific data" and "monitored user interactions" are used to update the parameters of the learning machine, not user-specific "past observations and experience." ('040 patent, col. 4 lines 23-29) Google's construction, even as amended, does not, however, include a "set of documents associated with the user," which the claims make clear are at least included in the user-specific data files that are used to estimate the parameters of the learning machine. ('040 patent, col. 32 lines 29-36; *id.* at col 4 lines 25-30)

The specification provides, "Based on the monitored interactions, parameters of the learning machine are updated." (*Id.*, col. 4 lines 63-64) The specification describes different modes in which the user functions, during which the learning machine operates to update the parameters. ('040 patent, col. 4 lines 57-64; *id.* at col. 21 line 64 to col. 22 line 27) The modes include "network searching mode," "email reading mode," "viewing 'pushed' information mode," "product purchasing mode," etc. (*Id.* at col. 4 lines 57-63) "Monitored user interactions," as offered by Google, could be viewed as excluding these various modes.

Additionally, the parameters may be updated based on other users in the same cluster as the User Model. For example, the specification provides that, "[d]istances between users are calculated to determine similar users, who are clustered into clusters of similar users. Parameters defining the User Model may include the calculated distances between the User Model and User Models of users within the user's cluster." ('040 patent, col. 5 lines 27-30; *see also id.*, col 5 lines 6-7 (discussing "cluster probability distribution")) Moreover, the parameters may be updated based on world knowledge, such as the number of users who have accessed a particular

document or saved a particular document in a favorites' list. (Tr. at 40)

In the Detailed Description, the patent notes that the User Model is "developed and updated using a variety of knowledge sources." ('040 patent, col. 8 lines 33-34) The "variety of knowledge sources" language captures the "world knowledge" concept, as well as information based on user behavior in a cluster.

The Court, then, accepts, for the most part, PUM's proposed construction of "learning machine," making some limited adjustments. Hence, a "learning machine" will be construed as "a mathematical function and/or model used to make a prediction, that attempts to improve its predictive ability over time by altering the values/weights given to its variables, depending on a variety of knowledge sources, including monitored user interactions with data and a set of documents associated with the user."

c. "User Model specific to the user"

The disputed term appears in the '040 patent in claims 1, 21, and 32.[17] PUM proposes that User Model specific to the user means "an implementation of a learning machine updated in part from data specific to the user." (D.I. 119 at 18) Google contends that the phrase should, instead, be construed to mean "model unique to the user, that is created and updated by the learning machine and stored in a data structure." (D.I. 116 at 8) The crux of the dispute is what the patent means when it says "specific" or "specific to the user." (Tr. at 61) To resolve this dispute, the Court must determine (1) whether the User Models must be "unique," as Google

[17]The parties also dispute the meaning of "user-specific learning machine," which arises in the '276 patent. The Court will address that term later in this Opinion.

suggests;[18] and (2) whether the User Model must be stored in a data structure.

(i) "unique" dispute

Google faults PUM's proposal for suggesting that the User Model is not unique to each user, which in Google's view is contrary to the entire patent. (D.I. 116 at 9) Under Google's understanding of the patent, each individual user has a different User Model that is unique to that individual user. Google maintains that PUM's construction merely copies verbatim the "specific to the user" claim language, resulting in only the "data" by which the learning machine is updated being "specific to the user." (D.I. 116 at 10) Google relies, in part, on the '040 patent inventor's testimony that purportedly demonstrates that each user has a unique user model. (D.I. 131 at 3-4)

PUM responds that, as used in the asserted claims, "specific" means "related to or associated with," not "unique to" the user. (D.I. 119 at 22) In support of this proposition, PUM relies on the "commonly understood" meaning of specific. PUM cites to *Webster's Third New International Dictionary*, which lists as one of the definitions, "restricted by nature to a particular individual, situation, relation, or effect." (*Id.* at 21) PUM argues that Google's construction would mean "that each user have his or her own personal learning machine/user model potentially resulting in millions and millions of learning machines/models with tens of millions of variables." (D.I. 136 at 2) In PUM's view, the specificity of the User Model comes from the parameters that are updated and tuned based on data specific to the user – including, for example, user interactions with data. (D.I. 136 at 13)

---

[18]After the hearing, Google apprised the Court that substituting the word "restricted" in place of the word "unique" would be an acceptable alternative. (D.I. 160 at 2) PUM responds that the word "restricted" results in the same problems as "unique." The Court agrees with PUM.

Resolution of this dispute turns on whether a User Model is specific because it has completely different ***variables*** than other User Models, or if, instead, a User Model is specific because it has completely different ***numerical values*** than other User Models. Hence, the Court agrees with PUM that the fundamental issue is the construction of parameters, which the Court addressed earlier.

Contrary to the parties' briefing and argumentation, the claims do not provide solid support for either side. The claims merely provide that the User Model must be specific to the user, but they do not clarify the meaning of specific.

The specification is more helpful. The specification notes that, "In some cases, initialization is performed without any user-specific information. . . . For such users, prototype users are supplied." ('040 patent, col. 20 lines 28-34) Prototype users, and the fact that a User Model can be initialized without any user-specific information, undercuts Google's argument. Elsewhere, the patent notes that these prototype users are called "hats," and that users can choose to try on different "hats" in certain circumstances. (*Id.*, col. 20 lines 40-43) Google's attempt to harmonize this embodiment of the patent with its proposed construction was unpersuasive. (*See* Tr. at 70) ("This is not a user model that is specific to the user. It's a general user model that might be used by a user but the language that they cite shows that it's not a user model specific to the user.")

Google argued at the hearing that when you have multiple users sharing a model, "[i]t is a group model, and that is absolutely not a specific model." (Tr. at 70) But the patent makes clear what is meant by group or cluster model: "Practically, the cluster model is computed from the User Models by averaging the different distributions of the individual User Models." ('040

patent, col. 25 lines 41-43) Averaging all of the values of the individual user models requires that the user models all have the same underlying variables, which corroborates PUM's contention that "a situation where you had a model that had 100 generic (a) (b) plus (a) (b) (c) (d) (e) times (x)" is a "template for a function." (Tr. at 119) Averaging would not seem possible under Google's construction because all of the User Models would have functions with different variables.

The Court concludes that Google's use of the word "unique" is inconsistent with the patent's specification.[19] According to PUM, "Now, in reality in our patent most of the models would probably be unique because the parameters define the user model and the user-specific learning machine would be likely quite different for each individual. They just don't have to be." (Tr. at 28) The Court agrees.

(ii) "data structure" dispute

Google also attempts to impose a limitation that the User Model specific to the user be "stored in a data structure." (D.I. 116 at 11) Google's argument, however, is circuitous and unpersuasive. Nowhere in the patent is there a requirement that the User Model be stored in a data structure. The specification seems to teach the opposite: "[t]he User Model 13 is a function that is developed and updated using a variety of knowledge sources and that is independent of a specific representation or data structure." (*Id.* at col. 8 lines 28-35) Google's reliance on the following quote is, therefore, unavailing: "the User Model is a function that may be implemented with any data structure." (D.I. 116 at 11) This statement refers not to where the User Model is

[19]Both parties rely on dictionary definitions of "specific" to suggest that their definition is the "plain and ordinary" meaning of the word. (D.I. 131 at 2; D.I. 119 at 21) Having reviewed the dictionary definitions, the Court accords this extrinsic evidence little weight.

stored or that the User Model must be stored; rather, the statement suggests that the User Model can predict a user's response to any data structure, such as an internet website, an advertisement, a product, or any other "data structure." The Court agrees with PUM that the patentee did not clearly and unambiguously disclaim or disavow the scope of the claim. *See Phillips*, 415 F.3d at 1319-20; *see also Liebel-Farsheim*, 358 F.3d at 904. There is no requirement in the patent that the User Model be stored in a data structure.

(iii) Conclusion on "User Model specific to the user"[20]

For the reasons given, the Court will construe "User Model specific to the user" to mean "an implementation of a learning machine updated in part by data specific to the user."

d. "user-specific learning machine"

The parties also dispute a related term found in the '276 patent: "user-specific learning machine." (D.I. 160 at 6) The disputes here are the same as those already addressed and the Court resolves them in the same manner. Thus, the Court will construe "user-specific learning machine" in the '276 patent as "a learning machine [as construed] specific to the user."

4. Limitation 1(d)

Limitation 1(d) provides as follows: "analyzing a document d to identify properties of the

---

[20]Relying on limitation c of claim 1, which provides that the "parameters of the learning machine define a User Model specific to the user," Google argues that "the User Model is created and updated by the learning machine." ('040 patent, col. 32 lines 33-34) Google disagrees with PUM's contention that the User Model is an "implementation of the learning machine." (D.I 116 at 11) Claim 1 goes on to provide, however, that "the learning machine [has] the parameters defined by the User Model." (*Id.* col. 32 lines 42-43) Thus, there is some ambiguity as to whether the User Model is an implementation of the learning machine's parameters or is instead created and updated by the learning machine. What is clear is that the process is dynamic and capable of occurring in "real time." Google's definition merely repeats that the User Model is a model, while PUM's construction provides more guidance. The Court will therefore use PUM's "implementation of a learning machine."

document." ('040 patent, col. 32 lines 37-38) The parties provide competing constructions of the term "document." (D.I. 116 at 19-20; D.I. 119 at 15-16)

a. "document"[21]

This term appears in numerous claims throughout both patents. ('040 patent, claims 1, 11, and 34; '276 patent, claims 1, 5-7, 14, 21-24) The term also appears throughout the written description. Google argues that document should be construed as an "electronic file," while PUM contends that it should mean "text or any type of media." (D.I. 119 at 15; D.I. 116 at 19; *see also* D.I. 160) Google eventually agreed to modify its construction to provide explicitly that documents could include text or any type of media. (D.I. 160 at 3) Likewise, PUM agreed that documents must be electronic. (Tr. at 37) The remaining dispute, then, is whether a document must be a file.

Google asserts that PUM's construction does not make sense given that the patent describes several different actions that are taken in reference to a document. For example, Google notes that the claims themselves refer to "linking" to a document and "identifying properties of a document." (D.I. 116 at 19) Similarly, the claims teach "crawling network" documents, as well as documents "having multiple distinct media types." (*Id.*) In Google's view, all of this only makes sense if documents are electronic files. Google also contends that PUM's construction would allow a single word to be considered a document, but such a construction would produce the absurd result that a webpage would be a "set of documents," "500 documents," and also a "single document." (D.I. 116 at 20) Nor could all of the actions

[21]The term "document" is also included among the terms that Google argues lack an antecedent basis. (D.I. 115) The Court addresses this general dispute toward the end of this Opinion.

taken in relation to a document be taken with respect to a single word. (*Id.* at 21)

PUM responds that Google's definition is "fatally flawed" because it does not define what a document is, but instead focuses on how a document is stored. (D.I. 136 at 16) PUM also contends that the specification expressly defines the document, stating "[t]he term 'document' includes not just ***text***, but ***any type of media***, including, but not limited to, hypertext, database, spreadsheet, image, sound, and video." ('040 patent, col. 9 lines 14-17) (emphasis in original) PUM contends that Google's construction improperly imports a non-existent storage limitation into the term.

The claims themselves provide little guidance about the meaning of the term document. Limitation 1(d) states that the patented method analyzes documents to identify properties of the document, and, eventually, those properties are used to estimate a probability that an unseen document would be of interest to a user. ('040 patent, col. 32 lines 37-42) Claim 2 refers to documents "of interest" and documents "not of interest." (*Id.* at col. 32 lines 46-47) Claim 3 explains that the documents may have "multiple distinct media types." Claim 7 states that the patent system identifies properties of a document, including an "author," "age," or "language." (*Id.* at col. 33 lines 4-15) Thus, the claim language does not answer definitively whether a document must be an electronic file or if a document may instead be simply text.

The specification is similarly inconclusive. As PUM emphasizes, the specification states that "[t]he term 'document' includes not just text, but any type of media, including, but not limited to, hypertext, database, spreadsheet, image, sound, and video." ('040 patent, col. 9 lines 14-17) While instructive, this language is not dispositive.

The Court concludes that Google's construction is correct. Words and phrases – in other

words, "text" – are not documents; instead, documents ***contain*** text and other kinds of data. The invention describes "documents of interest" and "documents not of interest" ('040 patent, col. 4 lines 29-30), which suggests a document is more than a single word, since it is unlikely the invention is trying to predict a user's interest in a single word. The invention also maintains a "word database" that contains "statistics of each word or phrase from all user documents. The word database contains the word ID, full word, and word frequency in all documents D." ('040 patent, col. 17 lines 6-10) Elsewhere, the specification notes that the invention maintains an "informative word and phrase list" that "contains the most informative words and phrases found in user documents." (*Id.* col 10 lines 10) These concepts also seem to contemplate that a document must be more than a single word or phrase.

PUM's main objection to Google's construction is that it purportedly would read out dynamically generated text and search results. Google disagrees with PUM's characterization: "And I'm not sure why they think that our construction would preclude a dynamically created page. We don't think it does, and they didn't explain how it does, and I don't think it does." (Tr. at 96) The Court is not persuaded by PUM on this point.

Therefore, the Court will adopt Google's amended construction. The term "document" will be construed to mean "an electronic file including text or any type of media."

5. Limitation 1(e)

Limitation 1(e) recites, "estimating a probability P(u/d) that an unseen document d is of interest to the user u, wherein the probability P(u/d) is estimated by applying the identified properties of the document to the learning machine having the parameters defined by the User Model." ('040 patent, col. 32 lines 39-43) This term raises three disputes: (1) estimating;

(2) probability; and (3) unseen document.[22] The Court will address each issue in turn.

a. "estimating"

The parties ask the Court to construe the term "estimating." (D.I. 136 at 8; D.I. 131 at 13) Google proposes that estimating should be construed as "calculating." (D.I. 116 at 16) Google points to several instances in which the specification appears to use the two terms interchangeably. ('040 patent, col. 5 lines 39-42; *id.* at col. 5 lines 49-52) Google argues that the patent's use of the word estimating is in the context of statistical calculations using mathematical formulas, not "ballpark guesses," as PUM's construction suggests. (D.I. 116 at 17)

PUM maintains that estimating should be construed as "approximating or roughly calculating." (D.I. 119 at 24) PUM argues that the specification teaches that "estimating" is intended to be understood in the broad context of whether a user would be interested in a document. (*Id.*) PUM also contends that "generally" the specification refers to "estimating," not calculating, as Google argues. (D.I. 136 at 18) In support of its construction, PUM relies heavily on dictionary definitions. (D.I. 119 at 24-25) PUM cites, for example, the *Oxford English Reference Dictionary*, the *Microsoft Press Computer Dictionary*, and the *Random House Webster's Unabridged Dictionary*.

Looking at the claim language, it is noteworthy that the claims use estimating in two separate limitations: estimating parameters and estimating a probability. ('040 patent, col. 32 lines 33-39) Normally, claim terms are used consistently throughout a patent. *See Phillips*, 415 F.3d at 1314. Thus, the Court's construction should apply equally when the patent speaks of

---

[22]The term "of interest" also implicates Google's indefiniteness argument, which the Court addresses elsewhere in this Opinion.

estimating parameters and estimating probabilities.

According to the specification, one of the objects of the invention is to "provide a method based on Bayesian statistics." ('040 patent, col. 4 lined 8-11; *see also id.* col. 8 lines 35-39) The Court is not persuaded that PUM's construction is inconsistent with Bayesian statistics. Further, the extrinsic evidence supplied by PUM supports PUM's contention that "estimating" was generally understood by one of ordinary skill in the art at the relevant time as a measurement that is not entirely precise.

Accordingly, the Court will construe "estimating" to mean "approximating or roughly calculating."

b. "probability"

Google contends that "probability" in the patent-in-suit means "percentage chance." (D.I. 116 at 16) For support, Google turns to the *American Heritage College Dictionary*, which describes probability in the context of statistics as a ratio. Google argues that PUM's "belief or likelihood" construction is imprecise and unscientific. Google also relies on the deposition testimony of Konig, one of the patent's inventors, which confirms, in Google's view, that the patent operates in a mathematical context. (D.I. 131 at 13) According to Google, PUM relies on overly expansive definitions of probability. (*Id.* at 14)

PUM construes probability as "numerical degree of belief or likelihood."[23] (D.I. 119 at 24) PUM points out that nowhere does the patent require that the probability be expressed as a percentage chance. PUM relies on various dictionaries that define probability as, for example,

[23]PUM agreed at the hearing that the probability must be expressed in numerical terms. (Tr. at 120-21; *see also* D.I. 161 at 4)

"the likelihood of something happening" or "[t]he likelihood that an event will happen." *The Oxford English Reference Dictionary* 1152 (2d ed. 1996) (D.I. 120 Ex. 11); *Microsoft Press Computer Dictionary* 382 (3d ed. 1997) (D.I. 120 Ex. 7); *see also Free Motion Fitness, Inc. v. Cybex Intern., Inc.*, 423 F.3d 1343, 1348-49 (Fed. Cir. 2005) ("[I]n those circumstances where reference to dictionaries is appropriate, the task is to scrutinize the intrinsic evidence in order to determine the most appropriate definition.").

In light of PUM's agreement to include the requirement that the probability be expressed in numerical format, it is unclear whether the parties still have a material dispute with respect to this term. (*See* Tr. at 85) (Google's counsel: "And, here again, there is kind of semantics here.") To the extent there is a dispute, the Court agrees with PUM's construction. Accordingly, the Court will construe "probability" to mean "numerical degree of belief or likelihood."

c. "unseen document"

Limitation 1(e) of the '040 patent provides that the invention estimates a probability that an "unseen document d" is of interest to the user. ('040 patent, col. 32 lines 39-40) The sole dispute is whether the unseen document refers to a document that is unseen by ***the*** (specific) user or if, on the other hand, the claim term refers to a document that has never been seen by ***any*** user. (D.I. 131 at 17; D.I. 136 at 19) The Court agrees with PUM that an unseen document is one that is unseen by the specific user.

PUM argues that the term unseen document is only relevant in the context of a specific user. The claims do not mention unseen documents in relation to "all users." (D.I. 119 at 25) PUM also finds support for its construction in the specification, which consistently refers to the user in the singular rather than the plural. (D.I. 119 at 26) The preferred embodiment, called

"Personal Web," applies the "User Model 13 to unseen documents . . . to determine the user's interest in the document;" importantly, in PUM's view, user is singular here. Likewise, the section of the patent entitled "Applying the User Model to Unseen Documents" also employs the singular. (*Id.*)

PUM also points to claim 7 of the '040 patent. According to PUM, claim 7 refers to identified properties of documents, including the number of times the document has been accessed by users, the number of users who have listed the document on their favorites list, and the number of users previously interested in a document. (D.I. 119 at 25-26) PUM indicates that these identified properties demonstrate that "such an unseen document to the user ***may*** have been accessed and/or saved by previous users;" if some of the unseen documents may have been accessed by previous users, then unseen document only makes sense if it means that the unseen document is unseen by the specific user. (D.I. 119 at 26)

Google contends that the correct construction is a "document not previously seen by any user." (D.I. 116 at 21) For its position, Google relies on the way that PUM distinguished the current invention from the prior art during the patent's prosecution. (D.I. 138 at 18) The specification notes that one of the drawbacks of prior art collaborative filtering methods was that "an item that has never been rated cannot be recommended or evaluated." ('040 patent, col. 3 lines 9-12) PUM purportedly relied on this distinction during the prosecution of the patent to distinguish its patent from the prior art. (D.I. 116 at 22) Thus, according to Google, "[i]f unseen document is construed as PUM suggests . . . then the '040 patent would not improve upon the prior art in the manner the applicants asserted, both in the specification and during prosecution." (D.I. 116 at 23)

As an initial matter, a document that is unseen by any user would by definition include a document that was unseen by the user. The parties thus agree that the term unseen document at least includes documents that have not been seen by the user. The question, then, is whether Google's additional limitation – that the document has also not been seen by any user – comports with the patent's claims and specification. The claims themselves do not contain any reference to "any user."

PUM relies heavily on claim 7 in support of its argument. Claim 7 provides that a series of "properties" are "identified" for each document d, such as an author of the document or an age of the document. The patented system conducts this analysis for each and every potential document, regardless of whether the document has been seen by any user. One of the properties is the "number of users who have accessed the document d;" that number may be zero. PUM rightly points out that this claim relates back to limitation 1(d), which discloses "analyzing a document d to identify properties of the document." ('040 patent, col. 32 lines 37-38) Limitation 1(d) does not make any distinction based on seen or unseen documents; the unseen document language does not appear until limitation 1(e). PUM's basic point is that because claim 7, which depends from claim 1, potentially analyzes documents that have in fact been viewed by some users – that, after all, is one of the identified properties – the only way this makes sense is if the term unseen document refers to a document unseen by this particular user and not unseen by all users. The Court agrees.

Google's argument about prosecution history is unavailing. Google relies heavily on the way in which PUM distinguished the Gerace prior art reference. (D.I. 116 at 22) However, a prosecution history disclaimer arises only when a patentee expressly disclaims subject matter that

would otherwise be within the scope of the patent's claim(s). *See Chimie v. PPG Indus.*, 402 F.3d 1371, 1377 (Fed. Cir. 2005). Here, PUM did not do so.[24]

The Court is persuaded that the overall logic of the patent compels the conclusion that unseen document is one that is unseen as to the specific user. Thus, "unseen document" will be construed as a "document not previously seen by the user."

6. Claim 11

Moving on from claim 1, the next claim that raises a disputed term is claim 11, which depends from claim 1. Claim 11 reads, "The method of claim **1** further comprising estimating a posterior probability P(u/d,q) that the document d is of interest to the user u, given a query q submitted by the user." ('040 patent, col. 34 lines 1-3) The claim includes several disputed terms that the Court has already discussed, such as "estimating" and "probability," and the Court incorporates its discussion of those terms here by reference. The claim also raises a new concept of "posterior probability," with respect to which the Court must decide between the parties' competing constructions. (D.I. 115; D.I. 160 Ex. 1)

a. "estimating a posterior probability P(u/d,q) that a document d is of interest to the user u given a query q submitted by the user"

The dispute about this claim term is rather unclear. It appears that the dispute involves Google's attempt to engraft a "new knowledge" limitation onto the claim. (Tr. at 46)

Google argues that posterior probability is a "new probability that results from using our

---

[24]The Court's conclusion is not altered by the PTO's non-final decisions, issued on May 31, 2011, to grant Google's *inter partes* reexamination of the '040 patent and reject the asserted claims of that patent on the bases on which Google sought reexam. (D.I. 276) Although now part of the prosecution history, these PTO actions are not final and were taken before the PTO received any input from PUM. (D.I. 281)

additional knowledge." (D.I. 116 at 17) Google relies for its construction on a treatise to suggest that a posterior probability requires additional information. *See* Christopher D. Manning & Hinrich Schutze, *Foundations of Statistical Natural Language Processing* 42 (1999). (D.I. 118 Ex. E)

PUM proposes that the term should mean "approximating or roughly calculating the degree of belief or likelihood that a document d is of interest to the user u given the information that is known about the document, and given a query q." PUM attacks Google's proposed construction, contending it merely introduces additional language that creates "ambiguity and confusion." (D.I. 119 at 25)

The claim itself does not provide guidance about the meaning of posterior probability. The specification is also of limited assistance. The specification does provide that, "A number of other probabilities can be calculated, such as a posterior probability P(u/d,q) that the document is of interest to the user, given a search query submitted by the user." ('040 patent, col. 5 lines 34-37) The specification goes on to explain, "estimating the posterior probability includes estimating a probability that the query is expressed by the user with an information need contained in the document." (*Id.* col. 5 lines 37-39) Nothing in the specification indicates anything about any "new knowledge" involving the document d. Rather, the additional information that Google's treatise states is necessary to distinguish prior probability from posterior probability may, it seems, include the search query submitted by the user.

The Court will not engraft any additional limitation onto the claim. *See Omega Eng'g, Inc., v. Raytek Corp.*, 334 F.3d 1314, 1322 (Fed. Cir. 2003); *Stanacard, LLC v. Rebtel Networks, AB*, 680 F. Supp. 2d 483, 493 (S.D.N.Y. 2010) (rejecting defendant's proposed construction, as it

"serves only to introduce additional terms into the claim and would result in confusion for the jury").

The Court will, therefore, construe "estimating a posterior probability P(u/d,q) that a document d is of interest to the user u given a query q submitted by the user" to mean "approximating or roughly calculating a numerical degree of belief or likelihood that a document d is of interest to the user u given the information that is known about the document, and given a query q."[25]

B. '276 patent

Claim 1 of the '276 patent presents one additional term for construction. It also requires the Court to determine whether certain claim terms are indefinite. The Court will first address the disputed term, "present," and then turn to the issue of indefiniteness.[26]

1. "present"

The final limitation of claim 1 of the '276 patent requires "using the estimated probabilities for the respective plurality of retrieved documents to present at least a portion of the retrieved documents to the user." ('276 patent, col. 31 lines 63-65) Essentially, claim 1 refers to a method for providing search results that are tailored to a specific user. Thus, once the system has received a "search query," according to the claim, the system retrieves a "plurality" of

---

[25]The Court construes the disputed term "estimating a probability P(u/d) that an unseen document d is of interest to the user u," which appears in claim limitation 1(e), as "approximating or roughly calculating a numerical degree of belief or likelihood that an unseen document d is of interest to the user u given the information that is known about the unseen document."

[26]The disputed terms "estimating," "parameters," "learning machine," and "user-specific" all also appear in claim 1 of the '276 patent. They will each be given the same construction already described above in connection with the '040 patent.

documents and "presents" a portion of documents that are personalized to the user. ('276 patent, col. 32 lines 54-65)

The parties dispute the construction of "present" or "presenting." (D.I. 119 at 27; D.I. 116 at 29) Google argues that the ordinary meaning of present is "display," and Google cites a dictionary definition – to "show" or "display" – as well as numerous uses of the word "display" in the specification to support its construction. (D.I. 116 at 29) In Google's view, claim 1 of the '276 patent discloses personalization services in the context of searching the internet in a web browser; in that context, common sense dictates that presenting documents means displaying those documents to the user. (*Id.*)

PUM's argument about the term present has three prongs. First, PUM contends that claim 24, which depends from claim 23, differentiates the meaning of the term present from display. (D.I. 119 at 27) PUM asserts this proves that Google's construction must be wrong. Second, PUM argues that the specification consistently uses presenting and displaying in different ways. Displaying documents is a subset of situations in which results are actually shown to a user. (*Id.* at 28) Finally, the extrinsic evidence (dictionaries) purportedly supports PUM's construction.

Looking first at the claim language, claims 23 and 24 clearly distinguish between "presenting" and "displaying." ('276 patent, col. 34 lines 19-23) When the claims themselves make such a distinction, a court may not ignore the difference. *See AllVoice Computing PLC v. Nuance Commc'ns, Inc.*, 504 F.3d 1236, 1247-48 (Fed. Cir. 2007) (holding that different words have different meanings when claims and specification consistently used different words). Google argues that PUM's claim differentiation argument is without merit because the difference

between claim 23 and claim 24 involves a "personal web page" limitation that is found in claim 24 but not claim 23; therefore, in Google's view, the use of the word displaying in claim 24 actually means the same thing as presenting in claim 23. (D.I. 131 at 23) This argument is unpersuasive.

Further, the patent speaks of presenting information to a ***user***. In one embodiment, the system "collects and presents personal information to a user based on the User Model 13." ('276 patent, col. 8 lines 15-16) As Google also acknowledges, the specification discusses that the patented system presents information to a user. (D.I. 119 at 28) Given that the Court's construction of the term "user" includes the person operating the computer ***or*** that person's associated representation, presenting must mean, as a matter of logic, making available to the user. Displaying documents to an associated representation – such as a tag or identifier – does not make sense.

Finally, the dictionaries upon which both parties rely suggest that presenting may include displaying, but that presenting is not necessarily synonymous with displaying. *See The American Heritage Dictionary of the English Language*, 1101 (4th ed.2002) ("1. a. To introduce, especially with a formal ceremony . . . 2. To bring before the public . . . 3. a. To make a gift or award of . . . 4. a. To offer for observation, examination or consideration; show or display: ***present one's credentials***.") (emphasis omitted).

The Court agrees with PUM and will construe "present" and "presenting" to mean "to provide or make available."

2. Indefiniteness

The third step in claim 1 of the '276 patent provides "estimating parameters of a user-

specific learning machine based at least in part on the documents of interest to the user." ('276 patent, col. 31 lines 50-52) The dispute about this term centers on the phrase "of interest."[27]

Google argues that "of interest" is indefinite. The Federal Circuit has held claim terms indefinite when the terms were "completely dependent on a person's subjective opinion." *Datamize, LLC v. Plumtree Software, Inc.*, 417 F.3d 1342, 1350 (Fed. Cir. 2005) (finding "aesthetically pleasing" indefinite). In Google's view, because "[n]othing in the intrinsic record provides an objective basis to determine whether a document is of interest or not of interest," the claim involves a purely subjective question. (D.I. 116 at 23) Google relies on *The Concise Oxford Dictionary of Current English* to argue that "interest" is a personal quality and "lies in the eye of the beholder." (D.I. 116 at 24)

PUM responds by pointing out that indefiniteness is a question of invalidity and, accordingly, some courts have at times refused to entertain indefiniteness arguments as part of the claim construction process. *See, e.g., Telecordia Techs., Inc. v. Alcatel USA, Inc.*, C.A. No. 04-874-GMS, Order at 2 (D. Del. Apr. 21, 2006) ("[T]he Court does not permit summary judgment arguments, including indefiniteness arguments, during the claim construction phase of the litigation.") (D.I. 120 Ex. 13); *NetRatings, Inc. v. Coremetrics, Inc.*, C.A. No. 05-314-GMS, Order at 1-2 (D. Del. June 7, 2006) (same) (D.I. 120 Ex. 14). PUM further contends that even if the Court were inclined to examine indefiniteness, a claim may be found indefinite only if it is "insolubly ambiguous," such that "no narrowing construction is possible." (D.I. 119 at 29; *see*

---

[27]Similarly, dependent claim 21 of the '040 patent discloses, "The method of claim **1** further comprising sending to a third party web server ***user interest information derived from the User Model***, whereby the third party web server may customize its interaction with the user." ('040 patent, col. 34 lines 42-45) (emphasis added) The phrase "user interest information" raises the same indefiniteness issue and the Court's discussion applies to both disputes.

also Exxon Research & Eng'g Co. v. United States, 265 F.3d 1371, 1375 (Fed. Cir. 2001)) Here, PUM argues that the specification supports a construction to the effect that positive and negative user actions illustrate whether a document is "of interest" to the user or not. (D.I. 116 at 30)

In order to withstand an argument that claims are indefinite, the terms in a claim must be "sufficiently definite to inform the public of the bounds of the protected invention." *Halliburton Energy Servs. v. M-I LLC,* 514 F.3d 1244, 1249-1250 (Fed. Cir. 2008); *see also* 35 U.S.C. § 112. *Halliburton* noted that there are three situations in which the Federal Circuit has recognized claims as indefinite: (1) a claim that "recites means-plus-function limitations without disclosing corresponding structure in the specification;" (2) a claim using a "numeric limitation without disclosing which of multiple methods of measuring that number should be used;" and (3) a claim that "contains a term completely dependent on a person's subjective opinion." *Id.* at 1249. The "common thread" in all three circumstances is that a person with "ordinary skill in the art could not determine the bounds of the claims," rendering the claim term "insolubly ambiguous." *Id.* Here, the only basis on which Google can argue indefiniteness is the third circumstance: that "of interest" and "not of interest" are entirely dependent on a person's subjective opinion.

As a general matter, proving indefiniteness requires meeting an "exacting standard." *Halliburton*, 514 F.3d at 1249. So long as one with ordinary skill in the art can ascertain a "workable objective standard," such that a claim term is not "completely dependent on a person's subjective opinion," the term will not be deemed indefinite. *Datamize*, 417 F.3d at 1350. "[T]his standard is met where an accused infringer shows by clear and convincing evidence that a skilled artisan could not discern the boundaries of the claim based on the claim language, the

specification, and the prosecution history, as well as her knowledge of the relevant art area." *Halliburton*, 514 F.3d at 1249-50.

The Court concludes that this standard has not been met with respect to the claim limitations under consideration here: documents "of interest," documents "not of interest," and "user interest information derived from the User Model." In *Datamize*, 417 F.3d at 1350, the claim limitation provided that the interface screen must be (1) uniform and (2) aesthetically pleasing. The patent provided no guidance as to what factors would be used to determine whether a screen was aesthetically pleasing. There was no way for artisans with ordinary skill to look at an interface screen to determine if it practiced the aesthetically pleasing limitation of the claim and infringed the patent.

An analogous situation might be presented in the instant case if the claims contained a limitation that the document be "interesting." Here, however, the claim term does not require one with ordinary skill to make any kind of subjective determination. Instead, the operative question is whether a document is of interest ***to a user***. The subjective component, to the extent that it can be characterized as subjective, is in reference to a third party. One having ordinary skill in the art would understand there is an objective standard for determining a user's reactions to documents.

As the specification explains:

> Through his or her actions, the user creates positive and negative patterns. Positive examples are documents of interest to a user: search results that are visited following a search query, documents saved in the user favorites or bookmarks file, web sites that the user visits independently of search queries, etc. Negative examples are the documents that are not of interest to the user, and include search results that are ignored although appear at the top of the search result, deleted bookmarks, and ignored pushed news or

email.

('040 patent, col. 22 lines 15-23) The specification goes on to note that the "degree of interest [in a document], for example, whether it was positive or negative, saved in the bookmarks file, how long the user spent viewing the document" are all recorded and stored in an "accessed buffer for subsequent analysis." ('276 patent, col. 22 lines 14-17) Based on all of these interactions, the system determines a "metric" for each document "to indicate whether it is a positive, negative, or neutral event." (*Id.* at col. 22 lines 20-21) The specification further discloses a mechanism for transforming user interactions into mathematical terms: 0 is a "completely negative event," and 1 is a "completely positive event." (*Id.* at col. 22 lines 22-24)

Google has not demonstrated by clear and convincing evidence that the term "of interest" is indefinite. Nor has Google identified other disagreements with PUM's proposed constructions of the terms under discussion. Hence, the Court will construe the term "documents of interest" to mean "documents [i.e., electronic files (including text or any type of media)] for which the user has a positive response," and "documents not of interest" to mean "documents [i.e., electronic files (including text or any type of media)] for which the user has a negative response or has ignored."[28]

C. Global Disputes

Two final issues presented by Google remain to be addressed. First, Google argues that many of the terms in the claims, including some terms that this Court has construed earlier, lack an antecedent basis. Second, Google contends that the steps of the claims must be performed in

---

[28]Consistent with this analysis, the Court will construe the disputed term "user interest information derived from the User Model" as "interests or other information inferred from the User Model."

a specific order. The Court now turns its attention to these disputes.

1. Antecedent Basis Terms

Certain terms and phrases contained in the '040 and '276 patents are first introduced with an indefinite article, "a" or "an," or with no article. These same terms and phrases are then sometimes referred to later in the claim with the definite article "the." For example, limitation 1(b) of the '040 patent recites, "updating ***user-specific data files***, wherein ***the user specific data files*** comprise. . ." ('040 patent, col. 32 lines 29-30) (emphasis added) The term "user-specific data files" has no article initially, but the claim then refers to "the" user-specific data files.[29] The parties' dispute is whether, in such situations, the latter introduced term necessarily refers to the former term. (D.I. 119 at 9; D.I. 116 at 24-25) In other words, does the latter term have the former term as its antecedent basis. *See Halliburton*, 514 F.3d at 1249 (explaining claim could be indefinite "if a term does not have proper antecedent basis where such basis is not otherwise present by implication or the meaning is not readily ascertainable").

Google argue that, "[a]s a matter of common sense, where these terms are first introduced with 'a,' and then later used with 'the,' the latter term must be referring to the former term." (D.I. 116 at 24-25) PUM responds that none of the disputed claim terms require construction. (D.I. 115 at 1-2) In PUM's view, the context of the terms makes clear when a term refers to an earlier term and when it does not. PUM contends that Google has not demonstrated any basis for applying a blanket rule for all of these so-called "antecedent basis" terms. (D.I. 119 at 11)

For the most part, the Court agrees with Google. For all except two of the antecedent basis terms – "document d" in limitation 1(e) and "estimated probability" in claim 1(f) – the

[29]The "antecedent basis terms" are set out at D.I. 161 Ex. A at 8-9.

Court finds that the terms using definite articles refer to the former terms in the same limitation. However, the Court sees no need to inject further language and potential confusion into the claims when, as here, the claims, the context, and the ordinary meaning make clear that terms refer to each other. Two claims terms – "document d" in limitation 1(e) and "estimated probability" in claim 1(f) – require further discussion.

a. "the document" in limitation 1(e)

Limitation 1(e) of the '040 patent states, "estimating a probability P(u/d) that ***an unseen document d*** is of interest to the user u, wherein the probability P(u/d) is estimated by applying the identified properties of ***the document*** to the learning machine having the parameters defined by the User Model." ('040 patent, col. 32 lines 39-43) (emphasis added) Here, "the document" refers to the same "unseen document" recited in the limitation. A related question, however, is whether document d in limitation 1(e) refers to the same document d recited in limitation 1(d). (D.I. 138 at 21)

Google argues that "the document" in limitation 1(e) does refer to the document previously recited in 1(d). Google argues that since both limitations, 1(d) and 1(e), use the notation "d," the patent clearly intends that the two terms refer to the same document.

In PUM's view, on the other hand, limitation 1(e) introduces a new "subset" of document d: that subset of documents that are unseen. Thus, since the document in limitation 1(e) clearly refers back to the unseen document from the same 1(e) limitation, and since not all of the documents in limitation 1(d) will be unseen, the two terms will not always refer back to one another. (*Id.*)

The claim language – particularly the structure of the claims – is instructive. Limitation

1(e) provides for "estimating a probability P(u/d) that an unseen document d . . . by applying the identified properties of the document." (*Id.* at col. 32 lines 39-42) The unseen document of 1(e) has gone through the analysis required by limitation 1(d) ("analyzing a document d to identify properties of the document"), but that does not mean that "the document" in limitation 1(e) always and only refers to the same document in limitation 1(d). As PUM points out, some of the documents analyzed in limitation 1(d) will ***not*** be unseen documents and, therefore, logically, the document in limitation 1(e), which ***will*** always be unseen, will in some cases not refer to the same document.

The Court concludes that the term "document" in limitation 1(e) refers to the unseen document recited in the same limitation. The document recited in limitation 1(e), however, does not necessarily refer to the same document in limitation 1(d). No particular construction is necessary as these conclusions are clear from the structure of the claim.

b. "estimated probability" in limitation 1(f)

Limitation 1(e) provides, "estimating a probability P(u/d) . . . wherein the probability P(u/d) is estimated by applying. . ." (*Id.* lines 39-42) As with several of the limitations, limitation 1(e) follows the structure of introducing a new term – a probability P(u/d) – and then proceeding to explain the same term later in the limitation. Thus, "the probability P(u/d)" clearly refers back to "a probability P(u/d)" that the limitation 1(e) recites.

The issue is whether the term "the estimated probability" found thereafter in limitation 1(f) has as its antecedent basis the same probability recited in the previous limitation, 1(e). A corollary issue is whether the estimated probability is different for each unseen document. The Court finds that the answer to both questions is yes.

PUM concedes that "'the estimated probability' of step (f) refers back to the estimated 'probability P(u/d)' of step (e)." (D.I. 119 at 10) PUM further concedes that this process is repeated for each unseen document: the patent teaches identifying the properties of a document, estimating the probability of an unseen document, and finally using the estimated probability to provide the personalized services to the user. The result, in PUM's view, is that there is not a "single probability" but, instead, a different probability for each unseen document. (*Id.*)

In Google's view, the parties essentially agree on the construction of this term. Google agrees with PUM that the "estimated probability of step (f) refers back to the estimated probability P(u/d) of step (e)." (D.I. 138 at 21) Google contends that the only dispute is PUM's fear Google will attempt to limit the probabilities to a single probability. Google insists it is not suggesting there must be only a single probability. (D.I. 138 at 21)

In some circumstances, claim terms with no clear antecedent basis may have an implicit antecedent basis. *See Energizer Holdings,* 435 F.3d at 1370-71 (finding "anode gel" is by implication the antecedent basis for "said zinc anode"). In other words, claim terms without an explicit antecedent basis may nevertheless be definite if the term "has a reasonably ascertainable meaning . . . decided in context." *Id.* The Court concludes – as both parties appear to agree – that the antecedent basis for the term "estimated probability" is the "probability that an unseen document d is of interest to the user," as the previous limitation 1(e) discloses. ('040 patent, col. 32 lines 39-40; *see also id.* col. 4 lines 31-36 ("The parameters are used to estimate a probability P(u/d) that a document is of interest to the user, and the estimated probability is ***then*** used to provide personalized information services to the user. The probability is estimated by analyzing properties of the document and applying them to the learning machine.") (emphasis added))

The patent also makes clear that this process of estimating a probability is dynamic, which requires there to be multiple probabilities as opposed to a single probability. According to the specification, the "User Model is a dynamic entity that is refined and updated based on all user interactions." ('040 patent col. 21 lines 63-64; *see also id.* col. 9 lines 6-8) ("In response to the services provided, the user performs a series of actions, and these actions are in turn monitored to further update the User Model.")) The Court agrees with PUM that there is "not a single probability," but instead "a probability for each unseen document."

Thus, to one with ordinary skill in the art, the antecedent basis of the term estimated probability in limitation 1(f) is reasonably ascertainable based on the claims and the specification. Specifically, "the estimated probability" of limitation 1(f) refers to the probability that was estimated in the previous limitation, 1(e). Additionally, the patent does not require a single probability. Because the parties effectively agree on all of this and it is clear from the claim language, the Court finds it unnecessary to provide any further construction.

2. Order of Steps

Google argues that certain claims – claims 1 and 32 of the '040 patent and claims 1 and 23 of the '276 patent – contain steps that must be performed in a specific sequence. Google, relying on language from the Federal Circuit, contends that when "each subsequent step" of a method claim "references something logically indicating the prior step has been performed," then the patent teaches that the steps must be performed in an order. (D.I. 119 at 26; *see also Interactive Gift Express, Inc. v. Compuserve Inc.*, 256 F.3d 1323, 1342 (Fed. Cir. 2001))

Limitation 1(b) of the '040 patent, for example, discloses updating user-specific data files "compris[ing] the monitored user interactions," and Google contends that this updating process

cannot be completed until after the user interactions have been monitored, which occurs during step 1(a). Google maintains that limitations 1(a) – 1(c) must be performed in a fixed order: a then b then c. Further, in Google's view, a through c must be completed before steps 1(e) and 1(f) occur. Also, Google's construction requires that 1(d) must be performed before 1(e) and 1(f).

PUM responds that no construction is necessary, but "[i]f the Court is inclined to address the issue, then it should hold that the steps may be performed in a consecutive manner, in an overlapping manner, or a combination of the two," with some exceptions. (D.I 119 at 6) In PUM's view, "[a]lthough it may be preferable that the steps occur in a particular order, with certain limited exceptions, such an order is not required." (*Id.* at 7) Basically, PUM contends that in the '040 patent, the properties of the document must be identified (limitations 1(d) and 32(d)) before the identified properties are applied to the learning machine as required in limitations 1(e) and 32(e); similarly, in the '276 patent, the search query is received (limitation 1(d)) before documents are retrieved based on the search query. (*Id.* at 7) To PUM, other than these exceptions the claims do not require a fixed order of steps.

Generally, "[u]nless the steps of a method actually recite an order, the steps are not ordinarily construed to require one." *Interactive*, 256 F.3d at 1342. Neither party contends that the patents explicitly provide that the steps must be performed in an order. In some cases, however, "the method steps implicitly require that they be performed in the order written." *Id.; see also Loral Fairchild Corp. v. Sony Corp.*, 181 F.3d 1313, 1322 (Fed. Cir. 1999). Courts use a two-part test to determine if, in the absence of an explicit sequential order, method steps must nevertheless be performed in order. First, a court looks "to the claim language to determine if, as

a matter of logic or grammar, [the steps] must be performed in the order written." *Altiris, Inc. v. Symantech Corp.*, 318 F.3d 1363, 1369 (Fed. Cir. 2003). If the logic or grammar of the claim language does not compel an order, courts "next look to the rest of the specification to determine whether it directly or implicitly requires such a narrow construction." *Id.* (internal quotation marks omitted). If the invention could potentially be performed without going in the specific sequence, then the claim does not require an order of steps. *See id.* at 1371.

Here, the method does not explicitly require an order. Unless the patent explicitly imposes an order, normally an order is not required. *See Interactive*, 256 F.3d at 1342. Second, as PUM points out, the invention is dynamic, and the specification emphasizes that the User Model will not be updated after every newly reviewed document or search. ('040 patent, col. 26 lines 4-15; *see also id.* col. 22 lines 55-63) As a matter of logic, if the specification teaches that some documents may be viewed without having the User Model updated, which is recited in an earlier step, the corollary is that limitation 1(d) (analyzing) may occur in some instances without limitation 1(b) (updating) having occurred in the same "cycle." Google appears to concede that some steps may not be performed in every "cycle." (Tr. at 109, 112-13) This concession supports PUM's position.

The Court, thus, agrees with PUM and will not impose an order of steps on the patent. To the extent that some steps presuppose that one of the earlier steps be performed, that (limited) required order of steps is clear from the claim language itself. No construction is necessary.

IV. CONCLUSION

For the reasons given above, the Court will construe the disputed terms as described. An appropriate Order follows.

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

PERSONALIZED USER MODEL LLP	:	
	:	
Plaintiff,	:	
	:	
v.	:	Civ. No. 09-525-LPS
	:	
GOOGLE INC.	:	
	:	
Defendant.	:	

ORDER

At Wilmington this 25th day of January, 2012:

For the reasons set forth in the Opinion issued this same date, the Court construes the disputed claim terms in the '040 and '276 patents as follows:

1. "user" and "user u" mean "a person operating a computer or the associated representation of the user."

2. "user-specific data files" means "the monitored user interactions with data and a set of documents associated with the user."

3. "monitored user interactions with the data" means "the collected information about the user's interactions with data."

4. "parameters" means "values or weights."

5. "estimating parameters of a learning machine" means "estimating values or weights of the variables of a learning machine."

6. "learning machine" means a "mathematical function and/or model used to make a prediction, that attempts to improve its predictive ability over time by altering the values/weights given to its variables, depending on a variety of knowledge sources, including monitored user

interactions with data and a set of documents associated with the user."

7. "User Model specific to the user" means "an implementation of a learning machine updated in part by data specific to the user."

8. "user-specific learning machine" means "a learning machine [as construed] specific to the user."

9. "document" means "an electronic file including text or any type of media."

10. "estimating" means "approximating or roughly calculating."

11. "probability" means "numerical degree of belief or likelihood."

12. "unseen document" means "document not previously seen by the user."

13. "estimating a probability P(u/d) that an unseen document d is of interest to the user u" means "approximating or roughly calculating a numerical degree of belief or likelihood that an unseen document d is of interest to the user u given the information that is known about the unseen document."

14. "estimating a posterior probability P(u/d,q) that a document d is of interest to the user u given a query q submitted by the user" means "approximating or roughly calculating a numerical degree of belief or likelihood that a document d is of interest to the user u given the information that is known about the document, and given a query q."

15. "present" and "presenting" mean "to provide or make available."

16. "documents of interest to the user" means "documents [i.e., electronic files (including text or any type of media)] for which the user has a positive response."

17. "documents not of interest to the user" means "documents [i.e., electronic files (including text or any type of media)] for which the user has a negative response or has ignored."

18. "user interest information derived from the User Model" means "interests or other

information inferred from the User Model."

The Court will also adopt the parties' agreed-upon constructions, as follows:

1. "set" means "group or collection."

2. "set of documents associated with the user" means "group or collection of documents associated with the user."

3. "automatic" means "without human intervention."

4. "central computer" means "computer on the server side of a client-server relationship."

The Court further concludes that no construction is necessary to resolve the parties' disputes with respect to the antecedent basis and order of steps issues.

UNITED STATES DISTRICT JUDGE