IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

XEROX CORPORATION, Plaintiff, v. GOOGLE INC., YAHOO! INC., RIGHT MEDIA INC., RIGHT MEDIA LLC, YOUTUBE, INC. and YOUTUBE, LLC, Defendants.	)	C.A. No. 10-136-LPS **PUBLIC VERSION**

DEFENDANTS' RESPONSE BRIEF ON CLAIM CONSTRUCTION

OF COUNSEL:

Charles K. Verhoeven
David A. Perlson
Brian C. Howard
QUINN EMANUEL URQUHART
& SULLIVAN, LLP
50 California Street, 22nd Floor
San Francisco, CA 94111
Tel.: (415) 875-6600
Google-Xerox@quinnemanuel.com

Andrea Pallios Roberts
QUINN EMANUEL URQUHART
& SULLIVAN, LLP
555 Twin Dolphin Dr., 5th Floor
Redwood Shores, California 94065
Tel.: (650) 801-5000
Google-Xerox@quinnemanuel.com

Richard L. Horwitz (#2246)
David E. Moore (#3983)
POTTER ANDERSON & CORROON LLP
Hercules Plaza 6th Floor
1313 N. Market Street
Wilmington, DE 19899
Tel: (302) 984-6000
rhorwitz@potteranderson.com
dmoore@potteranderson.com

Attorneys for Defendant Google Inc.

OF COUNSEL:

Matthew B. Lehr (#2370)
Anthony I. Fenwick
David J. Lisson
Jeremy Brodsky
DAVIS POLK & WARDWELL LLP
1600 El Camino Real
Menlo Park, CA 94025
650-752-2000

Jack B. Blumenfeld (#1014)
Maryellen Noreika (#3208)
Jeremy A. Tigan (#5239)
MORRIS, NICHOLS, ARSHT & TUNNELL LLP
1201 North Market Street
P.O. Box 1347
Wilmington, DE 19899
(302) 658-9200
jblumenfeld@mnat.com
mnoreika@mnat.com
jtigan@mnat.com

Attorneys for Defendants / Counterclaim Plaintiffs
Yahoo! Inc. and Right Media LLC

Dated: April 26, 2011
PUBLIC VERSION DATED: May 3, 2011
10114071/35374

TABLE OF CONTENTS

TABLE OF AUTHORITIES

CASES **PAGES**

OTHER AUTHORITIES

NOTE ON CITATIONS AND APPENDICES

1. References to the patent-in-suit U.S. Patent. No. 6,778,979 (the "979 patent") are indicated by column and line number, or by claim number. A reference to "3:15" therefore means column 3, line 15 of the patent.
2. Relevant excerpts of transcripts of the depositions of Gregory Grefenstette and James Shanahan, the named inventors of the '979 patent, are attached hereto as Exhibits E and F, respectively, and are indicated by "Grefenstette Dep." and "Shanahan Dep." Relevant exhibits from the deposition of James Shanahan are attached hereto as S-4 through S-8. All other exhibits referred to in the brief are attached hereto and indicated by "Ex." followed by the appropriate letter.
3. Citations to Defendants' Opening Brief on Claim Construction shall appear as "DOB," and then the page cite.
4. Citations to Plaintiff's Opening Brief on Claim Construction shall appear as "XOB," and then the page cite.

Introduction

There is no mystery in this case as to which claim construction dispute matters most to Xerox's infringement theories across the full range of accused products. The "restrict a search ... to the category" limitation is pivotal to the case, just as it was pivotal to the Patent Office's allowance of the '979 patent in the first place. Defendants candidly emphasized this point to the Court weeks ago[1] and elicited no disagreement from Xerox.

It is surprising, then, that Xerox buries this issue in the back half of its opening brief, addressing it only superficially in just a few paragraphs, and with no mention at all of the limitation's prominence in the prosecution history. It is as though Xerox seeks to minimize or avoid the issue, in much the same way that Xerox's proposed construction of the limitation conspicuously seeks to minimize and avoid the limitation's actual language ("restrict a search ... to the category"). Xerox's positions on a number of the claim construction issues before the Court reflect a similar aversion to the actual language of the claims and other important intrinsic evidence. Even as to the limitation on which Xerox spends the most time—"selected document content"—it ignores the central problem with that limitation Defendants had previously identified to Xerox; namely, what it actually means for document content to be "selected."

As described below, the depositions of the named inventors, taken after opening briefs were submitted, have only bolstered Defendants' positions, which should be adopted by the Court in their entirety.

---

[1] *See* February 22, 2011, Discovery Conference Transcript (Ex. D) at 22:24-23:10 ("the crux of many of the issues that are going to be brought before your Honor in this case going forward have to do with this notion in that limitation of restricting a search to the category of information that is identified by the assigned classification label, restricting a search to the category, and there is just nothing in these contentions that tells us what Xerox's theory is as to how our products are doing that. And we ask ourselves, why are we here? How can these patents possibly be asserted against these products? That question of that limitation is front and center.")

Argument

I. "TO RESTRICT A SEARCH . . . TO THE CATEGORY OF INFORMATION . . . IDENTIFIED BY THE ASSIGNED CLASSIFICATION LABEL." (1: (D); 18: (F))

Term	*Defendants' Construction*	*Xerox's Construction*
to restrict a search at the information retrieval system for information concerning the set of entities to the category of information in the information retrieval system identified by the assigned classification label	to confine a search at the information retrieval system to the category of information identified by the assigned classification label, where the search seeks information concerning the set of entities.	the set of data specifying search criteria includes data items corresponding to one or more entities identified in the "automatically identifying" step and one or more classification labels assigned in the "automatically categorizing" step.

The crux of the parties' dispute here is simple: Defendants' construction recognizes that a search is restricted to a category, whereas Plaintiff's does not. Xerox tries to wish away the critical aspect of the limitation – namely, the requirement that the "query" of the claim will cause a search at the information retrieval system to be restricted to the category of information identified by the assigned classification label. What the patent claims is a system where a query is generated for a search that will be restricted to—i.e. confined within—a particular defined subset of the overall body of information at the information retrieval system. Plaintiff, instead, attempts to rewrite the claims such that the search resulting from the query could be performed across even the entire body of information in the information retrieval system, contrary to Xerox's statements during prosecution. Defendants' construction stays true to the claim language, comports with the intrinsic evidence, and is supported by noncontroversial dictionary definitions and the testimony of the named inventors.

A. Xerox Fails to Justify its Construction That Is Contrary to the Plain Language of the Claims.

Xerox does not actually dispute that the plain meaning of "restrict a search . . . to the category" is "to confine a search . . . to the category," as Defendants' construction provides. Nor

does Xerox provide any evidence or argument that the inventors were their own lexicographer such that this ordinary meaning should not apply.

Instead, Xerox argues that its construction somehow comports with the plain meaning of the claims based on the following proposition found at page 15 of its brief:

> Because Claim 1 also requires that each classification label 'correspond[] to a category of information in an information retrieval system,' [citation omitted] the inclusion in the query of data corresponding to the classification label(s) will restrict a search at the information retrieval system . . . to the categories of information corresponding to the assigned classification labels.

(XOB, 15.) Beyond this conclusory statement and saying its construction is "easily comprehensible" or "easily understood," Xerox offers no support whatsoever for the proposition that merely including data corresponding to a classification label in a query will cause the resulting search to be restricted to the category of information identified by that classification label.

In fact, it is easy to think of queries where the mere inclusion of an assigned classification label would not result in the required restriction to the category of documents associated with the assigned label. For example, a query including both a data item for the set of entities and a data item for an assigned classification label could be framed as "seven-up OR genetics." But as a matter common sense, this search is not "restricted" to "genetics," as "genetics" is an alternative to "seven-up" in the search. Such a query would look for and produce results relating both to genetics and seven-up (i.e. the seven-up gene) and seven-up (i.e. the soft drink). Thus, the search would not be restricted to the category of the assigned classification label, "genetics" in this example. Yet, such a result seems to fall squarely within Xerox's proposed construction.

In actuality, simply including a "data item" that "corresponds" in some way to a category of information in the information retrieval system, as Xerox's construction provides, says nothing as to what effect, if any, that inclusion and correspondence may have on the search. It says nothing about what specific search criteria, if any, the query actually imposes in relation to the included

classification label. It does not articulate what other criteria and terms are, or are not, included in the query. And it says nothing regarding how the information system is to process queries that simply "include" a classification label.

Xerox seeks to justify its construction's failure to speak to the query's impact on a search at the information retrieval system by arguing that the claims only specify "what the query contains," not "how [the query] interacts with the information retrieval system." (*Id.* at 13 (emphasis in original), 14 ("the asserted claims are directed to formulating a query . . . The remainder of the Step (d) simply specifies what that set of query data must contain in terms of search criteria.").) But, there is language in the asserted claims addressing how the query is to interact with the information retrieval system. The claims require what the search will be for – "information concerning the set of entities." It also requires where the search will be performed. Not only must the search be performed "at the information retrieval system," but it must be restricted to—i.e. in, and only in—"the category of information in the information retrieval system identified by the assigned classification label." It is this very language on which Xerox specifically relied to distinguish its claimed invention from the prior art. (*See* DOB, 10). Yet, Xerox now seeks to improperly strip this language out of the claims entirely through a construction that does not in any way connect to the actual claim language, or have practical meaning in the field of information retrieval.

Finally, not only does Xerox ignore the impact of the query at the information retrieval system required by the claim, it also ignores any distinction between how the "set of entities" and the "[assigned] classification label" are to be used in the search called for in the claim. Xerox's construction treats them the same, requiring only that "corresponding data items" be "include[d]" for both, such that it would cover a keyword search for "seven-up AND genetics." It is as though, rather than construe the actual language, Xerox is construing a hypothetical limitation like the following:

automatically formulating the query to ~~restrict a search at the information retrieval system for information~~ concerning the set of entities ~~to the category of information in the information retrieval system identified by~~ and the assigned classification label.

But the search in the claims is not a search for information concerning the set of entities and the category. The search "concern[s]" the set of entities, but is "restrict[ed] to" the category. Xerox's attempt to rewrite the claims in this dramatic fashion is plainly improper.

B. The Inventors' Testimony and Corroborating Documentation Supports Defendants' Construction.

The named inventors [redacted] – were deposed during the week following the parties' submission of initial claim construction briefs. Their testimony supports Defendants' construction of the "restrict a search ... to the category" limitation. *See Voice Tech. Group, Inc. v. VMC Sys., Inc.*, 164 F.3d 605, 615 (Fed. Cir. 1999) (an inventor is "a competent witness to explain the invention and what was intended to be conveyed by the specification and covered by the claims").

1. [redacted]

Xerox's brief treats the query of Figure 39 (shown below) as though it merely adds the terms "science" and "biology" and "genetics" ("science+biology+genetics") as keywords to be searched for in addition to the phrase "seven up." Xerox proposes that the mere inclusion of these terms in the query in such a manner is sufficient to provide the "restriction" referenced in the claims. (*See, e.g.*, XOB, 1-2 and 15.)

QUERY WITH CONTEXT: Science + Biology + Genetics
(e.g., http://www.google.com/search?q=seven+up&cat=gwd%2Ftop%2FScience%2FBiology%2FGenetics) — 3930

FIG. 39

[redacted]

[redacted] Plainly, the query of Figure 39 goes well beyond the vague requirements of Xerox's proposed construction. Instead, it directs a search to a particular identified category of information within an information retrieval system, not the entire body of information in the information retrieval system as Xerox's construction seems to allow.

[redacted]

[redacted].[3] (Ex. H.) For example, in the section entitled "Category Search vs. Web Search," Google explained: "[w]hen you enter a search in the Google directory engine, only the category you are currently in will be searched." (*Id.* (emphasis added).) In the section: "[w]hen would I use the Google directory instead of Google's regular web search?," the FAQ also states, that:

> "Google's directory engine also lets you search within a category once you've decided on the specific subsection of the web that interests you. In this way, you'll get only responses that fall **within** that category. For example, you may want to search for teams named 'Cougars' within the college basketball section of the directory only, instead of across the entire web." (*Id.*) (emphasis added)).

Similarly, Google's FAQ explained that "a search over the entire web for 'lions' might return pages about lions (the animal), Lions (the football team), Lions (the public service organization), or any number of other subjects. By searching for 'Lions' within the category 'Sports > Football, American > Professional >', you will see only results related to the Detroit Lions football team." (*Id.* (emphasis added).)

[redacted]

[3] http://replay.waybackmachine.org/200012020034/http:/directory.google.com/dirhelp.html

[redacted]

C. Xerox Fails to Demonstrate That "The Category of Information in the Information Retrieval System" Can Be "One or More Classification Labels".

The claims require the search be restricted "to the category [singular] of information in the information retrieval system identified by the assigned classification label." Xerox's construction, however, provides that the search criteria include "one or more classification labels assigned in the 'automatically categorizing' step." Xerox provides no basis for this "one or more" language. When

[redacted]

discussing the categorizing step, Xerox points to the default rule that "a" means "one or more" and that this rule applies to subsequent uses of the term introduced by "the." (XOB, 9.) But that default rule does not apply to "the category" in step 1(d). While there is "a classification label" in step 1(b), the search in step 1(d) is not restricted to "the classification label." Instead, it is restricted to something else —"the category in the information retrieval system . . ." There is no preceding reference to "a category" that could justify invocation of the default rule to support a theory that "the category" actually means "one or more categories." Xerox's proposed construction should be rejected for what it is: an improper attempt to rewrite the claims.

II. "CATEGORIZING THE SELECTED DOCUMENT CONTENT . . . FOR ASSIGNING . . . A CLASSIFICATION LABEL." (1(C); 18(E))

Term	*Defendants' Construction*	*Xerox's Construction*
categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label.	using the organized classification of document content to categorize the selected document content and to assign to the selected document content a single classification label.	determining the subject matter of the selected document content using one or more of the categories defining the organized classification of document content and assigning the corresponding classification label(s) to the selected document content.

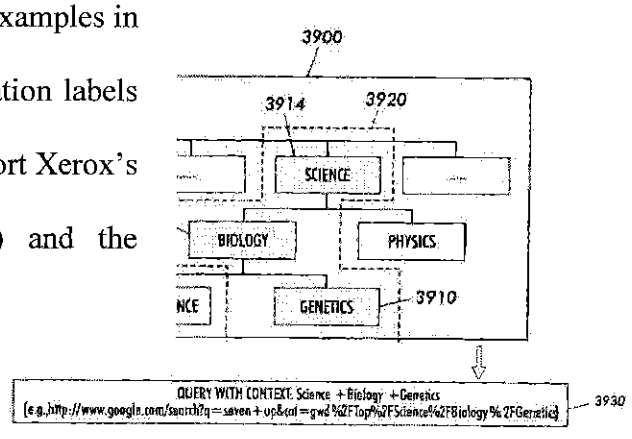
Rather than address the language of the claims in defending its construction, Xerox points to a default presumption and mischaracterizes the specification to argue that multiple classification labels may be assigned to the document content. Yet, even Xerox admits that the claims require that "the formulated query restricts the search 'to *the* category of information in the information retrieval system identified by *the* assigned classification label.'" (XOB, 15 (emphasis added).) To be clear, the issue with this limitation is not whether multiple classification labels are "used" or "identified" in the categorization process. Defendants' proposed construction fully allows for such an approach. Instead, the issue is whether, as Defendants' construction requires, only a single label is ultimately "assigned" for purposes of restricting the search to "the category" identified by "the assigned

classification label," or whether, as Xerox contends, multiple labels may be assigned. Only Defendants' construction can be reconciled with the plain meaning of this phrase.

The very case law cited by Xerox acknowledges that the default presumption that "a" means "more than one" does not apply "where the language of the claims themselves, the specification, or the prosecution history necessitate a departure from the rule." (XOB, 9 (quoting *Baldwin Graphic Sys., Inc. v. Siebert, Inc.*, 512 F.3d 1338, 1342-43 (Fed. Cir. 2008).) As explained in Defendants' Brief, the claim language makes sense only if a single classification label is ultimately assigned in the "categorizing" step. (DOB, 11-12.) That step culminates in the assignment of the classification label used in the "formulating" step to identify "the category" –i.e. the single assigned category – to which a search is to be restricted. Tellingly, Xerox does not address this claim language.

Further, while a document containing "seven up" could theoretically be assigned the "science/biology/genetics" label and also the labels "beverages," "soft drinks," and "advertisements," assigning these classification labels would not "improve the quality (e.g. in terms of precision recall) of information retrieval systems," as the patent purports to do. (XOB, 1 (quoting 48:37-39).) Yet, Xerox's construction allows an unlimited number of labels to be assigned and included in a query, which would wholly undermine this supposed goal of "precision recall."

Xerox does correctly state there are several examples in the specification of multiple categories or classification labels being "used" or "identified." But this does not support Xerox's construction. As Figure 39 (shown on right) and the corresponding text of the patent contemplate, and as Defendants' construction allows, a document may relate to nested categories and subcategories, with the result that a query is "contextualized using"



multiple "classification labels." (50:1-2.) In this sense, the "categorizer" of the patent uses document content to "identif[y]" multiple relevant "classification labels." (50:4.) As the specification makes clear, however, document content is ultimately "assigned" a single classification label after a determination regarding which one is most appropriate: "in this example, the entities 'seven' and 'up' are determined by categorizer 3610 to relate most appropriately to the class of documents found in the directory science>biology>genetics." (50:6-9.) The specification goes on to explain, "the search is focused on documents found in the single node of the document hierarchy genetics." (50:9-11; *see also* 49:22-28 ("given a classification scheme such as a class hierarchy . . . in which documents are assigned class labels (or assigned to nodes in a labeled hierarchy), a classification profile is derived that allows document content to be assigned to an existing label or to an existing class." (emphasis added)); 48:67-49:02 ("Advantageously, the query may be contextualized at different levels: first, the query is set to be directed in a specific category of an information retrieval system that may, for example, be hierarchically organized").)

Finally, Xerox's suggestion that Defendants' construction is inconsistent with the "set of categories" in the agreed-upon meaning of "organized classification of document content" is misplaced. (XOB, 9, 11.) The organized classification scheme consists of a set of multiple categories, and Defendants' construction does not say otherwise. However, the fact that the classification scheme has multiple categories does not mean that multiple categories must be assigned to specific document content. Rather, common sense and the claim language dictate just the opposite — that a document is assigned a single classification label that then directs a search to the single corresponding category in the information retrieval system.

III. QUERY (1: PREAMBLE, 1 (D); 18: (B), (F))

Defendants' Construction	***Xerox's Construction***
Request for search results.	a set of data specifying search criteria.

Xerox claims that the construction of "query" should not include any notion of how the query operates in use. Instead, Xerox argues, "[t]he query in Figure 39 . . . is depicted solely with reference to its contents: the entities identified in document content ('seven up') and the result of categorizing that content ('science/biology/genetics')." (XOB, 13.) As discussed above in § I(A), however, the query in Figure 39 is not depicted solely in reference to its contents, independent of the functional implications of those contents. Rather, the specification and claims describe a query that directs that a search at the information retrieval system be restricted to the specified category. Xerox's construction is yet another attempt to eliminate this requirement.

Xerox also claims that "the only query activity mentioned by the claims is 'generating' or 'formulating' the query," and that "the claims cover what the generated query must contain (i.e., entity data restricted by classification label data). The claim language stops there . . . There is no reference to 'search results' in the claim." (XOB, 11-12 (emphasis added).) But, as Defendants noted in their Opening Brief, claims 11 and 17 discuss ranking "results" from the "query." (DOB, 14.) Since there is no dispute that a query is used to "search," there can be no legitimate dispute that the "results" referred to in the claims are "search" results from the query, just as Defendants' construction provides. (*See* XOB, 13 ("[t]o be sure, the intrinsic evidence indicates that the formulated query will be used in performing a search at an information retrieval system.").)

Finally, Xerox points to Figure 38 and the corresponding portion of the specification (48:41-51), which does not indicate how the query interacts with the information retrieval system. (XOB, 12-13.) This makes sense, however, given that this portion of the specification merely describes what a generated query "may include." It does not seek to define "query." In any event, this excerpt is perfectly consistent with Defendants' construction that a query is a request for search results.

IV. CLASSIFICATION LABEL, 1: (A), (C), (D); 18: (C), (E), (F)

Defendants' Construction	*Xerox's Construction*
classifying word or phrase.	a label in any format that identifies a category in the organized classification of document content.

In the claims, a classification label is associated with "each class in the organized classification of document content." (1(a), 18(c).) As Defendants' construction provides, the plain meaning of a label in this context is "a descriptive, classifying, or identifying word or phrase." *Webster's Third New Int'l Dictionary, Unabridged* (2002); *Random House Unabridged Dictionary* (2d Ed. 1993) ("a word or phrase indicating that what follows belongs in a particular category or classification.")

Nevertheless, Xerox argues that the ordinary meaning of a "label" is "[a]n item used to identify something or someone," citing *The American College Dictionary*. (XOB, 7.) But Xerox omits the remainder of the definition referring to a "paper" or "cloth" to designate origin:

Xerox omits

la•bel (lā′bəl) *n.* **1. An item used to identify something or someone, as a small piece of paper or cloth attached to an article to designate its origin, owner, contents, use, or destination.**

Xerox cites

Obviously, the claims do not concern a paper "label," such as on a Campbell's soup can.

Xerox also contends the meaning of label in computer science is "[a]n identifier within or attached to a set of data elements," citing the IBM Computer Dictionary. (XOB, 7). But Xerox omits that this is just one of the twelve definitions in the cited IBM dictionary. (XOB, Ex. J at 374.) And Xerox does not even try to explain how its chosen definition fits in the context of the claims and specification. As the Federal Circuit has explained, however, "in those circumstances where reference to dictionaries is appropriate, the task is to scrutinize the intrinsic evidence in order to determine the most appropriate definition." *Free Motion Fitness Inc. v. Cybex Intern., Inc.*, 423 F.3d 1343, 1349 (Fed. Cir. 2005); *see also Phillips v. AWH Corp.*, 415 F.3d 1303, 1321 (Fed. Cir. 2005) (en banc) ("[t]he problem is that if the district court starts with the broad dictionary definition in

every case and fails to fully appreciate how the specification implicitly limits that definition, the error will systematically cause the construction of the claim to be unduly expansive.")

Instead of finding support in the specification, Xerox points to the alphanumeric code in the Library of Congress classification system and a number in the Dewey Decimal System as purported support for its construction. (XOB, 7.) In actuality, these systems use words as labels, as is the case in every embodiment in the specification. *See* Library of Congress Classification Outline, *available at* http://www.loc.gov/catdir/cpso/lcco/ (disclosing such class labels as "History of the Americas," "Law," and "Education.") Each label in these library systems does have a corresponding "call number" that can be placed on the spine of the book and/or used for reference. These call numbers, however, are not the actual "labels" of the class—which are words or phrases— but are distinct numbers associated with the descriptive label.

Moreover, even if a label could be something other than a word or phrase, it must still be used for "classification." While Xerox's construction uses the word "classification," it actually omits any requirement that the label classify anything. Instead, Xerox argues that "the claim simply requires that a 'classification label' be associated with, and therefore identify, a descriptive category in the OCDC." (XOB, 8.) Thus, according to Xerox, the "classification label" would not be "Genetics," but would be something else that somehow "identifies" the descriptive category Genetics. This makes no sense and is inconsistent with every disclosure in the specification.[5]

[5] Xerox argues the following excerpt from the specification supports its construction: "classifier accepts as input a document 'Doc' and predicts the target *value* C, or a classification label." (XOB, 8 (emphasis added).) Initially, the word "label" does not actually appear in that portion of the specification. In any event, this portion of the specification does not address what a classification label is in the classification system, but rather concerns how the process of classifying a document can be referred to mathematically in a particular machine learning technique, naïve bayesian, that may be used predict the class for a document. (43:47-51.)

Further, even in the Library of Congress system Xerox points to, "Q" corresponds to Science and "H" corresponds to the Biology subcategory in Science. The "Q" and "H" show what kind of information is in the associated category. The same is true with the Dewey Decimal Number of "576," in which "500" is Science, "570" is Life Science; "576" is Genetics.[6] In contrast, if the "label" can be "4375816" and there is no Dewey-like system that ascribes classification information to different parts of that number—as Xerox's construction seems to allow—then the number does not "classify" the document content. It just identifies it. The claims, however, require a "classification label," not an "identification label."

V. SELECTED DOCUMENT CONTENT (1: PREAMBLE, (B), (C); 18: (B), (D), (E))

Defendants' Construction	*Xerox's Construction*
Indefinite.	all or part of the content of a document in electronic form.

A. Xerox Ignores the Insoluble Ambiguity of "Selected Document Content"

As discussed in Defendants' Brief, the term "selected document content" is insolubly ambiguous – it has no antecedent basis, and there is no way to determine who does the selecting or how the selecting occurs. (DOB, 16-17.) Although Defendants raised both of these issues during the parties' meet and confers regarding claim construction, Xerox's Brief ignores them entirely.

Instead, Xerox argues that the "intrinsic evidence demonstrates that 'selected document content' comprises an input to the claimed method." (XOB, 3; *see also* XOB, 4 ("'selected document content' simply comprises the document content that serves as an input to the claimed method.").) Obviously, document content must be somehow "inputted" for the patented method to take place. But Xerox's "input-based" interpretation does not distinguish between "selected document content" and the "document content," and Xerox's construction does not even actually

[6] *See* "List of Dewey Decimal Classes," *available at* http://en.wikipedia.org/wiki/List_of_Dewey_Decimal_classes#500_.E2.80.93_Science

require the document content to be inputted. The claims clearly do not allow for just any type of document content to be inputted, but rather require that the document content must be "selected." The notion of "inputting" therefore adds nothing to the meaning of "selected document content" – the term remains insolubly ambiguous and incapable of construction.[7]

B. Xerox's Leap from "Document Content" to "All or Part of the Document" Is without Merit.

Even assuming the phrase "selected document content" could be construed, it is not "all or part of the content of a document" as Xerox contends. For one thing, Xerox's construction still does not explain how the "selecting" occurs. In other words, Xerox's construction does nothing to resolve the ambiguity of "selected document content," as discussed above.

Moreover, Xerox's "all or part" construction would mean that "document content" could be a single word, such that a 500 word document could consist of 500 separate "selected document contents." But this makes no sense in the context of how "document content" is used in the patent. For instance, Claim 9 recites "extracting noun phrases" from the selected document content after determining how frequently those noun phrases occurred, and the specification teaches that "only terms appearing in the document content more than a certain number of times will be annotated . . ." (33:62-65; *see* also Fig. 28 ("annotate entity when it appears in document content with a frequency greater than: 4").) If the document content could be a single word – as Xerox's construction allows – it would make no sense to speak of noun phrases occurring more than a certain number of times

[7] Xerox suggests that Defendants should be precluded from arguing indefiniteness because Defendants did not mention indefiniteness in their prior interrogatory responses. (XOB, 3). Xerox's position is hypocritical given that Xerox withheld the very claims it was actually asserting in the case until the day initial claim constructions were due. (D.I. 118.) Xerox also points to no prejudice on this point, nor could it, as Defendants explained the basis for indefiniteness during the parties' meet-and-confer regarding claim construction. Also, at Xerox's request, Defendants have supplemented their interrogatory responses to address the indefiniteness.

within the document content. (*See also* Figures 38 and 39, which show an entire document, not merely part of a document, as "document content.")

Second, while Xerox states that the specification teaches that the "selected document content" can be all or part of the document, its cited passages do not. Xerox cites the following:

> In operation as shown in FIG. 38, the document content 3612 or alternatively limited context (i.e., words, sentences, or paragraphs) surrounding the entity 3808 is analyzed by categorizer 3610 to produce a set of categories 3620. (48:52-55.)

(XOB, 5 (emphasis added).) Yet, this quote actually shows that "document content" is something different than the "limited context" surrounding the entity. Xerox similarly points to the specification's explanation that a query may include "terms relating to context information surrounding the set of entities in the selected document content" and that "[p]roducing an aspect vector contextualizes queries related to the entities by examining a portion of the document content that may range from all of it to one or more paragraphs and/or segments around the entity." (50:21-25.) This language merely indicates that some or all of the entire document may be examined for the purposes of constructing the query. It in no way indicates that "selected" document content can be all or a portion of a document.

Finally, Xerox contends that "selected document content" must be in electronic form. The specification, however, specifically allows for "document content" to be in a hardcopy document: "[r]eferring again to FIG. 6, the hardcopy document includes document content 616 and embedded data 612." (14:60-61 (emphasis added).) Xerox's own inability to ascribe a supportable meaning to the term further shows it to be indefinite.

VI. ORDER OF STEPS

Defendants' Construction	***Plaintiff's Construction***
Claim 1: Step (a) must be performed before steps (c) and (d); Step (b) must be performed before step (d); Step (c) must be	**Claim 1:** Step (a) must be performed before steps (c) and (d); Step (b) must be performed before the completion of step (d); Step (c) must be performed

Defendants' Construction	*Plaintiff's Construction*
performed before step (d). **Claim 18:** Step (c) must be performed before steps (e) and (f); Step (d) must be performed before step (f); Step (e) must be performed before step (f). **Claim 2:** The steps of claim 1 must be performed before the step of 2. **Claim 19:** The steps of claim 18 must be performed before the step of 19	before the completion of step (d). **Claim 18:** Step (c) must be performed before steps (e) and (f); Step (d) must be performed before the completion of step (f); Step (e) must be performed before the completion of step (f). **Claim 2:** The step of Claim 2 must be performed during or after the completion of step (d) of Claim 1. **Claim 19:** The step of Claim 19 must be performed during or after the completion of step (f) of Claim 18.

Xerox does not dispute that method claims must be construed so that the steps are performed in the order recited if logic or the rules of grammar so demand. *Altiris Inc. v. Symantec Corp.*, 318 F.3d 1363, 1369 (Fed. Cir. 2003); *Interactive Gift Express, Inc. v. Compuserve, Inc.*, 256 F.3d 1323, 1342 (Fed. Cir. 2001). This is precisely the case here. Indeed, Xerox acknowledges:

> Step (b) of Claim 1 ("automatically identifying a set of entities") must precede the completion of Step (d) ("automatically formulating the query"), and Step (c) ("automatically categorizing the selected document content") must precede the completion of Step (d). This is necessarily so because, as explained above, the query formulated in Step (d) utilizes the data generated in both Steps (b) and (c).

(XOB, 19.) Xerox also agrees that in claim 1, step (a) must be performed before steps (c) and (d) and that in Claim 18, step (c) must be performed before steps (e) and (f). Where "most of the steps" of a "method claim refer to the completed results of the prior step," the steps must be "performed in order." *E-Pass Techs., Inc. v. 3Com Corp.*, 473 F.3d 1213, 1222 (Fed. Cir. 2007).

Nevertheless, Xerox argues that in independent claims 1 and 18 "the categorization step could find one category at a time, and pass each one to the query formulation step to add the classification label to the query." (XOB, 19-20.) This is just a retread of Xerox's argument that more than one classification label may be assigned in the categorizing step. Contrary to Xerox's argument, however, the categorization step could not assign multiple classification labels in succession because only a single classification label is assigned to each document. (*See supra*, § II.)

Xerox also asserts that "nothing in the claim language or the specification precludes an iterative process whereby, for example, the entity identification step finds an entity, the query formulation step adds that entity to the query, then the entity identification step finds another entity, and so on." (XOB, 19.) But the claim language does preclude it. Step 1(b) requires "identifying a set of entities," which indicates that the entities are identified as a "set," not in a piecemeal, "iterative" fashion. That some of the steps of the tasks could, as a technical matter, conceivably be done concurrently (as Xerox suggests) is irrelevant. That is not the claim Xerox drafted.

Notably, Xerox's own construction of "to restrict a search . . . to the category of information . . . identified by the assigned classification label" is "the set of data specifying search criteria includes data items corresponding to one or more entities identified in the 'automatically identifying' step and one or more classification labels assigned in the 'automatically categorizing' step." (emphasis added). Thus, even under Xerox's own construction, the entities must be already "identified" and the classification label(s) "assigned" before the query is formulated in Step 1(d).

Xerox further argues "nothing in the claim language requires that these 'terms' [of claims 2 and 19] be added to the query as search criteria only after data corresponding to the set of entities and the classification labels is present in the query." But the plain language of claims 2 and 19 note that they "further comprise" the step of "limiting the query." In other words, the query is "formulated" in steps 1(d) and 18(f), and then limited in claims 2 and 19.

Finally, Xerox states "the specification expressly teaches an embodiment in which the entities from Step (b) ('automatically identifying a set of entities') of Claim 1 and 'terms' of Claim 2 are present in the query being formulated before the addition of classification labels derived in Step (c) ('automatically categorizing the selected document content') of Claim 1." (XOB, 19.) Xerox quotes the following portion of the specification: "classification labels in one embodiment are

appended to the query 3812 by query generator 3810 to restrict the scope of the query (i.e., the entity 3808 and the context vector 3822) . . ." (*Id.* (quoting 49:31-35).) Contrary to Xerox's characterization, this quote in no way suggests that the 'terms' of Claim 2 (allegedly embodied in the context vector 3822) are present in the query before the classification labels are appended. Rather, it merely states that all three listed elements – classification label, entity, and context vector – may ultimately be used in the query.

VII. "CHARACTERISTIC VOCABULARY"

Defendants' Construction	*Xerox's Construction*
one or more words or phrases that describe the category of information corresponding to the class.	one or more words or phrases that describe a class in the organized classification of document content.

The specification is clear that a "characteristic vocabulary" is the same as a "category vocabulary" and that that a category vocabulary "consists of one or more terms that describe the category." (49:44-45; 51:36-37.) Xerox argues that the word "category" in the phrase "category vocabulary" does not refer to the categories of information in the information retrieval system, but rather to the "set of classes in the OCDC that are used to categorize document content." (XOB, 17.) This switch from "category" to "class" is inconsistent with the language of the claims and the specification. The intrinsic evidence that Xerox cites shows that its construction is unsupportable. Figure 36 (cited by Xerox) concerns the manner in which documents are assigned a class and a classification vocabulary by the categorizer. Thus, the category vocabulary is not simply a description of the class labels, as Xerox's construction wrongly posits, but is assigned based on actual documents in the category.

Conclusion

For these reasons, and those in Defendants' Opening Brief, the Court should adopt Defendants' constructions of the disputed terms.

POTTER ANDERSON & CORROON LLP

By: */s/ Richard L. Horwitz*
Richard L. Horwitz (#2246)
David E. Moore (#3983)
Hercules Plaza 6th Floor
1313 N. Market Street
Wilmington, DE 19899
Tel: (302) 984-6000
rhorwitz@potteranderson.com
dmoore@potteranderson.com

Attorneys for Defendant Google Inc.

OF COUNSEL:

Charles K. Verhoeven
David A. Perlson
Brian C. Howard
QUINN EMANUEL URQUHART
& SULLIVAN, LLP
50 California Street, 22nd Floor
San Francisco, CA 94111
Tel.: (415) 875-6600
Google-Xerox@quinnemanuel.com

Andrea Pallios Roberts
QUINN EMANUEL URQUHART
& SULLIVAN, LLP
555 Twin Dolphin Dr., 5th Floor
Redwood Shores, California 94065
Tel.: (650) 801-5000
Google-Xerox@quinnemanuel.com

MORRIS, NICHOLS, ARSHT & TUNNELL LLP

By: */s/ Maryellen Noreika*
Jack B. Blumenfeld (#1014)
Maryellen Noreika (#3208)
Jeremy A. Tigan (#5239)
1201 North Market Street
P.O. Box 1347
Wilmington, DE 19899
(302) 658-9200
jblumenfeld@mnat.com
mnoreika@mnat.com
jtigan@mnat.com

Attorneys for Defendants / Counterclaim Plaintiffs Yahoo! Inc. and Right Media LLC

OF COUNSEL:

Matthew B. Lehr (#2370)
Anthony I. Fenwick
David J. Lisson
Jeremy Brodsky
DAVIS POLK & WARDWELL LLP
1600 El Camino Real
Menlo Park, CA 94025
650-752-2000

Dated: April 26, 2011
PUBLIC VERSION DATED: May 3, 2011
1011407/35374

IN THE UNITED STATES DISTRICT COURT
FOR THE DISTRICT OF DELAWARE

CERTIFICATE OF SERVICE

I, Richard L. Horwitz, hereby certify that on May 3, 2011, the attached document was electronically filed with the Clerk of the Court using CM/ECF which will send notification to the registered attorney(s) of record that the document has been filed and is available for viewing and downloading.

I hereby certify that on May 3, 2011, the attached document was electronically mailed to the following person(s)

Lawrence C. Ashby
John G. Day
Lauren E. Maguire
ASHBY & GEDDES
500 Delaware Avenue, 8th Floor
Wilmington, DE 19899
lashby@ashby-geddes.com
jday@ashby-geddes.com
lmaguire@ashby-geddes.com
Attorneys for Plaintiff Xerox Corporation

Richard J. Stark
Andrei Harasymiak
Peter A. Emmi
Scott A. Leslie
CRAVATH, SWAINE & MOORE LLP
Worldwide Plaza
825 Eighth Avenue
New York, NY 10019
rstark@cravath.com
aharasymiak@cravath.com
pemmi@cravath.com
sleslie@cravath.com
Attorneys for Plaintiff Xerox Corporation

Jack B. Blumenfeld
Maryellen Noreika
Jeremy A. Tigan
MORRIS, NICHOLS, ARSHT & TUNNELL LLP
1201 North Market Street
Wilmington, DE 19899
jblumenfeld@mnat.com
mnoreika@mnat.com
jtigan@mnat.com
Attorneys for Defendants Right Media LLC and Yahoo! Inc.

Matthew B. Lehr
Anthony I. Fenwick
David J. Lisson
Jeremy M. Brodsky
DAVIS POLK & WARDELL LLP
1600 El Camino Real
Menlo Park, CA 94025
matthew.lehr@davispolk.com
anthony.fenwick@davispolk.com
david.lisson@davispolk.com
Jeremy.brodsky@davispolk.com
Attorneys for Defendants Right Media LLC and Yahoo! Inc.

/s/ Richard L. Horwitz

Richard L. Horwitz
David E. Moore
POTTER ANDERSON & CORROON LLP
(302) 984-6000
rhorwitz@potteranderson.com
dmoore@potteranderson.com