UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF TEXAS
MARSHALL DIVISION

NORTHEASTERN UNIVERSITY and JARG CORPORATION	§	
vs.	§	CASE NO. 2:07-CV-486-CE
GOOGLE, INC.	§	

MEMORANDUM OPINION AND ORDER

After considering the submissions and the arguments of counsel, the court issues the following order concerning the claim construction issues:

I. Introduction

In this case, plaintiffs Northeastern University[1] and Jarg Corporation contend that Google infringes various claims of U.S. Patent No. 5,694,593 ("the '593 patent"). The patent covers search database architecture and includes both method and system claims. Defendant is in the business of providing search services which Plaintiffs claim infringe both the method and system claims.

This memorandum addresses the parties' claim construction disputes. The memorandum will first briefly summarize the subject matter of the patent before addressing the merits of the parties' claim construction positions.

II. Technical Summary

The patent in suit relates to a distributed computer database system. Specifically, the patent describes and claims a non-relational database system wherein information is stored on individual

[1] Northeastern is the patent owner and Jarg Corporation is its licensee.

computers called nodes, those nodes are networked together, and queries to access the stored information are divided and into smaller fragments that are each analyzed using a hash function. After each query fragment is hashed, the computer analyzing the query sends a request for information to another computer on the network based on a portion of that hashed query fragment. The computer that receives the request uses part of the hashed query fragment to retrieve information and return it to the computer analyzing the query. The claims refer to the node analyzing the query as the "home node" and the nodes receiving the hashed query parts as "query nodes." The claims additionally specify that the hashed query fragment has a "first portion" and a "second portion," and that the first portion is used to rout the hashed query fragment to a query node where the second portion is used to retrieve an "object identifier" corresponding to the relevant stored data.

The patent abstract is quoted below:

> A distributed computer database system including a front end computer and a plurality of computer nodes interconnected by a network into a search engine. A query from a user is transmitted to the front end computer which forwards the query to one of the computer nodes, termed the home node, of the search engine. The home node fragments the query and hashes the fragments of query to create an index by which the hashed query fragments are transmitted to one or more nodes on the network. Each node on the network which receives a hashed fragment uses the fragment of the query to perform a search on its respective database. The results of the searches of the local database are then gathered by the home node.

Plaintiffs have asserted independent claims 1, 8, and 13 against Defendant. Claim 1 is quoted below:

> 1. A method for information retrieval using fuzzy queries in a non-relational, distributed database system having a plurality of home nodes and a plurality of query nodes connected by a network, said method comprising the steps of:
>
> randomly selecting a first one of said plurality of home nodes;

> fragmenting, by said selected home node, a query from a user into a plurality of query fragments;
>
> hashing, by said selected home node, each said query fragment of said plurality of query fragments, said hashed query fragment having a first portion and a second portion;
>
> transmitting, by said selected home node, each said hashed query fragment of said plurality of query fragments to a respective one of said plurality of query nodes indicated by said first portion of each said hashed query fragment;
>
> using, by said query node, said second portion of said respective hashed query fragment to access data according to a local hash table located on said query node; and
>
> returning, by each said query node accessing data according to said respective hashed query fragment, an object identifier corresponding to said accessed data to said selected home node.

III. General Principles Governing Claim Construction

"A claim in a patent provides the metes and bounds of the right which the patent confers on the patentee to exclude others from making, using or selling the protected invention." *Burke, Inc. v. Bruno Indep. Living Aids, Inc.*, 183 F.3d 1334, 1340 (Fed. Cir. 1999) (quoting *Corning Glass Works v. Sumitomo Elec. U.S.A., Inc.*, 868 F.2d 1251, 1257 (Fed. Cir. 1989)). Claim construction is an issue of law for the court to decide. *Markman v. Westview Instruments, Inc.*, 517 U.S. 370, 391 (1996).

To ascertain the meaning of claims, the court looks to three primary sources: the claims, the specification, and the prosecution history. *Markman v. Westview Instruments, Inc.*, 52 F.3d 967, 979 (Fed. Cir. 1995), *aff'd*, 517 U.S. 370 (1996) (quoting *Unique Concepts, Inc. v. Brown*, 939 F.2d 1558, 1561 (Fed. Cir. 1991)). Under the patent law, the specification must contain a written description of the invention that enables one of ordinary skill in the art to make and use the

invention. 35 U.S.C. § 112; *id.* at 978. A patent's claims "must be read in view of the specification, of which they are a part." *Markman*, 52 F.3d at 979. "For claim construction purposes, the description may act as a sort of dictionary, which explains the invention and may define terms used in the claims." *Id*. "One purpose for examining the specification is to determine if the patentee has limited the scope of the claims." *Watts v. XL Sys., Inc.*, 232 F.3d 877, 882 (Fed. Cir. 2000).

Nonetheless, it is the function of the claims, not the specification, to set forth the limits of the patentee's claims. Otherwise, there would be no need for claims. *SRI Int'l v. Matsushita Elec. Corp*., 775 F.2d 1107, 1121 (Fed. Cir. 1985) (en banc). The patentee is free to be his own lexicographer, but any special definition given to a word must be clearly set forth in the specification. *Intellicall, Inc. v. Phonometrics*, 952 F.2d 1384, 1388 (Fed. Cir. 1992). And, although the specification may indicate that certain embodiments are preferred, particular embodiments appearing in the specification will not be read into the claims when the claim language is broader than the embodiments. *Electro Med. Sys., S.A. v. Cooper Life Scis., Inc.*, 34 F.3d 1048, 1054 (Fed. Cir. 1994).

This court's claim construction decision must be informed by the Federal Circuit's decision in *Phillips v. AWH Corp.*, 415 F.3d 1303 (Fed. Cir. 2005) (en banc). In *Phillips*, the court set forth several guideposts that courts should follow when construing claims. In particular, the court reiterated that "the *claims* of a patent define the invention to which the patentee is entitled the right to exclude." *Id.* at 1312 (quoting *Innova/Pure Water, Inc. v. Safari Water Filtration Sys., Inc.*, 381 F.3d 1111, 1115 (Fed. Cir. 2004)) (emphasis added). To that end, the words used in a claim "are generally given their ordinary and customary meaning." *Id*. (quoting *Vitronics Corp. v. Conceptronic, Inc.*, 90 F.3d 1576, 1582 (Fed. Cir. 1996)). "[T]he ordinary and customary meaning

of a claim term 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 1313. This principle of patent law flows naturally from the recognition that inventors are usually persons who are skilled in the field of the invention. *Id.* The patent is addressed to and intended to be read by others skilled in the particular art. *Id*.

The primacy of claim terms notwithstanding, *Phillips* made clear that "the person of ordinary skill in the art is deemed to read the claim term not only in the context of the particular claim in which the disputed term appears, but in the context of the entire patent, including the specification." *Phillips*, 415 F.3d at 1313. Although the claims themselves may provide guidance as to the meaning of particular terms, those terms are part of "a fully integrated written instrument." *Id*. at 1315 (quoting *Markman*, 52 F.3d at 978). Thus, the *Phillips* court emphasized the specification as being the primary basis for construing the claims. *Id*. at 1314-17. The Supreme Court stated long ago that "in case of doubt or ambiguity it is proper in all cases to refer back to the descriptive portions of the specification to aid in solving the doubt or in ascertaining the true intent and meaning of the language employed in the claims." *Bates v. Coe*, 98 U.S. 31, 38 (1878). In addressing the role of the specification, the *Phillips* court quoted with approval its earlier observations from *Renishaw PLC v. Marposs Societa' per Azioni*, 158 F.3d 1243, 1250 (Fed. Cir. 1998):

> Ultimately, the interpretation to be given a term can only be determined and confirmed with a full understanding of what the inventors actually invented and intended to envelop with the claim. The 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.

Phillips, 415 F.3d at 1316. Consequently, *Phillips* emphasized the important role the specification plays in the claim construction process.

The prosecution history also continues to play an important role in claim interpretation. The prosecution history helps to demonstrate how the inventor and the PTO understood the patent. *Id.* at 1317. Because the file history, however, "represents an ongoing negotiation between the PTO and the applicant," it may lack the clarity of the specification and thus be less useful in claim construction proceedings. *Id*. Nevertheless, the prosecution history is intrinsic evidence. *Id.* That evidence is relevant to the determination of how the inventor understood the invention and whether the inventor limited the invention during prosecution by narrowing the scope of the claims. *Id.*

Phillips rejected any claim construction approach that sacrificed the intrinsic record in favor of extrinsic evidence, such as dictionary definitions or expert testimony. *Id.* The *en banc* court condemned the suggestion made by *Texas Digital Systems, Inc. v. Telegenix, Inc.*, 308 F.3d 1193 (Fed. Cir. 2002), that a court should discern the ordinary meaning of the claim terms (through dictionaries or otherwise) before resorting to the specification for certain limited purposes. *Phillips*, 415 F.3d at 1319-24. The approach suggested by *Texas Digital*–the assignment of a limited role to the specification–was rejected as inconsistent with decisions holding the specification to be the best guide to the meaning of a disputed term. *Id*. at 1320-21 (quoting *Vitronics*, 90 F.3d at 1582). According to *Phillips*, reliance on dictionary definitions at the expense of the specification had the effect of "focus[ing] the inquiry on the abstract meaning of words rather than on the meaning of claim terms within the context of the patent." *Id*. at 1321. *Phillips* emphasized that "[t]he patent system is based on the proposition that the claims cover only the invented subject matter." *Id.* What is described in the claims flows from the statutory requirement imposed on the patentee to describe and particularly claim what he or she has invented. *Id*. The definitions found in dictionaries, however, often flow from the editors' objective of assembling all of the possible definitions for a

word. *Id*. at 1321-22.

Phillips does not preclude all uses of dictionaries in claim construction proceedings. *Phillips*, 415 F.3d at 1322. Instead, the court assigned dictionaries a role subordinate to the intrinsic record. *Id.* at 1317-19. In doing so, the court emphasized that claim construction issues are not resolved by any "magic formula." *Id.* at 1324. The court did not impose any particular sequence of steps for a court to follow when it considers disputed claim language. *Id*. at 1323-25. Rather, *Phillips* held that a court must attach the appropriate weight to the intrinsic sources offered in support of a proposed claim construction, bearing in mind the general rule that the claims measure the scope of the patent grant. *Id.* at 1324.

IV. Agreed Terms

The parties have stipulated to the construction of the following term:

- The parties agree that "fuzzy queries" means "imprecise or inexact requests for information from a database, the result of which does not necessarily contain each term in the query." The term "fuzzy queries" is found in claim 1 of the '593 patent.

V. Disputed Terms

The parties dispute the construction for several terms relating to the database system. Specifically, the parties dispute whether the computers in the system can have specialized functions, how computers are selected for each query, how to define a "non-relational database," what constitutes "hashing," whether and to what extent the first and second portions of the hashed query fragment can overlap, what is meant by "transmitting" information from one computer to another, and what properties are required of the hash tables stored on the "query nodes."

A. "query fragment"

- Plaintiffs' proposed construction: "a sub-part or piece of a query"
- Defendant's proposed construction: "a part of a query consisting of a limited number of attributes and attribute values joined by relationships, specified in the same formal, artificial language and ontology which describes the attributes values of objects in the database"

The term "query fragment" appears in asserted claims 1, 8 and 13 of the '593 patent. The parties agree that this term describes a sub-part or piece of a "query," but disagree as to what a constitutes a "query." Plaintiffs argue that no further construction is necessary while Defendant seeks to incorporate attributes of the prior art described in the background section of the specification. In the context of "fuzzy queries," the parties agreed that a "query" is a "request for information from a database."

Defendant draws support from the background of the invention, which recites many limitations of conventional database systems, as well as from the prosecution history of U.S. Patent No. 6,505,191 ("the '191 patent"). The '191 patent is an unrelated patent filed by the same inventor, but with different prosecuting attorneys. As part of a response to an obviousness-type double-patenting rejection over the '593 patent, the inventor argued that a passage from column 1, lines 27-31 of the '593 patent defines the term "query fragments." Defendant does not argue directly that the passage from the background of the invention defines the term "query fragment," but instead relies on the statements made during the prosecution of the '191 patent.

As legal support for its argument that the prosecution history of the '191 patent can be used to limit the '593 patent, Defendant cites *Omega Engineering, Inc. v. Raytek Corp.*, 334 F.3d 1314

(Fed. Cir. 2003). Defendant maintains *Omega* holds that "a later patent can inform the understanding of terms in an earlier patent." *Omega* reached no such holding, however. Rather, *Omega* addresses the situation where the prosecution history of a parent patent or a previously issued patent with a common parent contains a narrowing definition or disavowal of claim scope of a term common to both patents. 334 F.3d at 1333-1334. *Omega* does not address the issue whether a parent patent can be limited by the prosecution history of a child application, much less whether a patent should be limited by the prosecution history of a later, unrelated application. *Id.* Accordingly, the Court treats the prosecution history of the '191 patent as extrinsic evidence in construing the claims of the '593 patent and gives the inventor's statements there less weight. *See Phillips*, 415 F.3d at 1317-18.

The inventor's statement in the prosecution of the '191 patent contradicts the intrinsic record in the '593 patent. The relevant passage of the '593 patent makes no mention of "query fragments," but instead discusses limitations of queries in relational databases. After discussing these limitations, the background section concludes by stating that "the present invention avoids these limitations." Further, the hashing process described and claimed in the '593 patent would be useless if all queries and query fragments had to be written according to a "formal, artificial language and ontology." The data location and retrieval system described by the patent is designed to eliminate that limitation present in the prior art. Defendant's proposed definition would provide for the opposite of what the patentee intended when he described and claimed his invention.

Plaintiffs' proposed construction, on the other hand, does not embrace the parties' agreed construction of "fuzzy queries." The parties have agreed that "fuzzy queries" means "imprecise or inexact requests for information from a database, the result of which does not necessarily contain

each term in the query." The portion of the "fuzzy query" definition corresponding to the term "query" is "request for information from a database." If the Court is adopting a definition for "query" as it has in the agreed construction of "fuzzy queries" then that definition should be consistent. Accordingly, the Court construes "query fragment" to mean "a part or piece of a request for information from a database."

B. "randomly selecting"

- Plaintiffs' proposed construction: "selecting without apparent pattern"
- Defendant's proposed construction: "selecting by chance, independently of preceding selections, where each item in the set has equal probability of being chosen"

The term "randomly selecting" appears in asserted claim 1 of the '593 patent. The parties disagree over how to define "randomly." Defendant attempts to use a general dictionary definition to define "randomly" to mean a very specific kind of random selection that is not supported by the intrinsic record and far more specific than would be understood by one of ordinary skill in the art.

The field of the invention is computer science. "Random" has a different meaning in computer science than in the general vernacular. It is generally understood that computers do not perform "randomly," but may be programmed to behave pseudorandomly.[2] In other words, "random" in computer science means that something approximates random, not that it is random. Further, there are many different kinds of "randomness" in computer science and mathematics.

---

[2] Mathematician and computer scientist John von Neumann famously said, "Anyone who considers arithmetical methods of producing random digits is, of course, in a state of sin. For, as has been pointed out several times, there is no such thing as a random number – there are only methods to produce random numbers, and a strict arithmetic procedure of course is not such a method." *Various Techniques Used in Connection with Random Digits*, 3 J. OF RES. OF THE NAT'L BUREAU OF STANDARDS, APPLIED MATHEMATICS SERIES 36, 36-38 (1951) (remark made at a symposium on Monte Carlo method).

Some "randomness" has uniform distribution and appears as the result of independent trials, as is suggested by Defendant's coin flip analogy. Other "randomness" has a non-uniform distribution, such as the heights of individuals in a population. "Randomness" may also appear in variables that have "memory," such as a deck of cards at a blackjack table. The specification and the cited portions of the prosecution history do not specify what type of "randomness" is employed by the claimed invention, and it would be improper to arbitrarily limit the invention to a specific type of random behavior.

Plaintiffs' construction is also improper. The term "apparent pattern" requires an observer, and whether a pattern appears to the observer would be a subjective inquiry. Over a long enough period to an observer with sufficient memory, no computer-generated "randomness" will lack an apparent pattern.

In light of the above concerns, the Court construes "randomly selected" to mean "selected using random computer methods." The parties are not to argue to the jury that the pseudorandom number generators commonly employed as random number generators in modern computers are not, in fact, random. Additionally, the Court rejects any suggestion that a specific probability distribution is required to meet this limitation.

C. "a first portion and a second portion"

- Plaintiffs' proposed construction: the claim language has its plain and ordinary meaning; no further construction is necessary
- Defendant's proposed construction: "a first part separate and distinct from a second part"

The "first portion" and "second portion" limitations appear in asserted claims 1, 8, and 13

of the '593 patent. In the main, the parties dispute whether the first portion and second portion of a hashed query fragment[3] may overlap, and, if they may overlap, to what extent they may overlap.

Central to the parties' disagreement on this point is their interpretation of *Linear Tech. Corp. v. ITC*. 566 F.3d 1049 (Fed. Cir. 2009). In *Linear Tech*., the Federal Circuit reviewed a claim construction for a claim that recited a "second circuit" and a "third circuit" where the patent holder argued that the circuits may share components and structure so long as both performed the required functions while the defendant argued that the circuits must be entirely separate and distinct. *Id.* at 1055. The Federal Circuit held that, absent an explicit indication that the second and third circuits must be separate or distinct, the circuits need not be distinct so long as they performed the claimed functions. *Id*. Plaintiffs argue that *Linear Tech.* supports its construction while Defendant argues that *Linear Tech*. is inapposite to the general rule that separately mentioned claim terms require separate and distinct structures. Defendant's argument runs contrary to the clear language of *Linear Tech.* and relies upon cases where substantial intrinsic evidence supported the construction of separate and distinct elements.[4] The claims of the '593 patent do not require the hashed query

---

[3] As will be explained later, a "hashed query fragment" is a number. A portion of a hashed query fragment is some set of the digits of that number. For example, if a hashed query fragment contained 37 digits, the first five digits would be a portion of the hashed query fragment. The parties appear to agree on this point.

[4] For example, in *Engel Industries, Inc. v. Lockformer Co.*, the Federal Circuit construed "return portion" and "second portion"to be separate structures where the "return portion" was claimed to be "defined by" the second portion and the separate and distinct return portion was specifically relied upon to distinguish the claimed invention over prior art during prosecution. 96 F.3d 1398, 1404-05 (Fed. Cir. 1996). Defendant also relies on *Bicon, Inc. v. Straumann Co.*, but in that case the claimed structures were described in explicit detail in reference to an "abutment" recited and described in the preamble and that patentee was seeking to avoid those limitations explicitly in the claim. 441 F.3d 945, 949-50 (Fed. Cir. 2006). In contrast, here Defendant is attempting to read in additional limitations not made explicit in the specification or the claims.

fragment to be "divided" or "separated" into a first and second portion, and this requirement is not explicitly stated anywhere in the patent or prosecution history.

Defendant's sole support in the intrinsic record is a preferred embodiment where the first portion and second portion are separate and distinct. However, importing this limitation from the specification to the claims would severely limit the invention in ways not anticipated or intended by the inventor. For example, in the 37 bit system discussed in the patent where the first five digits are used to send the hashed query fragment to a query node, the system must have exactly 32 query nodes to function properly. If the system has more than 32 query nodes, then some nodes will be without an address. If the system has fewer than 32 query nodes then, on at least one node, some distinct query fragments will have the same last 32 digits and return incorrect data. In contrast, if some overlap is allowed then the system can grow dynamically as capacity changes. Without overlap, every increase in capacity would require the number of query nodes to double. The intrinsic record offers no evidence that the inventor contemplated, much less intended such a limitation.

Absent support from the intrinsic record or the language of the claims, requiring the "first portion" of the hashed query fragment to be distinct and separate from the "second portion" would be improper. At the claim construction hearing, Plaintiffs conceded that the first and second portions could not be coextensive. Thus, according to Plaintiffs, the first and second portions must be distinct but need not be separate. The Court agrees. Overlap may occur but at least one portion must contain at least one bit not contained in the other portion. Additionally, both the first portion and the second portion must be distinct from the whole hashed query fragment. Otherwise, the Court would fail to give effect to the word "portion" in the claim. The Court construes "a first portion and a second portion" to mean "a first portion and a second portion, the first portion distinct from the second

portion. The portions may overlap, but neither portion may encompass the entire hashed query fragment."

D. "hashing"

- Plaintiffs' proposed construction: "a computer technique whereby one or more functions are used to transform values into corresponding values"
- Defendant's proposed construction: "performing a mathematical function on a key value to generate the address of the location of data associated with the key value"

The term "hashing" appears in claims 1, 8 (as "hashes"), and 13 of the '593 patent. Plaintiffs seek a broad construction not limited to the use of hash tables to store information while Defendant seeks to narrow the term to require several of the features of hash tables.

Because the remainder of the claim describes how the hashed value is used, Defendant's proposed definition fails to define the term in the context of the claim. For example, the term "key value" is not used in the patent and is an artifact of a dictionary definition. Additionally, Defendant's definition requires the result of a hash to be the location of data, but hashing is also used in other contexts, such as checksums. While in this case the hash is ultimately used to locate data, it is not necessarily the case that the hashed values are the addresses of the location of the data. The claims describe how the hashed values are used to locate data and limiting the claims beyond that would be improper.

Plaintiffs' proposed construction, on the other hand, is too vague. Their own dictionary definition is more specific. Any function may "transform values into associated values," and not all functions are hash functions. For example, a function that multiplied inputs by a constant would transform values into associated values, but multiplying inputs by a constant would not, to one of

skill in the art, constitute hashing.

Based on the dictionary definitions submitted by the parties, the Court concludes that hashing is the process by which an input, often of variable size and complexity, is mapped into a compact space or range. For example, the modulus function could be used as a hash function on integer inputs and the entire set of integers could be mapped into a compact space determined by the chosen modulus. The courts employ such a technique when referring to patents by the last three digits of their patent number, mapping all patents to a set of cardinality 1000 using *modulo* 1000. In the embodiment described in the specification, the "query fragment" input, which could be any size, is mapped onto the range of integers $[0, 2^{37}-1]$.

In light of the above, the Court construes "hashing" to mean "applying a hash function to" and "hashes" to mean "applies a hash function to." Further, the court defines "hash function" to mean "a mathematical function that converts inputs into a value within a predetermined range." Although the dictionary definitions submitted by the parties suggest that these inputs are often "naturally occurring, diverse, ill-structured, and/or scattered data," the Court declines to so limit the term here. A hash function remains a hash function when applied to organized data.

E. "local hash table"

- Plaintiffs' proposed construction: "a table that associates hash values with other data"
- Defendant's proposed construction: "a table resident on and unique to a particular query node in which the unique location of the information in the table is determined by hashing a key value"

This term appears in asserted claims 1, 8, and 13. The parties' principal disagreements over this term relate to whether the table must be "unique" to each query node and the apparent

requirement in Defendant's definition that the query nodes perform hashing.

Defendant argues that, because the term "local hash table" appears within the phrase "local hash table located on said query node," "local" cannot simply mean "located on said query node." Defendant's leap from that position to requiring a "unique" table on each query node appears to lack support, however. Additionally, Defendant's attempt to read its definition of hash or hashing into this term does little to clarify and would potentially confuse the jury into believing that a second hash function must be executed at the query node. This is unsupported by the intrinsic record and in its briefing Defendant clarified that it had not intended to require a second hashing.

Plaintiffs' construction would be acceptable except it appears to read the term "local" out of the claim. While the tables need not be "unique," "local" must have some meaning. Accordingly, the Court construes "hash table" to mean "a table that associates hash values with other data" and "local hash table" to mean "a hash table containing at least the hash values and data associated with said first portion of said hashed query fragment." Redundancy within the system or between systems is not precluded by this construction. The Court does not read into this term a requirement that each first portion of the hashed query fragment must identify one and only one query node.

F. "non-relational, distributed database system"

- Plaintiffs' proposed construction: "a database not using a relational model that is distributed among a plurality of interconnected computer nodes"
- Defendant's proposed construction: "a database, stored across multiple computers on a network, wherein data objects exist independently of their attribute values and wherein data is not extracted using relational algebra"

This term appears in asserted claims 1, 8, and 13 of the '593 patent. The parties agree that

a distributed database system is stored across multiple computers on a network or distributed among a plurality of interconnected computer nodes, which the Court interprets to have equivalent meaning. The parties disagree on how to construe the "non-relational" limitation in this term.

Plaintiffs ask the Court to adopt a construction that essentially restates "non-relational" using more words. Although a "non-relational database" is undoubtedly "a database not using a relational model," this construction adds no clarity without first defining "relational model." The patentee did explain what features are present in a database using a relational model, but never offered a concise definition of the term.

Defendant seeks to import two of the characteristics the patentee used to describe relational model databases as negative limitations. However, such an application of the prosecution history would be erroneous. The patentee did not clearly define "non-relational database" in the prosecution history. Instead, he explained relational databases generally and made specific distinctions between the claimed invention and the prior art in traversing the examiner's rejections.

In the office actions cited by Defendant, the patentee does not rely on the distinctions that Defendant now seeks to import as negative limitations. For example, the patentee explained that the "fragments" described in the cited references were the distinct, non-overlapping pieces of the distributed database and that the "fragment queries" were queries, written in the relational algebra of the database, that uniquely identified one and only one record from within the database fragment. In contrast, the "query fragments" of the claims were potentially overlapping portions of the query to the search engine of the invention, and each query may return multiple information objects. The patentee explained that fragment queries are also portions of larger queries in a relational database, but that they are non-overlapping and are derived by factoring the query using the relational algebra

of the database. Additionally, the patentee distinguished his invention from the cited references based on the invention's use of hashing. Neither reference cited by the patent examiner used hashing and, as explained by the patentee, hashing a query or a fragment query would be antithetical to the function of these software structures within a relational database. Hashing would destroy the relational algebra of the query and defeat the organizational scheme of the database. Further, the patentee explained that, in the relational databases of the prior art, the volume of information stored and indexed by the database is limited by the relational structure of the database. In the patentee's example, an entry may be limited in character length of 40 for name or 60 for address. In contrast, the claimed invention has no limit beyond the capability of hardware for the size of any one information object stored in the database. Additionally, the patentee distinguished his invention over the relational databases in the prior art by explaining that the relational model explicitly rejects object identity while the claims require it.

The Court, reading the prosecution history and the claims, concludes that the differences relied upon in distinguishing between the non-relational database of the claims and the relational databases of the Chaturvedi and Houtsma references are already in the claims, either explicitly or implicitly. Because both parties seek construction of this term, however, the Court will provide one. The Court first adopts Plaintiffs' construction of "a database not using a relational model." To further clarify, the Court construes "relational model" to mean "a database model wherein the data stored in the database is accessed by algebraic queries that uniquely identify each data entry sought, information stored across multiple computers is stored in non-overlapping sub-databases called fragments, queries to the entire database are factored into non-overlapping sub-queries using the database's relational algebra and sub-queries are then sent to fragments, the queries and sub-queries

are not hashed, and the data stored in the database lacks object identity." A database using a model that does not comply with all of these limitations is a non-relational database.

G. "a plurality of home nodes and a plurality of query nodes connected by a network"

- Plaintiffs' proposed construction: the claim language has its plain and ordinary meaning; no further construction is necessary
- Defendant's proposed cosntruction: "a plurality of home nodes and query nodes connected by a network arranged with no central server and wherein, for any given query, any node may be defined as a home node or a query node"

This term appears in asserted claims 1, 8, and 13 of the '593 patent. Plaintiffs seek plain meaning while Defendant seeks to import limitations based on the prosecution history. The Court finds that Defendant's arguments lack merit, and that the term should be given its ordinary meaning.

Defendant seeks to impose the limitation that every node must be capable of being a query node or a home node for any query. To support this argument, Defendant cites portions of the summary of the invention and the detailed description that state that the node that receives the query becomes the home node of the search engine for that query. 2:3-8; 3:21-24. In the embodiment disclosed in the specification, one node is chosen as a home node while other nodes are query nodes. 2:66-3:50; Figure 1. Importantly, the hashed query fragments generated by the home node are distributed as probes to the other nodes and not to the home node. 3:46-50. The home node cannot, then, also be a query node because it reserves no hashed query fragments to itself for searching a local hash table, as required by the claims. Additionally, claims 8 and 13 recite the plurality of home nodes and the plurality of query nodes as separate elements within the body of the claim. Requiring

the plurality of home nodes and the plurality of query nodes to be the same plurality of nodes would contradict the plain language of the claims. However, a system wherein the same plurality of nodes could serve as both home and query nodes would also meet these claim limitations.

Defendant's argument for the negative limitation "with no central server" draws support from the prosecution history where the patentee distinguished his claims from the Chaturvedi reference. In Chaturvedi, there is a single central server node and a plurality of user nodes. Search queries are entered on the user nodes and sent to the central server node. The patentee contrasted this against the system described and claimed in the patent, where queries may enter through a "front end" computer that randomly distributes the queries across a plurality of home nodes which in turn fragment and hash the queries. The patentee's comments about a central server and hierarchical structure were directed at the Chaturvedi reference and cannot be read as a disclaimer of any kind of central computer (such as the "front end" computer described in the patent) or any kind of structure (such as the two classes of nodes recited in the claims).

Accordingly, the Court declines to construe "a plurality of home nodes and a plurality of query nodes." The plain meaning of the term is sufficiently clear and the patentee never clearly and unambiguously adopted a different construction during prosecution.

H. "transmitting, by said selected home node, each said hashed query fragment of said plurality of query fragments to a respective one of said plurality of query nodes indicated by said first portion of each said hashed query fragment"

- Plaintiffs' proposed construction: the claim language has its plain and ordinary meaning; no further construction is necessary.
- Defendant's proposed construction: "the selected home node sends each hashed query fragment to exactly one node on the network, that node being identified by said

first portion of the hashed query fragment"

This term appears[5] in asserted claims 1, 8, and 13 of the '593 patent. Plaintiffs seek only the plain meaning of the term and Defendant seeks to limit the possible architecture of the claimed network. The Court concludes that the prosecution history relied upon by Defendant does not support its proposed definition, and that the term is clear and does not require construction.

Defendant relies upon the May 4, 1997 response submitted by the patentee following the March 21, 1997 office action. In the response, the patentee distinguishes over the Kuechler reference. According to the patentee, Kuechler teaches a distributed database technique where the whole query is distributed to each processing node and each processing node deciphers the whole query to determine whether the relevant records are stored on that node. The patentee also argues that Kuechler teaches a single, fixed home node and uses the relational model. The patentee did not argue that his claimed invention requires that each hashed query fragment be sent to one and only one query node, as Defendant now seeks to limit the claims. The patentee did not even discuss the hashed query fragments of the claims in the office action cited by Defendant, but instead described the features of Kuechler. The office action cited by Defendant contains no clear, unambiguous disavowal of claim scope that would require Defendant's proposed construction. The claims at issue use an open transition, so reading the term itself as a negative limitation on other activity not recited in the claim would be improper without support in the intrinsic record. Accordingly, the Court gives this term is plain and ordinary meaning.

---

[5] This term appears in slightly different form in claims 1, 8, and 13. The differences between the analogous terms in these claims were not deemed to be relevant to claim construction by the parties, and at this time the Court agrees.

I. "using, by said query node, said second portion of said respective hashed query fragment to access data according to a local hash table located on said query node"

- Plaintiffs' proposed construction: the claim language has its plain and ordinary meaning; no further construction is necessary.
- Defendant's proposed construction: "each query node receiving a hashed query fragment uses the second portion of the hashed query fragment as a key value to identify the address of data according to a local hash table stored on that query node"

This term appears in asserted claims 1, 8, and 13 of the '593 patent. Plaintiffs assert that the term requires no construction and plain meaning is sufficient. Defendant asserts that this term should be construed to require that every node that receives a hashed query fragment uses the second portion of the hashed query fragment to access a local hash table. To support this construction, Defendant claims that a portion of the summary of the invention should be read into this claim. The Court disagrees.

Defendant argues that a portion of the summary of the invention stating that "each node on the network which receives a hashed query fragment uses the fragment of the query to perform a search" should be read into the claims. 2:12-13. To support this, Defendant argues that no embodiments in the specification allow for a node to receive a hashed query fragment but to not perform a search. However, the Court's reading of Figure 1 and the related description yields an embodiment where a hashed query fragment may be passed through one node to reach another, which would be excluded by Defendant's construction. Additionally, the patent makes little disclosure about the necessary network architecture. Common network architectures at the time of

the invention, such as Ethernet, transmitted all packets to each node on a hub. In such an arrangement, each node would read the packets it received and determine which packets were addressed and intended for that node. As a result, one of ordinary skill in the art would understand that, where the summary refers to transmitting a hashed query fragment to a node or having each receiving node perform some action, all nodes may in fact receive the hashed query fragment, but only the nodes with the appropriate address information would perform the required action. The claim language makes clear that this is what was intended. This term makes reference to a prior term and requires that the node indicated by the first portion of the hashed query fragment access data according to its local hash table. Additionally, the patentee chose to use an open transition in these claims. Importing negative limitations that restrict the scope of the claims to exclude embodiments that perform additional functions would be improper.

Accordingly, the Court declines to construe this term and will apply its plain and ordinary meaning.

J. "returning, by each said query node"

- Plaintiffs' proposed construction: the claim language has its plain and ordinary meaning; no further construction is necessary.
- Defendant's proposed construction: "each query node that accesses data returns an object identifier to the home node"

This term appears in asserted claims 1, 8, and 13 of the '593 patent. Plaintiffs seek no construction while Defendant seeks to import limitations from elsewhere in the claims and the summary of the invention. As discussed previously, the patentee chose to claim using an open transition and crafted this limitation to apply to the query node indicated by the first portion of the

hashed query fragment. As long as that node accesses and returns data as required by the claims, the system meets those limitations. If other nodes also access data, they are not required to return an object identifier for the system to meet this limitation. The plain language of the "returning" limitations in claims 1, 8, and 13 is sufficiently clear without construction. Accordingly, the Court declines to construe this term.

K. "predetermined degree of relevance"

- Plaintiffs' proposed construction: "a degree of relevance that is determined before returning accessed data to the user"
- Defendant's proposed construction: "a predefined degree of similarity; only results meeting or exceeding a predetermined level are returned to the user after the object identifier has been returned"

This term appears in asserted claims 3 and 9 of the '593 patent. Though not clear from their proposed terms, the parties' principal disagreement over this term focuses on the word "predetermined." Plaintiffs seek a construction that would allow the threshold degree of relevance to be chosen after the query is entered but before the search result is returned while Defendant seeks a construction that would require the threshold degree of relevance to be determined prior to receiving the query. The parties appear to agree that the threshold in question may be a number of results or a score as described in the specification. 4:2-7.[6] The Court declines to construe the term as a whole, but instead will construe the term "predetermined."

"Predetermined" appears in claims 3, 9, 13, 14, and 17 of the '593 patent. Defendant argues

[6] The Court does not read the specification or prosecution history to limit the claims to specific measures of relevance, and reads those in the specification as illustrative examples.

that predetermined means determined prior to receipt of the query while Plaintiffs argue that predetermined means determined prior to returning a result in response to a query. The Court agrees with Defendant. Plaintiffs' construction would impose no limitation in the context of claims 3 and 9, and Defendant's construction makes sense in the context of claims 13[7], 14, and 17. The Court does note, however, that the degree or measure of relevance could be predetermined by some prior or concurrent interaction with the user. For example, Internet searches often have tens of thousands of potentially relevant results, but only a handful are presented to the user at any given time. In many systems, the user can change how many are presented, and that information may reach the search provider at the same time as the query. In this example, the degree of relevance is determined by the user selecting the number of results to be displayed, which occurs before the search engine receives the query, and not by the command that communicates the user's selection, which may occur concurrently with the query. In that context, the Court construes "predetermined" to mean "determined prior to the receipt of the query by the home node."

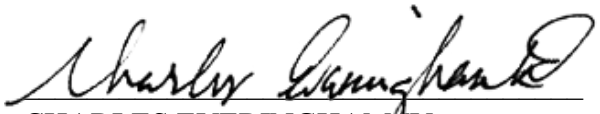
VI. Conclusion

The court adopts the constructions set forth in this opinion for the disputed terms of the '593 patent. The parties are ordered that they may not refer, directly or indirectly, to each other's claim construction positions in the presence of the jury. Likewise, the parties are ordered to refrain from mentioning any portion of this opinion, other than the actual definitions adopted by the court, in the presence of the jury. Any reference to claim construction proceedings is limited to informing the jury of the definitions adopted by the court.

---

[7] "Predetermined" in claims 13, 14, and 17 refers to either a predetermined response to a request or predetermined data sent in response to a query request.

SIGNED this 9th day of November, 2010.



CHARLES EVERINGHAM IV
UNITED STATES MAGISTRATE JUDGE