EXHIBIT A

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

In re Patent of:	Grefenstette
U.S. Patent No.:	6,778,979
Issue Date:	Aug. 17, 2004
Serial No.:	09/683,235
Filing Date:	Dec. 5, 2001
Title:	SYSTEM FOR AUTOMATICALLY GENERATING QUERIES

Mail Stop Inter Partes Reexam
Central Reexamination Unit
Commissioner for Patents
P.O. Box 1450
Alexandria, VA 22313-1450

REQUEST FOR INTER PARTES REEXAMINATION

Reexamination under 35 U.S.C. § 311 and 37 C.F.R. § 1.913 is requested for claims 1-20 of U.S. Patent No. 6,778,979 (the '979 patent) which issued on August 17, 2004 to assignee Xerox Corporation.

TABLE OF CONTENTS

Page

APPENDICES

APPENDIX A – U. S. Patent No. 6,778,979 to Grefenstette et al.

APPENDIX B – Relevant portions of the '979 patent file history

APPENDIX C – PCT Application Publication No. WO 01/44992 to Wieser et al.

APPENDIX D – U.S. Patent No. 6,236,768 to Rhodes et al.

APPENDIX E – U.S. Patent No. 7,089,236 to Stibel

APPENDIX F – "Syskill & Webert: Identifying interesting web sites" by Michael Pazzani, Jack Muramatsu and Daniel Billsus

APPENDIX G – U.S. Patent No. 6,546,386 to Black et al.

APPENDIX H – U.S. Patent No. 7,225,180 to Donaldson

APPENDIX I – U.S. Patent No. 5,748,954 to Mauldin

APPENDIX J – "HyPursuit: A Heirarchical Network Search Engine that Exploits Content-Link Hypertext Clustering" by Ron Weiss, Bienvenido Velez, Mark A. Sheldon, Chanathip Namprempre, Peter Szilagyi, Andrzej Duda and David K. Gifford

APPENDIX K – U.S. Patent Application Publication No. 2002/0147738 to Reader

APPENDIX L – "Experimental Simulation for Automatic Patent Categorization" by Hisuo Mase, Hiroshi Tsuji, Hiroshi Kinukawa, Yoshinori Hosoya, Kazuya Koutani, and Kenichi Kiyota

APPENDIX M – Claim chart showing the correspondence between the Reader reference and the '979 patent.

APPENDIX N – Claim chart showing the correspondence between the Reader and Mase references and the '979 patent.

APPENDIX O – Claim chart showing the correspondence between the Wieser reference and the '979 patent.

APPENDIX P – Claim chart showing the correspondence between the Wieser and Stibel references and the '979 patent.

APPENDIX Q – Claim chart showing the correspondence between the Wieser, Stibel and Syskill references and the '979 patent.

APPENDIX R – Claim chart showing the correspondence between the Rhodes reference and the '979 patent.

APPENDIX S – Claim chart showing the correspondence between the Rhodes and Stibel references and the '979 patent.

APPENDIX T – Claim chart showing the correspondence between the Rhodes and Syskill references and the '979 patent.

APPENDIX U – Claim chart showing the correspondence between the Rhodes, Black, Syskill, and Mauldin references and the '979 patent.

APPENDIX V – Claim chart showing the correspondence between the Black and Hypursuit references and the '979 patent.

APPENDIX W – Claim chart showing the correspondence between the Black, Hypursuit, and Syskill references and the '979 patent.

APPENDIX X – Claim chart showing the correspondence between the Black and Donaldson references and the '979 patent.

APPENDIX Y – Claim chart showing the correspondence between the Black, Donaldson, and Stibel references and the '979 patent.

APPENDIX Z – Claim chart showing the correspondence between the Black, Donaldson, Stibel, and Syskill references and the '979 patent.

I. INTRODUCTION

Pursuant to 35 U.S.C. §§ 311 *et seq.* (2002), reexamination of U.S. Patent No. 6,778,979 (the '979 patent) is requested by Google Inc. The '979 Patent issued on August 17, 2004 to Grefenstette et al. and was assigned to Xerox Corporation (collectively the "Applicants"). The Requestor is a named defendant in a proceeding in the United States District Court for the District of Delaware (case number 1:10-cv-00136-UNA) where the Patent Owner has asserted claims of the present patent against Requestor (the "pending litigation"). In light of pending litigation pursuant to 35 U.S.C. § 314 and MPEP § 2660, 8th Ed., Rev. 6 (September 2007), Requestor respectfully seeks expedited consideration of this reexamination request

A. Identification of Prior Art References Presented to Show Substantial New Questions of Patentability

Reexamination of claims 1-20 of the '979 patent is requested in view of the following prior art references.

1. PCT Application Publication No. WO 01/44992 to Wieser et al. (Appendix C)
2. U.S. Patent No. 6,236,768 to Rhodes et al. (Appendix D)
3. U.S. Patent No. 7,089,236 to Stibel (Appendix E)
4. "Syskill & Webert: Identifying interesting web sites" by Michael Pazzani, Jack Muramatsu and Daniel Billsus (Appendix F)
5. U.S. Patent No. 6,546,386 to Black et al. (Appendix G)
6. U.S. Patent No. 7,225,180 to Donaldson (Appendix H)
7. U.S. Patent No 5,748,954 to Mauldin (Appendix I)
8. "HyPursuit: A Heirarchical Network Search Engine that Exploits Content-Link Hypertext Clustering" by Ron Weiss, Bienvenido Velez, Mark A. Sheldon, Chanathip Namprempre, Peter Szilagyi, Andrzej Duda and David K. Gifford (Appendix J)
9. U.S. Patent Application Publication No. 2002/0147738 to Reader (Appendix K)
10. "Experimental Simulation for Automatic Patent Categorization" by Hisuo Mase, Hiroshi Tsuji, Hiroshi Kinukawa, Yoshinori Hosoya, Kazuya Koutani, and Kenichi Kiyota (Appendix L)

It is believed that the references at Appendices C-L were not before the Examiner during ex parte prosecution of the '979 patent.

B. Identification of Substantial New Questions of Patentability

Substantial new questions of patentability exist with respect to claims 1-20 of the '979 patent, based on references and combinations laid out in detail below and not before the Patent Office during original examination. The '979 patent is generally directed to automatically generating a search engine query based on selected document content and a specific document category and classification. Specifically, the '979 patent discloses a technique that identifies entities in selected document content and automatically generates a query to find additional information about the identified entities. The '979 patent contends that the disclosed technique improves information retrieval systems by automatically categorizing the selected document content into a specific category in an information retrieval system and then restricting application of the automatically-generated query to the specific category of the information retrieval system. (*Appendix A at 48/21 to 49/8*)

The claims of the '979 patent were allowed in response to arguments asserted by the Applicant in an Appeal Brief filed on 4/23/04 (See *Appendix B at 4/23/04*). In the Appeal Brief, the Applicant argued that the then-applied prior art failed to disclose or suggest :

(1) "**[automatically] formulating a query using a document**" (*Appendix B at 4/23/04, page 6*); and

(2) that "**the automatically formulated query is restricted to a category of information [assigned by categorizing the document]**" (*Appendix B at 4/23/04, page 7*)

However, these very features, along with all the limitations of the allowed claims, are disclosed in references that were not before the examiner at the time of allowance, and that form the specific proposed rejections of this Request. Based on the specific proposed rejections discussed below, this Request raises substantial new questions of patentability with respect to claims 1-20 of the '979 patent.

Briefly, U.S. Patent Application Publication 2002/0147738 to Reader (hereinafter "Reader" or "the Reader reference") teaches a system that (1) automatically generates a query from a selected patent document of interest, where the patent document of interest has been automatically categorized within one or more classes in a patent classification scheme of a U.S. or an international patent office. (*Appendix K at abstract and paragraphs 0004 and 0015*) The query is formulated to (2) constrain a search for Internet content to only particular information (i.e., company Website information) that has also been classified within the same one or more classes assigned to the patent document of interest. (*Appendix K at paragraph 0015*) As set forth in the included claim charts, Reader is shown to anticipate all of the independent claims (1, 14 and 18) and the dependent claims 5, 10 and 13 of the '979 patent. Reader also was not considered by the USPTO during the original prosecution.

The article "Experimental Simulation for Automatic Patent Categorization" by Hisuo Mase et al. (Advances in Production Management Systems, Kyoto, Japan, November 1996, pages 377-382, hereinafter "Mase" or "the Mase reference") was not before the Examiner during original prosecution of the '979 patent. Mase describes an automatic patent categorization system that analyzes the text of patent documents and automatically assigns the patent documents to different categories/classes within a patent classification scheme based on the results of the analysis. (*Appendix L at page 377, columns 1 and 2)* A person of ordinary skill in the art would have recognized that Reader's techniques could be advantageously combined with Mase's techniques to achieve a system that would provide an examiner at a patent office with both a patent application that is automatically classified in accordance with Mase's teachings and Web documents automatically determined as potential prior art relevant to the patent application in accordance with Reader's teachings. As discussed in further detail in Sections IV and V below, such a combined system would make the review of patent applications at the patent office more efficient by decreasing manual indexing efforts in accordance with Mase's express teachings and by decreasing manual Internet searches for prior art for new patent applications in accordance with Reader's express teachings. As set forth in the included claim charts, Reader in view of Mase is shown to render obvious all of the independent claims (1, 14 and 18) and the dependent claims 5, 10 and 13 of the '979 patent.

Additionally, PCT Application Publication No. WO 01/44992 to Wieser et al. (hereinafter "Wieser" or "the Wieser reference") was not considered by the Examiner during original prosecution. Weiser discloses a system that analyzes the content of a document, such as a Web page, to (1) automatically generate a query that is used to find product offers (e.g., advertisements) to be embedded into the Web page prior to presentation to a user. (*Appendix C at page 5, lines 8-12, page 2, lines 16-19*) The Web page is automatically categorized by identifying one or more groups/clusters of product offers deemed to match the contents of the Web page[1], and the query that is then formulated (2) constrains the search for matching product offers to those product offers that are within the identified one or more groups/clusters. (*Appendix C at page 17, lines 15-32, page 18, lines 16-22*) As set

[1] Requestor applies the Patent Owner's apparent claim construction in the related litigation when asserting that Wieser's clusters correspond to the present claims' categories/classes and that Wieser's set of clusters corresponds to the present claims' organized classification of document content. As described further in Section III. B., in a reexamination proceeding, the USPTO applies a broadest reasonable interpretation consistent with the specification. Requestor neither admits nor acquiesces as to any construction of any claim to be used in litigation by either its analysis of the patent claims or by operation of the particular prior art references applied herein, including the Patent Owner's asserted claim constructions. *See In re Trans Texas Holdings Corp.*, 498 F.3d 1290 (Fed. Cir. 2007).

forth in the claim charts, Wieser is shown to anticipate all of the independent claims (1, 14 and 18) and the dependent claims 5, 6 and 10-13 of the '979 patent.

U.S. Patent No. 6,236,768 to Rhodes et al. (hereinafter "Rhodes" or "the Rhodes reference") also discloses a system having the features relied upon by the Applicant to justify allowance of the claims of the '979 patent. Rhodes discloses a system that analyzes a current document, such as, for example, an e-mail displayed in an e-mail interface window, that a user is viewing and (1) automatically generates a query by identifying keywords in the text information of the current document to thereby find documents that are similar to the current document. In formulating the query, the selected document is automatically categorized into one or more wordcode groupings of document content (*Appendix D at 4/20-27, 4/45-55 and 5/12-28*); and the (2) application of the query is then constrained by the information retrieval system to the one or more wordcode groupings identified for the selected document. (*Appendix D at 5/12-28*) As detailed in the included claim charts, Rhodes is shown to anticipate all of the independent claims (1, 14 and 18) and the dependent claims 2, 5-7, 10-13, 15, 17 and 19 of the '979 patent.

Neither of U.S. Patent No. 6,546,386 to Black et al. (hereinafter "Black" or "the Black reference") nor "HyPursuit: A Heirarchical Network Search Engine that Exploits Content-Link Hypertext Clustering" (Proceedings of the Seventh ACM Conference on Hypertext, Washington, DC, March 1996, pages 180-193, hereinafter "HyPursuit" or "the HyPursuit reference") was considered by the Examiner during the original prosecution. A system formed as a result of the combination of the teachings of Black and HyPursuit also has the features relied upon by the Applicant to justify allowance of the claims of the '979 patent. That is, Black's teachings combined with HyPursuit's teachings disclose a system that (1) automatically generates a query from selected document content in accordance with Black's teachings, and (2) formulates the query to constrain a search to those documents within one or more categories/clusters automatically identified for the selected document content in accordance with HyPursuit's teachings.[2]

Specifically, Black describes a system that (1) automatically generates a keyword query by analyzing the body of a selected document to identify a query term corresponding to the main topic or subject matter of the document. (*Appendix G at abstract, 1/47-67, 2/10-14, 2/42-61, 3/37-57*) HyPursuit describes a hierarchical network search engine that groups documents into clusters based on both document content similarities and document hyperlink similarities. (*Appendix J at abstract*) Each

[2] Requestor applies the Patent Owner's apparent claim construction in the related litigation without waiver. See *n.* 1.

cluster of documents may be handled by a different content router of HyPursuit's search engine. (*Appendix J at page 181, column 1*) HyPursuit's search engine is configured to receive a keyword query, identify one or more clusters of documents that match the keyword query, and then (2) constrain the search to only those documents within the identified cluster(s) by routing the query to only those content routers that handle the identified cluster(s). (*Appendix J at page 182, column 1; page 185, column 1 and page 186, column 2*) HyPursuit's stated goal is to promote scalable and efficient query processing for keyword queries just like those produced by Black's system (*Appendix J at page 181, column 2*) and, therefore, as discussed in further detail in Sections IV and V below, one of ordinary skill in the art would have easily and readily applied HyPursuit's search engine to process Black's keyword queries to achieve the benefits expressly taught by HyPursuit. In such a combined system, HyPursuit's automatic identification of one or more clusters that match Black's keyword query corresponds to an automatic categorization of Black's document, and HyPursuit's constraining of the search to only those documents within the identified cluster(s) corresponds to formulating Black's query to restrict the search to only information in the identified category(ies). As set forth in the included claim charts, Black in view of HyPursuit is shown to render obvious all of the independent claims (1, 14 and 18) and the dependent claims 2, 4-7, 10-13, 15-17, 19 and 20 of the '979 patent.

U.S. Patent No. 7,225,180 to Donaldson et al. (hereinafter "Donaldson" or "the Donaldson reference") was not considered by the Examiner during the original prosecution and, in combination with Black, teaches the features used by the Applicant to justify allowance. A system formed as a result of the combination of the teachings of Black and Donaldson has the features Applicant alleged were missing from the art before the Examiner that the Applicant used to justify allowance of the claims of the '979 patent. That is, Black's teachings combined with Donaldson's teachings disclose a system that (1) automatically generates a query from selected document content in accordance with Black's teachings, and (2) formulates the query to constrain a search to those documents within one or more categories identified for the selected document content in accordance with Donaldson's teachings.

In particular, as previously stated, Black describes a system that (1) automatically generates a keyword query by analyzing the body of a selected document to identify a query term corresponding to the main topic or subject matter of the document. (*Appendix G at abstract, 1/47-67, 2/10-14, 2/42-61, 3/37-57*) Donaldson describes a system that automatically determines a category (e.g., a "non-offensive" category) for one or more search terms in a keyword query, and then (2) restricts the application of the keyword query to a data source corresponding to the determined category (e.g., to an electronic information store that only includes content categorized as "non-offensive") in order to

ensure that only desired search results are produced. (*Appendix H at 18/47-54, 19/51-65, and 20/3-18 and 19-29*)

As discussed further in Sections IV and V below, a person of ordinary skill in the art would have readily appreciated that Donaldson's system is designed to handle and process keyword queries that are just like the queries produced by Black's system. Accordingly, such a person would have easily recognized that Donaldson's techniques could be used to process Black's query in order to enable parental controls or to otherwise ensure that only desired search results are provided, as expressly taught by Donaldson, in response to Black's query. (*Appendix H at 19/51-61, 20/19-29*) In such a combined system, Donaldson's automatic identification of one or more categories as matching the term(s) of Black's keyword query corresponds to an automatic categorization of Black's document, and Donaldson's constraining of the search to only those documents within the identified category(ies) corresponds to formulating Black's query to restrict the search to only information in the identified category(ies). As set forth in the included claim charts, Black in view of Donaldson is shown to render obvious all of the independent claims (1, 14 and 18) and the dependent claims 5, 6 and 10-13 of the '979 patent.

Certain other references (i.e, Stibel, Syskill, and Mauldin) that were not considered by the USPTO during original prosecution and that are described further below render certain of the claims obvious according to the table below. To be specific, the following table shows the various substantial new questions of patentability (SNQs) set forth in this Request and described in the claim charts included as Appendices M-Z:

Reference	Questions Raised Alone or In Combination with Other References	SNQ	Claim Chart (Appendix)
Reader	Claims 1, 5, 10, 13, 14 and 18 are anticipated by Reader	1	M
	Claims 1, 5, 10, 13, 14 and 18 are rendered obvious by Reader taken in view of Mase	2	N
Wieser	Claims 1, 5, 6, 10-14 and 18 are anticipated by Wieser	3	O
	Claims 2, 7, 15 and 19 are rendered obvious by Wieser taken in view of Stibel	4	P
	Claims 3 and 8 are rendered obvious by Wieser taken in view of Stibel and Syskill	5	Q
Rhodes	Claims 1, 2, 5-7, 10-15 and 18-19 are anticipated by Rhodes	6	R
	Claims 4, 16-17, and 20 are rendered obvious by Rhodes taken in view of Stibel	7	S
	Claims 3 and 8 are rendered obvious by Rhodes taken in view of Syskill	8	T
	Claim 9 is rendered obvious by Rhodes taken in view of Black, Syskill, and Mauldin	9	U
Black and HyPursuit	Claims 1, 2, 4-7 and 10-20 are rendered obvious by Black taken in view of HyPursuit	10	V
	Claims 3 and 8 are rendered obvious by Black taken in view of HyPursuit and Syskill	11	W
Black and Donaldson	Claims 1, 5, 6, 10-14 and 18 are rendered obvious by Black taken in view of Donaldson	12	X
	Claims 2, 4, 7, 15-17, 19 and 20 are rendered obvious by Black taken in view of Donaldson and Stibel	13	Y
	Claims 3 and 8 are rendered obvious by Black taken in view of Donaldson, Stibel and Syskill	14	Z

C. Claims for Which Reexamination is Requested

In view of the forgoing discussion, the Reader, Mase, Wieser, Rhodes, Black, HyPursuit, Donaldson, Stibel, Syskill, and Mauldin references were not considered by the USPTO during prosecution of the '979 patent. The SNQ table shown in Section I of this Request and the claim charts provided as Appendices M-Z show how the features recited by claims 1-20 of that patent are met by the specific proposed rejections based on these references. Consequently, reexamination is hereby requested for claims 1-20 of the '979 patent in view of the specific proposed rejections based on the patents and publications discussed below. A copy of the '979 patent is attached as Appendix A of this document; copies of the relevant portions of the '979 patent prosecution history are attached at Appendix B.

II. SUMMARY OF THE '979 PATENT

A. Brief Overview

The '979 patent is generally directed to automatically generating a search engine query based on selected document content and restricting that query using a specific document category and classification. Specifically, the '979 patent discloses a technique that identifies entities in selected document content and automatically generates a query to find additional information about the identified entities. The '979 patent contends that the disclosed technique improves information retrieval systems by automatically categorizing the selected document content into a specific category in an information retrieval system and then restricting application of the query to the specific category of the information retrieval system. (*Appendix A at 48/21 to 49/8*)

B. The Claims of the '979 Patent

Independent claim 1 recites:

> A method for automatically generating a query from selected document content, comprising:
>
> > defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;
> >
> > automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;
> >
> > automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and
> >
> > 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 the assigned classification label.

Claims 2-13, which depend on claim 1, recite the following features:

- "limiting the query by adding terms relating to context information surrounding the set of entities in the selected document content" (claim 2);
- "the number of terms added is limited to a predefined number" (claim 3);

- "limiting the query by adding terms defining the assigned classification label" (claim 4);
- "the organized classification of document content is defined using a hierarchical organization" (claim 5);
- "using a text categorizer to assign the classification label assigned from the organized classification of content" (claim 6);
- "extracting with the text categorizer a set of terms relating to the document content" (claim 7);
- "appending to the query ones of the set of terms extracted by the text categorizer to contextualize the query" (claim 7);
- "abbreviating the set of terms extracted by the text categorizer to a predefined number of terms" (claim 8);
- "abbreviating comprises: extracting noun phrases from the selected document content; ranking the noun phrases by those that occur most frequently in the document content; defining a subset of noun phrases by identifying those ranked noun phrases that occur more frequently than a first predefined frequency; ranking those words in the subset of noun phrases by their frequency of occurrence to define an ordered list of words; defining a subset of the ordered list of words by identifying those ranked words that occur more frequently than a second predefined frequency; re-ranking the subset of words in inverse frequency to their use in the category of information in the information retrieval system identified by the assigned classification label; using only those highest ranked words in the re-ranked subset of words to define the set of terms appended to the query" (claim 9);
- "each class in the organized classification of document content has associated therewith a characteristic vocabulary" (claim 10);
- "ranking results from the query performed at the information retrieval system in accordance with one of the assigned classification label and the characteristic vocabulary" (claim 11);
- "using the method in a system for enriching selected content of a document with personalities that identify enrichment themes" (claim 12); and
- "automatically identifying the set of entities using a service that recognizes entities of a predefined type" (claim 13).

Independent claim 14 is a system claim that is similar to claim 1. Claims 15-17, which depend on claim 14, recite the following features:

- "a short length aspect vector generator for generating terms relating to context information surrounding the set of entities in the selected document content; wherein

the query generator adds the terms relating to the context information to limit the query" (claim 15);

- "the query generator further limits the query by adding terms defining the selected classification label provided by the categorizer" (claim 16); and

- "a content manager for enriching the selected document content with results provided from the information retrieval system using the query" (claim 17).

Independent claim 18 is an article of manufacture claim that is similar to claim 1. Claims 19-20 are similar to claims 2 and 4 above.

C. The Prosecution History of the '979 Patent

On December 5, 2001, the Applicant filed utility application serial number 09/683,235, claiming priority to a provisional application filed August 13, 2001. Utility application serial number 09/683,235 was allowed on May 18, 2004, and was issued as the '979 patent on August 17, 2004.

In a first office action mailed December 23, 2002, the Examiner rejected all of the claims (claims 1-20) of the utility application 09/683,235 as rendered obvious by U.S. Patent No. 6,393,427 (the Vu reference) in view of U.S. Patent No. 6,374,274 (the Myers reference). *(Appendix B at 12/23/02 action)* The Examiner identified no allowable subject matter in the action.

In response, the Applicant argued that the Vu reference singly or in combination with the Myers reference did not teach or disclose the claimed invention described by claims 1-20 of the utility application 09/683,235. Specifically, the Applicant argued that the Vu reference did not teach or disclose "the automatic formulation of a query". *(Appendix B at 3/24/03 response, page 5)* To this end, the Applicant advocated that the Vu reference singly or in combination with the Myers reference did not teach or disclose "a query[which] is automatically formulated that restricts a search for information that concerns a set of entities identified in selected document content to a category of information at an information retrieval system, where the category, assigned to the selected document content, is identified by a classification label assigned from a classification of content where each classification label of the classification corresponds to a category in the information retrieval system." *(Appendix B at 3/24/03 response, page 5)*

In the next action mailed June 6, 2003, the Examiner rejected all of the claims (claims 1-20) of the utility application 09/683,235 as rendered obvious by the Vu reference in view of the Myers reference and further in view of U.S. Patent No. 6,430,588 (the Delano reference). *(Appendix B at 6/6/03 action)* Here again, the Examiner identified no allowable subject matter in the action.

This time, in response, the Applicant amended independent claims 1, 14 and 18 to recite automatically "generating a query **from selected document content**" (1, 14 and 18), "**automatically** identifying a set of entities in the selected document content" (1, 14 and 18), "**automatically categorizing the selected document content using the organized classification of document content**." (1 and 18), and "the categorizer **automatically** assigning the selected document content a classification label from the organized classification of content" (14). *(Appendix B at 9/8/03 response, pages 2-6)* The Applicant argued that the Vu reference in view of the Myers reference and further in view of the Delano reference did not teach or disclose the invention claimed in the newly amended independent claims 1, 14 and 18.

In support of the amended claims, the Applicant argued that the references alone or in combination did not teach or disclose automatic query generation based on a set of automatically identified entities from selected document content that is automatically categorized according to the organized classification of the document. Specifically, the Applicant argued that the references alone or in combination failed to disclose "a query [that] is automatically formulated … to restrict a search at an information retrieval system for information concerning a set of entities (automatically identified from selected document content) to a category of information in the information retrieval system identified by the assigned classification label (assigned by automatically categorizing the selected document content according to the organized classification of document content*)." (Appendix B at 9/8/03 response, page 9)* Additionally, the Applicant amended dependent claim 8, as follows:

> 8. The method according to claim 7, further comprising abbreviating the set of terms extracted by the text categorizer to a predefined number of terms.

In the next action mailed November 21, 2003, the Examiner rejected claims 1-8 and 10-20 of the '979 patent as obvious in light of U.S. Patent No. 6,154,213 (the Rennison reference). In addition, the Examiner objected to claim 9 as dependent upon a rejected base claim, but the Examiner otherwise deemed claim 9 allowable (if rewritten in independent form including all of the limitations of the base claim and any intervening claims). *(Appendix B at 11/21/03 action)* This action was made final.

In response, the Applicant argued that the Rennison reference discloses users navigating through a hierarchy of topics to find a particular topic of interest and the documents associated with the particular topic. Applicant contrasted this characterization of Rennison against aspects of independent claims 1, 14, and 18 said to describe categorizing a document, assigning the document a classification label, and formulating a query to restrict a search to the category of information assigned to the classification label. Specifically, the Applicant argued that:

> "[u]nlike Rennison which discusses in column 18 systems in which 'users navigate through [a] hierarchy to find a particular Topic of interest and the associated documents' (lines 36-38), Applicant claims in independent claims 1, 14, and 18: (a) the categorization of selected document content with an organized classification of document content to assign the selected document content a classification label, and (b) the formulation of a query to restrict a search to the category of information at an information retrieval system related to the assigned classification label concerning a set of entities automatically identified in the selected document content." *(Appendix B at 1/23/04 response, page 3)*

In the next action mailed February 9, 2004, the Examiner issued an advisory action stating the request for reconsideration was considered but did not place the application in condition for allowance because claims 1-20 of the application were rendered obvious by the Rennison reference. The Examiner further indicated, for Appeal purposes, that claims 1-20 were rejected. *(Appendix B at 2/9/04 action)* By indicating that claims 1-20 were rejected, the Advisory Action contradicted the Examiner's earlier indication, in the previous office action mailed November 23, 2003, that the subject matter of claim 9 was deemed allowable.

In response, the Applicant filed a Notice of Appeal *(Appendix B at 2/23/04 response)* and a subsequent Appeal Brief. *(Appendix B at 4/23/04 Appeal Brief)* In the Appeal Brief, the Applicant (the Appellant) traversed the final rejection of claims 1-8 and 10-20 as obvious by the Rennison reference. Additionally, the Applicant advocated that claim 9 included allowable subject matter, as indicated by the Examiner in the final action mailed November 21, 2003, to which the Appeal is directed. The Applicant grouped the pending rejected claims of the utility application 09/683,235 into three groups, stating that claims in each group would stand or fall together, while addressing claim 17 independently.

Specifically, in the Appeal Brief, the Applicant argued that a first group of claims, independent claims 1, 14, and 18 and dependent claims 3, 5-8, and 10-13 were patentable over the Rennison reference because "Rennison Fails to Disclose or Suggest Restricting A Query To A Category Of Information In An Information Retrieval System Using A Classification Label." *(Appendix B at 4/23/04 Appeal Brief, page 6)* In the Appeal Brief, independent claim 1 was deemed representative of the first group, based on its recitation of "a method for formulating a query using a document, not to use a query to define a set of documents that are used to build a graph [as disclosed in the Rennison reference]." *(Appendix B at 4/23/04 Appeal Brief, page 6)* With respect to that first group, the Applicant argued:

> "… Appellant's claimed invention concerns the automatic formulation of a query in which the automatically formulated query is restricted to information concerning a set of entities identified in document content and a category of information in an information retrieval system identified by a classification label assigned by categorizing the document content. "*(Appendix B at 4/23/04 Appeal Brief, page 7)*

The Applicant further argued that "Appellant does not dynamically generate an information space in response by expanding terms of a user query using a knowledge base as taught by Rennison." *(Appendix B at 4/23/04 Appeal Brief, page 7)* The Applicant continued to argue that the first group of claims was patentable over the Rennison reference since "Rennison Fails to Disclose or Suggest Categorizing a Document to Formulate A Query". *(Appendix B at 4/23/04 Appeal Brief, page 7)* In trying to distinguish the disclosure of the Rennison reference, the Applicant argued:

> "[The] Appellant categorizes document content to identify a classification label to restrict a query to a category of information in an information retrieval system, where each classification label corresponds to a category of information in the information retrieval system. Appellant's invention as recited in independent claim 1 concerns automatic query formulation, where the formulated query restricts a search at an information retrieval system to information concerning a set of entities (automatically identified in selected document content) to a category of information in the information retrieval system identified by a classification label (assigned by categorizing the selected document content using an organized classification of document content)." *(Appendix B at 4/23/04 Appeal Brief, page 8)*

Next, in the Appeal Brief, the Applicant argued that a second group of claims, claims 2, 15, and 19 which depend from claims 1, 14, and 15, respectively, were also patentable over the Rennison reference. Using claim 2 as representative of the group, the Applicant argued "while Rennison identifies additional terms to improve (i.e., expands the possible) mappings between concepts extracted from a document and concepts in a knowledge base, Appellant further constrains the formulation of a query to be applied to a category of information in an information retrieval system." *(Appendix B at 4/23/04 Appeal Brief, pages 9-10)* The applicant further argued:

> "In contrast, Appellant's claim 2 recites formulating a query by further adding terms defining an assigned classification label. As set forth above, Rennison fails to describe automatically generating a query from selected document content by, in part, (a) *categorizing* the selected document content, and (b) formulating a query to restrict a search to a category of Information at an information retrieval system, and (c) adding terms to the query made up of an identified set of entities from context information surrounding the set of entities in the selected document content." *(Appendix B at 4/23/04 Appeal Brief, page 10)*

Third, in the Appeal Brief, the Applicant argued that a third group of claims, claims 4, 16, and 20, which depend from dependent claims 2, 15, and 19, respectively, were also patentable over the Rennison reference. Using claim 4 as representative of the group, the Applicant argued "… Rennison in 'finding documents that match a user's query performs query *expansion* (see Rennison column 26. lines 25-35), and in 'organizing the results in a structured space' categorizes all of the documents that match the user query to build the structured space (see Rennison column 26, lines 42-46)." *(Appendix B at 4/23/04 Appeal Brief, page 10)* The Applicant further argued:

"In contrast, Appellant's claim 4 recites formulating a query by further adding terms defining an assigned classification label. As set forth above, Rennison falls to describe either when finding documents that match a users query or when thereafter organizing the documents in a structure space to automatically generate a query from selected document content by, in part, (a) *categorizing* the selected document content, and (b) formulating a query to *restrict* a search to a category of information at an information retrieval system, (c) adding terms to the query made up of an identified set of entities from context information surrounding the set of entities in the selected document content, and (d) further adding terms to the query defining an assigned classification label identifying the category of information in the information retrieval system." *(Appendix B at 4/23/04 Appeal Brief, page 10)*

Finally, the Applicant argued that claim 17, which depends from claims 16, 15, and 14 is patentable over the Rennison reference, based on the following assertion:

"However, the sections of Rennison cited in the Office Action fail to disclose or suggest, as recited by Appellant in claim 17, the 'enrichment' or 'annotation' of document content with search results provided from an information retrieval system using an automatically formulated query. Moreover, as discussed above Rennison further fails to describe or suggest, as recited by Appellant in claim 17 which read together with claims 16, 15, and 14, the automatic formulation of a query that is used to query an information provider for the results that are used to enrich the document content." *(Appendix B at 4/23/04 Appeal Brief, pages 13-14)*

Thereafter, the Examiner issued a Notice of Allowance on May 18, 2004 (*Appendix B at 5/18/04 Notice of Allowance)* allowing all pending claims (claims 1-20) of the utility application 09/683,235. The application issued as U.S. Patent No. 6,778,979 on August 17, 2004. In the Notice of Allowance, **the Examiner's reasons for allowance indicated that he believed that the prior art of record did not describe or suggest various recitations of independent claim 1**:

"The prior art does not teach the combination steps, as follows:
Automatically identifying a stet *(sic)* of entities in the selected document content for searching additional information related, automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content, and 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 the assigned classification label." *(Appendix B at 5/17/04 Notice of Allowance, pages 2-3)*

III. LAW GOVERNING REEXAMINATION

A. Any Printed Publication, Including Electronic Publications, Can Raise a Substantial New Question of Patentability

In support of this Request, Requestors have identified patents and published patent applications that raise substantial new questions of patentability. All of these references are appropriate for

consideration since any patent or printed publication bearing on the patentability of a claim of a particular patent can justify reexamination. *See* 35 U.S.C. §§ 301 & 311(a). Any document qualifies as a "printed publication" if it is accessible to the public,[3] and electronic publications are treated the same way as traditional printed publications.[4]

B. Claims Are Broadly Interpreted During Reexamination

The proper claim interpretation standard used by the PTO in an *inter partes* reexamination proceeding differs from those applied by a district court in a litigation context. During an *inter partes* reexamination, claims are to be interpreted according to the "broadest reasonable interpretation" consistent with the specification. MPEP § 2258(I)(G) ("During reexamination, claims are given the broadest reasonable interpretation . . ."); *see also* MPEP § 2658 (incorporating the standards for ex parte reexamination (MPEP § 2258) for *inter partes* reexamination). This standard differs from those used when determining validity and infringement during litigation.

The Federal Circuit has recognized this difference in the treatment of claims before the PTO and district courts. As explained in *Atlantic Thermoplastics Co. v. Faytex Corp.*, 970 F.2d 834, 23 U.S.P.Q.2D (BNA) 1481 (Fed. Cir. 1992):

> This court already distinguishes treatment of claims for patentability before the PTO from treatment of claims for validity before the courts. This court permits the PTO to give claims their broadest reasonable meaning when determining patentability. During litigation determining validity or infringement, however, this approach is inapplicable. Rather the courts must consult the specification, prosecution history, prior art, and other claims to determine the proper construction of the claim language.

Id. at 846 (internal citations omitted). In other words, the PTO "is required to use a different standard for construing claims than that used by district courts," and vice versa. *In re American Academy of Science Tech. Center*, 367 F.3d 1359, 1369, 70 U.S.P.Q.2D (BNA) 1827, 1834 (Fed. Cir. 2004) (reiterating that "it is error for the [PTO] to apply the mode of claim interpretation that is used by courts in litigation, when interpreting the claims of issued patents in connection with determinations of

---

[3] *See* MPEP 2128 (A reference is proven to be a "printed publication" "upon a satisfactory showing that such document has been disseminated or otherwise made available to the extend that persons interested and ordinarily skilled in the subject matter or art, exercising reasonable diligence, can locate it." (citing *In re Wyer,* 655 F.2d 221, 227 (CCPA 1981)).

[4] *See id.* ("An electronic publication, including an on-line database or Internet publication, is considered to be a 'printed publicaton' within the meaning of 35 U.S.C. § 102(a) and (b) provided the publication was accessible to persons concerned with the art to which the document relates.").

infringement and validity"); *see also Atlantic Thermoplastics*, 970 F.2d at 846, 23 U.S.P.Q.2D (BNA) at 1491.

The claim analysis employed in this Request generally sets forth claim interpretations asserted by the Patent holder during litigation. In analyzing the claims in this Request by setting forth these interpretations, and otherwise, Requestor neither admits nor acquiesces as to any construction of any claim to be used in litigation by either its analysis of the patent claims or by operation of the particular prior art references applied herein, including the Patent Owner's asserted claim constructions. *See In re Trans Texas Holdings Corp.*, 498 F.3d 1290 (Fed. Cir. 2007).

IV. THE REFERENCES RELIED UPON HEREIN PROVIDE NEW, NON-CUMULATIVE TECHNICAL TEACHINGS

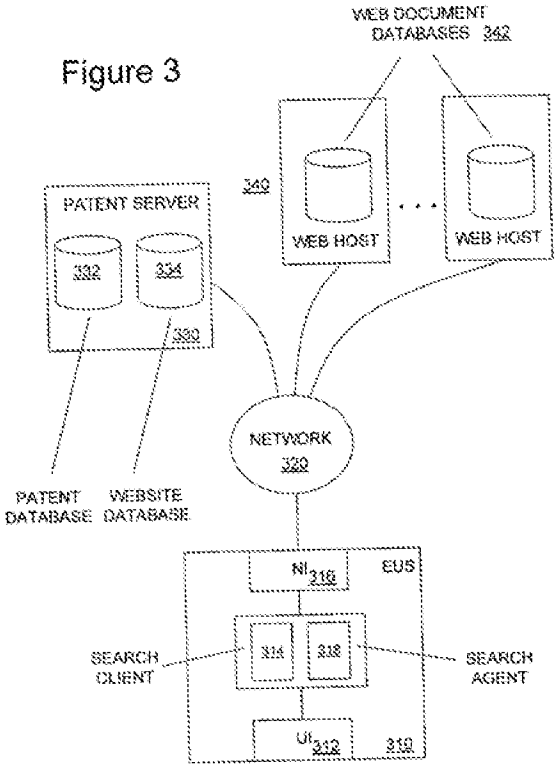
A. Reader

U.S. Patent Application Publication No. 2002/0147738 to Reader qualifies as prior art under 35 U.S.C. § 102(e) because it was filed on April 6, 2001 prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001) but it was published on October 10, 2002 after the filing date of the '979 patent (i.e., August 13, 2001). The Reader reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Reader reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

The Reader reference describes an electronic search system that automatically generates queries from patent documents to find patent-relevant publications on the Internet. Specifically, Reader recognizes that the continued expansion of the Internet has made the Internet a valuable resource for patent practitioners when attempting to find "prior art" publications to invalidate a patent. (*Appendix K at paragraphs 0001-002)* Reader notes that practitioners have historically relied on relatively manual Internet search techniques to find patent-relevant publications on the Internet. (*Appendix K at paragraph 0002)* Reader proposes a system that automates the searching process for such "prior art" publications by enabling a user to select the content of a patent (e.g., the full text of the patent or one or more claims of the patent) and then, using a computer, automatically generates a query from the selected patent content to find relevant publications on the Internet. (*Appendix K at paragraph 0004)*

Reader's system, as shown in Fig. 3 of Reader (reproduced below), includes an end user station (EUS) 310 that communicates with a patent server 330 having a patent database 332 and a Website database 334. (*Appendix K at paragraph 0014)*

Figure 3



As shown in Fig. 3, Reader's patent database 332 stores entries that associate patent-identifying (PI) attributes, such as a patent number and optionally claim numbers, inventor names, and assignee names, to patent language (e.g., claim text or the full text of the patent). The entries in the patent database 332 also associate the PI attributes to the patent classification(s) assigned by the U.S. patent office[5] or by an international patent office to the patent document(s) identified by the PI attributes. (*Appendix K at paragraph 0014)* The Website database 334 stores entries that associate company Website identifiers (e.g., URLs of company home pages) to patent classifications assigned by the U.S. or by an international patent office to the patents held by those companies. (*Appendix K at paragraph 0014)*

In operation, a user of the EUS 310 selects PI attributes for a patent document of interest and the EUS 310 provides the PI attributes to the patent server 330 for processing. (*Appendix K at paragraph 0015, corresponding to operations 405 and 410 of Fig. 4)* The patent server 330 uses the PI attributes to access the patent database 332 to find the corresponding patent language (e.g., the full text of the patent) identified by the PI attributes and to automatically categorize the patent document

[5] Requestor refers to the U.S. Patent and Trademark Office (USPTO) as the "U.S. patent office" at various locations in the Reader portions of this Request solely for ease of reference.

by retrieving one or more patent classification(s) set forth by a U.S. or international patent classification scheme for the patent document. The patent server 330 sends the patent language and the patent classification information to the EUS 310. (*Appendix K at paragraph 0015, corresponding to operation 415 of Fig. 4)* In response to a further automatic request from the EUS 310, the patent server 330 also automatically identifies, using the Website database 334, one or more company Website identifiers (e.g., URLs) that have the same classification(s) as the patent document corresponding to the PI attributes. (*Appendix K at paragraph 0015, corresponding to operations 420-430 of Fig. 4)* The company Website identifiers are also returned to the EUS 310. The EUS 310 then automatically abstracts the patent language to identify Web Document Identifying (WDI) attributes that are used to form a query for related Web documents. (*Appendix K at paragraph 0015, corresponding to operation 435 of Fig. 4)* The WDI attributes may be, for example, keywords extracted from the full text of the patent content separated by Boolean operators. The EUS 310 then automatically forms a query to search only those Web sites identified by the company Website identifiers for Web documents that match the WDI attributes of the patent document of interest. (*Appendix K at paragraph 0015, corresponding to operation 440-445 of Fig. 4).* The results of the search are outputted by the EUS 310 to a user as a list of Web document identifiers (e.g., URLs) corresponding to Web documents in the identified Web sites that are believed to be relevant to the patent document of interest.

Accordingly, Reader describes a system that automatically generates a query from a selected patent document of interest, where the patent document of interest has been automatically categorized within one or more classes in a patent classification scheme like that of a U.S. or an international patent office. The query is formulated to constrain a search for Internet content to only particular information (i.e., company Website information) that has also been classified within the same one or more classes assigned to the patent document of interest.

The Reader disclosure teaches the features, i.e., the restricting of a query (for patent-relevant Internet publications) to the category of the (patent) document, deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, as set forth in the SNQ table presented in Section I of this Request, these teachings, and thus, the Reader reference, present substantial new questions of patentability with respect to the claims of the '979 patent.

B. Mase

The article "Experimental Simulation for Automatic Patent Categorization" by Hisuo Mase, Hiroshi Tsuji, Hiroshi Kinukawa, Yoshinori Hosoya, Kazuya Koutani, and Kenichi Kiyota qualifies as prior art under 35 U.S.C. § 102(b) because it was published in 1996, more than one year prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001). The Mase reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Mase reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

Mase describes an automatic patent categorization system that analyzes the text of patent documents and automatically assigns the patent documents to different categories/classes within a patent classification scheme based on the results of the analysis. (*Appendix L at page 377, columns 1 and 2)* Mase recognizes that a patent office may receive hundreds of thousands of patent applications in a given year and may have thousands of different patent categories/classes to which the received patent applications have to be assigned for processing. (*Appendix L at page 377, column 1)* Mase asserts that patent applications "should be appropriately indexed as soon as possible" (i.e., assigned to their appropriate categories/classes in the patent classification scheme as soon as possible) "to manage the review, search and citation of patents at the patent office efficiently." (*Appendix L at page 377, column 1)* However, classifying the patent applications quickly has proven difficult due to the large number of patent applications received by a patent office and the time intensive nature of the manual classification process currently performed by indexing experts at the patent office. (*Appendix L at page 377, columns 1 and 2)* To address this problem, Mase proposes an automated patent classification system that employs a classification tool called "FLUTE." (*Appendix L at page 377, column 2)*

As shown in Fig. 1 of Mase (reproduced below), FLUTE includes a Keyword Extraction Engine that receives the text of a new patent application and analyzes the text to identify keywords deemed to represent the content of the new patent application. (*Appendix L at page 378, column 1)*

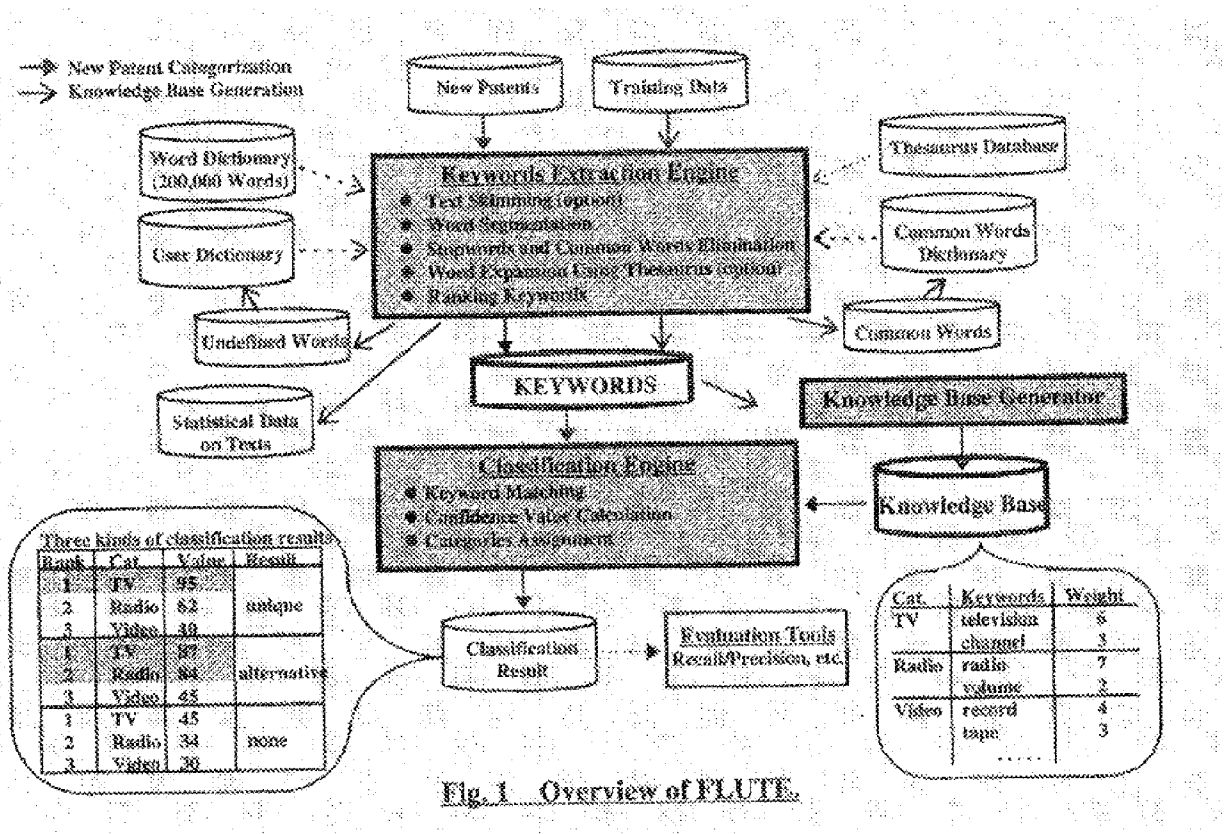


Fig. 1 Overview of FLUTE.

The identified keywords are then used by a Classification Engine to identify one or more categories/classes within a patent classification scheme for the new patent application. Specifically, the Classification Engine identifies the one or more categories/classes by calculating the similarity between the keywords extracted from the new patent application to keywords stored in a Knowledge Base that characterize the various categories/classes of the patent classification scheme. (*Appendix L at page 379, column 1)* Those categories/classes identified for the new application with a high degree of confidence are automatically assigned to the new application, while those identified with a lower degree of confidence may require further review by an indexing expert. (*Appendix L at page 378, columns 1 and 2)* To the extent that the Reader reference is considered not to disclose or suggest an automatic categorization of a selected patent document, it would have been obvious to one of ordinary skill in the relevant art to have incorporated in Reader's system features taught by Mase to arrive at what is claimed. In particular, Mase teaches a system that may be used by a patent office to automatically categorize received patent applications using a patent classification scheme. Mase indicates that such an automatic categorization is desirable to "manage the review, search, and citation of patents at the patent office efficiently". (*Appendix L at page 377, column 1).*

A person of ordinary skill in the art would have readily recognized that Reader's system also could be used to manage the review of patents at a patent office more efficiently. Specifically, a person of ordinary skill in the art would have recognized that Reader's techniques could be advantageously combined with Mase's techniques to achieve a system that would provide an examiner at the patent office with both a patent application that is automatically classified in accordance with Mase's teachings and Web documents automatically determined as potential prior art relevant to the patent application in accordance with Reader's teachings. Such a combined system would make the review of patents at the patent office more efficient by decreasing manual indexing efforts in accordance with Mase's express teachings and by decreasing manual Internet searches for prior art for new patent applications in accordance with Reader's express teachings. Moreover, Mase expressly suggests such a combined system by identifying the related work of Yoden: "Yoden developed a patent search system, which supports finding relevant patents to an inputted one. It extracts weighted keywords...it then compares these keywords with those of [other] patents." (*Appendix L at page 382, column 1*) Thus, Mase expressly contemplates combining the disclosed patent classification system with a patent search system akin to that of Reader (i.e., a patent search system that supports finding documents related to an input patent by extracting keywords from the patent in question).

Furthermore, the combination of Mase's teachings with those of Reader would involve merely combining and/or substituting known prior art elements to yield predictable results. One skilled in the art could have easily combined the known elements as claimed by known methods, with each element in the combination performing the same function as it does separately, and the combination yielding nothing more than predictable results to one skilled in the relevant art.

The Mase disclosure, in combination with the Reader disclosure, teaches features, i.e., the restricting of a query (for patent-relevant Internet publications) to the category of the (patent) document, deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, as set forth in the SNQ table presented in Section I of this Request, these teachings, and thus, the Mase reference, present substantial new questions of patentability with respect to the claims of the '979 patent.

C. Wieser et al.

PCT Application Publication No. WO 01/44992 to Wieser et al. qualifies as prior art under 35 U.S.C. § 102(a) because it was published on June 21, 2001, prior to but within one year of the earliest

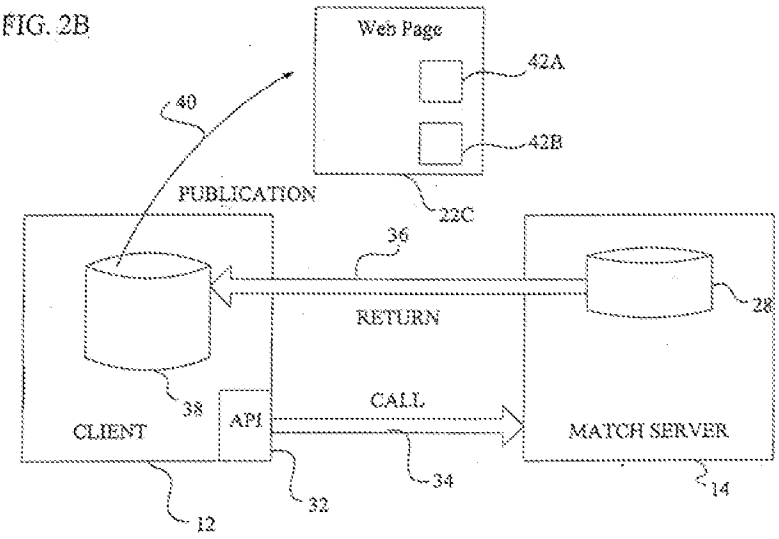
claimed priority date of the '979 patent (i.e., August 13, 2001). The Wieser reference also qualifies as prior art under 35 U.S.C. § 102(e) because it was filed on October 31, 2000, claiming priority to a U.S. Provisional Application (60/170,974) filed on December 15, 1999, which is about one year and 8 months prior to the August 13, 2001 earliest claimed priority date of the '979 patent. The Wieser reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Wieser reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

Wieser, like the 979 patent, is directed to an electronic search system that automatically generates queries from documents to annotate the documents. In particular, the Wieser reference teaches an electronic searching system that automatically generates a query from all or a portion of a selected document, such as, for example, a Web page. The Wieser system generates a query context vector for the selected document content by identifying a set of keywords in the selected document content for retrieval of offers (e.g., advertisements) related to the keywords. (*Appendix C at page 2, lines 16-19 and page 5, lines 8-12*) The selected document content (e.g., the Web page content) is automatically categorized in that, as part of the matching process performed by the Wieser system to find offers related to the keywords, one or more spheres or clusters of product offer vectors are matched to the query context vector.[6] (*Appendix C at page 15, lines 8-11; page 17, lines 15-32*) The Wieser system then restricts the application of the query to only those product offers that have vectors within the one or more spheres or clusters matched to the query context vector. (*Appendix C at page 18, lines 16-22*) The product offers deemed most similar to the selected document content may be included in the selected document content when the selected document content is presented to a user. (*Appendix C at page 5, line 22 to page 6, line 15; page 8, lines 7-15; page 9, lines 9-16*)

Fig. 2B of Wieser, reproduced below, shows an example of a Web page that is displayed on a browser 22C that has been annotated to include product offers 42A and 42B as embedded data.

[6] As stated previously, Requestor applies the Patent Owner's apparent claim construction in the related litigation without waiver. See *n.* 1.

FIG. 2B



Specifically, in the example shown in Fig.2B of Wieser, a query is automatically generated from a Web page selected for display on a browser 22C since a query context vector is automatically generated by a match server 14 by extracting keywords from the text of the Web page. The Web page is then automatically categorized, based on the claim construction that appears to be advanced by the Patent Owner in the related litigation, since the match server 14 automatically identifies spheres or clusters of product offer vectors that match the query context vector generated for the Web page. The application of the query is then restricted to the category or categories assigned to the Web page in that the comparison of the query context vector to product offer vectors is restricted to only those product offer vectors within the identified spheres or clusters. In the example shown in Fig. 2B, the comparison of the query context vector to the product offer vectors in the identified spheres or clusters yields matched product offer vectors that correspond to product offers 42A and 42B. Having found the matching product offers, the match server 14 returns data for the Web page along with additional embedded data for the product offers 42A and 42B to the client 12 for display to the user via the browser 22C.

As described above, Wieser's disclosure teaches the features, i.e., the restricting of a query (for product offers) to the category of the (Web) document, deemed missing from prior art during original prosecution, and responsible for the Examiner's allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, these teachings, and thus, the Wieser reference, present substantial new questions of patentability with respect to the claims of the '979 patent set forth in the SNQ table presented in Section I of this Request.

D. Rhodes et al.

U.S. Patent No. 6,236,768 to Rhodes et al. qualifies as prior art under 35 U.S.C. § 102(e) because it was filed on May 1, 1998 prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001). The Rhodes reference also qualifies as prior art under 35 U.S.C. § 102(a) because it was published on May 22, 2001, prior to but within one year of the earliest claimed priority date of the '979 patent (i.e., August 13, 2001). The Rhodes reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Rhodes reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

Rhodes, like the '979 patent, is directed to an electronic search system that automatically generates queries from documents to annotate the documents. In particular, the Rhodes reference teaches an electronic searching system that automatically analyzes the text information of a selected document with which the user is interacting (e.g., the body text of an e-mail displayed in an Emacs window) and automatically generates a query from the text information of the selected document. (*Appendix D at 13/19-34*) In developing the query, words appearing in the selected document are identified and used to determine one or more wordcodes using a wordvec file format, and the query is then restricted to searching only those documents corresponding to the wordcode(s). (*Appendix D at 4/20-27, 4/45-55 and 5/12-28*) For example, a selected document including the word "pickle" is matched against and produces a numeric wordcode. The wordcode then yields a vector, revealing all documents that include at least one occurrence of the corresponding word "pickle" against which more complex processing can be performed. In this process, derivation of the wordcode for the word "pickle," is done for the selected document, such that the selected document is automatically categorized based on its content. The query formed for the selected document is then restricted to searching those documents corresponding to the wordcode for "pickle", and, thus, to searching those documents that, like the selected document, include at least one occurrence of the word "pickle."

The grouping of documents by wordcodes in a wordvec file format is shown by Rhodes in a table, reproduced below:

(int)	(width*uns int)	(int)	(uns int)	(uns int)		(uns int)
NUM_WORDS,	WORDCODE-1, WORDCODE-2, etc.	NUM_DOCS=N1, NUM_DOCS=N2,	DOC-1, DOC-1,	DOC-2, . . . , DOC-2, . . . ,		DOC-N1, DOC-N2,

In particular, as shown in the above table, the wordcode "WORDCODE-1" is a category of documents that contains "N1" documents listed as "DOC-1, DOC-2, … , DOC-N1." A query having a word corresponding to "WORDCODE-1" will be restricted to searching the corresponding documents listed in the same row of the wordvec file as ""DOC-1, DOC-2, … , DOC-N1."

After finding the documents responsive to the query, a short one line summary of the few most relevant documents may be inserted into the selected document. (*Appendix D at 2/5-5-9*) This disclosure teaches the features, i.e., the restricting of a query (for similar documents) to the category of the (current) document, deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office.

Moreover, these teachings, and thus, the Rhodes reference, present substantial new questions of patentability with respect to the claims of the '979 patent set forth in the SNQ table presented in Section I of this Request.

E. Black et al.

U.S. Patent No. 6,546,386 to Black et al. qualifies as prior art under 35 U.S.C. § 102(e) because it was filed on October 31, 2000 prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001) but it was published on April 8, 2003 after the filing date of the '979 patent (i.e., August 13, 2001). The Black reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Black reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

Black, like the 979 patent, is directed to an electronic search system that automatically generates queries from documents to annotate the documents. Black discloses a Brilliant Query System that automatically generates queries from documents presented to a user. The queries are automatically generated by analyzing the text of the document to determine the main topic or subject matter of the document and the resulting queries are then included as links in the document presented to the user. Accordingly, Black describes a system that analyzes the text of a selected document and automatically generates a query that includes a term corresponding to the main topic or subject matter of the document (referred to as a "hook") and an additional term. (*Appendix G at abstract, 1/47-67, 2/10-14, 2/42-61, 3/37-57*) The automatically generated query can then be provided to a search engine for processing to determine search results. (*Appendix G at 3/55-3/57*)

This disclosure teaches features, i.e., the formulation of a query using a document, deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims.

As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, as set forth in the SNQ table presented in Section I of this Request, these teachings, and thus, the Black reference, present substantial new questions of patentability with respect to the claims of the '979 patent.

F. HyPursuit

The article "HyPursuit: A Heirarchical Network Search Engine that Exploits Content-Link Hypertext Clustering" by Ron Weiss, Bienvenido Velez, Mark A. Sheldon, Chanathip Namprempre, Peter Szilagyi, Andrzej Duda and David K. Gifford qualifies as prior art under 35 U.S.C. § 102(b) because it was published in 1996, more than one year prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001). The HyPursuit reference was not before the Examiner during original prosecution of the '979 patent, and as such, the HyPursuit reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

HyPursuit describes a hierarchical network search engine that clusters hypertext documents for purposes of searching. For this clustering, HyPursuit uses a combination of term similarities and content-link similarities (referred to as content-link hypertext clustering). (*Appendix J at abstract*). HyPursuit also contemplates organizing the resulting clusters based on various document grouping techniques, such as based on the Library of Congress catalog subjects. (*Appendix J at abstract and at page 184, column 1*). To support scalable query processing, the contents of each cluster of documents are summarized using a content label and each cluster is managed by a particular content router, with the documents themselves handled by leaf information servers. (*Appendix J at page 181, column 1 to page 182, column 1; Fig. 3 and page 184*). These content routers and leaf information servers together form a hierarchical network of information servers that is shown in Fig. 3 of HyPursuit.

To achieve efficient and scalable query processing, the HyPursuit system leverages the clusters and corresponding content labels during query processing by identifying one or more relevant clusters for a given query and forwarding the query to only those information servers corresponding to those clusters. The results received from these disparate information servers are then merged into a single result set produced in response to the query. (*Appendix J at page 182, column 1; page 185, column 1; page 186, column 2*). Using the Patent Owner's apparent claim construction in the related litigation, HyPursuit's identification of one or more relevant clusters (and thus content labels) for the given query and the routing of the query to only those information servers corresponding to those clusters for processing corresponds to an automatic categorization of the query and to a restriction of the

associated search to the category of information in the HyPursuit system identified by the associated content labels. [7]

Black's query is a keyword query, just like those processed by HyPursuit's search engine. (*Appendix J at Fig. 6*). Accordingly, one of ordinary skill in the art would have easily recognized that Black's system is ready for improvement through use of HyPursuit's techniques. Specifically, a person of ordinary skill in the art would have readily appreciated that HyPursuit's techniques could be used to process Black's keyword query to enable scalable and efficient query processing. In applying HyPursuit's techniques to Black's query, Black's document would be automatically categorized via identification of one or more clusters as corresponding to Black's query, and the corresponding category or categories would be used to restrict application of Black's query to only those information servers corresponding to the identified cluster(s). Furthermore, the combination of HyPursuit's teachings with those of Black would involve merely combining and/or substituting known prior art elements to yield predictable results. One skilled in the art could have easily combined the known elements as claimed by known methods, with each element in the combination performing the same function as it does separately, and the combination yielding nothing more than predictable results to one skilled in the relevant art.

The HyPursuit disclosure, in combination with the Black disclosure, teaches features, i.e., the restricting of a query to the category of the document, deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, as set forth in the SNQ table presented in Section I of this Request, these teachings, and thus, the HyPursuit reference, present substantial new questions of patentability with respect to the claims of the '979 patent.

G. Donaldson et al.

U.S. Patent No. 7,225,180 to Donaldson et al. qualifies as prior art under 35 U.S.C. § 102(e) because it was filed on December 28, 2000 prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001) but it was published on May 29, 2007 after the filing date of the '979 patent (i.e., August 13, 2001). The Donaldson reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Donaldson reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

[7] As stated previously, Requestor applies the Patent Owner's apparent claim construction in the related litigation without waiver. See *n.* 1.

Donaldson describes a system that intelligently filters search results provided by an electronic search system in response to a query. The intelligent filtering enables parental controls to be applied and ensures that only desired search results are provided to a user. In particular, the Donaldson reference teaches automatically determining a category that corresponds to one or more search terms in a query (e.g., a "non-offensive" category, a "legal" category, and a "sports" category) and restricting the application of the query to an electronic information store that only includes document content that has been classified to be within that category in order to ensure that only desired search results are produced. (*Appendix H at 18/47-54, 19/51-65, and 20/3-29*)

As shown in Figs. 9b and 9c of Donaldson, Donaldson's assignment of a category to the one or more terms is used to improve the search results generated by the query by restricting the application of the search query to an electronic information store classified under that category. Specifically, a search term is categorized as either being in a first category or a second category. If the search term is categorized as being in the first category, the search is restricted to searching data stored within the first electronic information store 992. In contrast, if the search term is categorized as being in the second category, the search is restricted to searching data stored within the second electronic information store 994.

For example, if the category of the one or more search terms is determined to be "sports," the application of the query is restricted to only documents within an electronic information store (e.g., store 992 or 994 of Fig. 9c) classified as being within the "sports" category. In another example, if the category of the one or more search terms is determined to be "non-offensive," the application of the query is restricted to only documents within an electronic information store (e.g., store 992 or 994 of Fig. 9c) classified as being within the "non-offensive" category. (*Appendix H at 20/3-29*) Donaldson's techniques for categorizing search terms and restricting the application of the corresponding search query improve search systems by providing more relevant search results and, in some implementations, effecting parental controls. (*Appendix H at 20/19-29*)

Donaldson's system operates on keyword queries just like those produced by Black's system. Accordingly, one of ordinary skill in the art would have easily recognized that Black's system is ready for improvement through use of Donaldson's techniques. Specifically, a person of ordinary skill in the art would have readily appreciated that Donaldson's techniques could be used to automatically categorize Black's document via categorization of the term corresponding to the main topic or subject matter of Black's document and then could be used to restrict application of Black's query to a data source corresponding to that category to ensure that only desired search results are provided. Furthermore,

the combination of Donaldson's teachings with those of Black would involve merely combining and/or substituting known prior art elements to yield predictable results. One skilled in the art could have easily combined the known elements as claimed by known methods, with each element in the combination performing the same function as it does separately, and the combination yielding nothing more than predictable results to one skilled in the relevant art.

The Donaldson disclosure, in combination with the Black disclosure, teaches features, i.e., the restricting of a query to the category of the document, deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, as set forth in the SNQ table presented in Section I of this Request, these teachings, and thus, the Donaldson reference, present substantial new questions of patentability with respect to the claims of the '979 patent.

H. Stibel

U.S. Patent No. 7,089,236 to Stibel qualifies as prior art under 35 U.S.C. § 102(e) because it was filed on October 13, 1999 prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001) but it was published on August 8, 2006, later than the filing date (i.e., August 13, 2001) of the '979 patent. The Stibel reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Stibel reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

In particular, the Stibel reference teaches a search engine interface system that receives a search query having one or more query terms and that processes the search query to generate a new search query capable of more effectively retrieving information that is relevant to the original search query. (*Appendix E at 2/42-47*) The processing includes identifying a meaning or Sense of one or more query terms, identifying related terms that are likely to appear on a Web page when a user is querying for information associated with that identified meaning, and then modifying the original query to form the new search query by appending the meaning and one or more of the related terms to the original query. For example, if the original query includes only the term "java" and is determined to correspond to the meaning "coffee" such that it also relates to the terms "espresso" and "beverage," the system may modify the search query to be the new search query "java + coffee + espresso + beverage." *(Appendix E at 11/56-67, column 12, table 2)*

Table 2 of Stibel, reproduced below, illustrates an example of a list of terms that Stibel's system may generate to improve the user query "java."

TABLE 2

Search Terms Generated by the Relational Knowledgebase for User Query "java"

java (required)	decaffeinated coffee	ice coffee
coffee (meaning)	decaf	mocha
cafe au lait	espresso	Turkish coffee
cafe noir	capuccino	cafe royale
demitasse	coffee capuccino	beverage
	iced coffee	coffee royal
		Irish coffee

As shown in Table 2, Stibel's system may generate the meaning term "coffee" and additional related terms "café au lait," "café noir," "demitasse," etc. for the user query "java." Any one or more of these generated terms may then be used by Stibel's system to further refine the user query "java" as described above.

Stibel's system operates on keyword queries to allow the keyword queries to "more effectively retrieve information from the database that is relevant to the query of the user" and to "increase the likelihood that the Internet search engine will return a meaningful hit list." (*Appendix E at 2/45-47 and 12/18-21*) The system of Wieser generates keyword queries just like those that are processed by Stibel's system. Therefore, one of ordinary skill in the art would have easily recognized that the system of Wieser is ready for improvement through use of Stibel's techniques. Specifically, a person of ordinary skill in the art would have readily recognized that Stibel's techniques for processing a search query could be easily applied to the search queries of Wieser to improve the effectiveness of the search queries by further limiting the queries through the addition of meanings and related terms. Furthermore, the combination of Stibel's teachings with those of Wieser would involve merely combining and/or substituting known prior art elements to yield predictable results. One skilled in the art could have easily combined the known elements as claimed by known methods, with each element in the combination performing the same function as it does separately, and the combination yielding nothing more than predictable results to one skilled in the relevant art.

The system of Rhodes also generates keyword queries just like those that are processed by Stibel's system. Accordingly, for the reasons stated above with respect to Wieser, a person of ordinary skill in the art would also have readily recognized that Stibel's techniques for processing a search query could be easily applied to the search queries of Rhodes. Furthermore, the combination of Stibel's teachings with those of Rhodes would involve merely combining and/or substituting known prior art elements to yield predictable results.

The system of Black, as modified based on the teachings of Donaldson in the manner set forth in Section IV(G), also generates keyword queries just like those that are processed by Stibel's system. Accordingly, for the reasons stated above with respect to Wieser, a person of ordinary skill in the art would also have readily recognized that Stibel's techniques for processing a search query could be easily applied to the search queries of Black, as modified based on Donaldson's teachings. Furthermore, the combination of Stibel's teachings with those of Black and Donaldson would involve merely combining and/or substituting known prior art elements to yield predictable results.

This disclosure teaches items deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, as described above and as set forth in the SNQ table presented in Section I of this Request, Stibel justifies multiple substantial new questions of patentability with respect to the claims of the '979 patent.

I. Syskill

The article "Syskill & Webert: Identifying interesting web sites" by Michael Pazzani, Jack Muramatsu and Daniel Billsus qualifies as prior art under 35 U.S.C. § 102(b) because it was published in 1996, more than one year prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001). The Syskill reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Syskill reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

In particular, the Syskill reference describes a system that develops a user profile for a user that browses the World Wide Web. To this end, Syskill receives a rating from the user for each visited Web page and analyzes the information on each Web page. (*Appendix F at abstract*) The Syskill system is then able to generate a query based on the user profile that is submitted to a search engine to find Web pages that may be of interest to the user. Syskill's system limits the number of terms in the search query submitted to the search engine to a predetermined number (i.e., 14 terms) in recognition of the fact that search engines cannot accept very long queries. (*Appendix F at page 56, second column, lines 3-6*)

Syskill operates on keyword queries just like those generated by the system of Wieser. Syskill expressly teaches what a person of ordinary skill in the art would have already known – i.e., that any realistic implementation of an electronic search engine, like that of Wieser, has a limited amount of computing resources and, therefore, would limit the number of terms/keywords in a keyword query inputted into the search engine for processing. Accordingly, a person of ordinary skill in the art would

have readily recognized that Syskill's limitation of terms in a query to a predetermined number could be easily applied to the search queries of Wieser (as modified by Stibel in the manner set forth in Section IV(H)) in recognition of the limited processing capabilities of search engines, as expressly disclosed by Syskill.

The system of Rhodes also generates keyword queries just like those that are processed by Syskill's system. Accordingly, for the reasons stated above, a person of ordinary skill in the art would also have readily recognized that Syskill's limitation of terms in a query to a predetermined number could be easily applied to the search queries of Rhodes.

The system of Black, as modified based on the teachings of HyPursuit in the manner set forth in Section IV(F), also generates keyword queries just like those that are processed by Syskill's system. Accordingly, for the reasons stated above, a person of ordinary skill in the art would also have readily recognized that Syskill's limitation of terms in a query to a predetermined number could be easily applied to the search queries of Black, as modified based on HyPursuit's teachings.

The system of Black, as modified based on the teachings of Donaldson and Stibel in the manner set forth in Section IV(H), also generates keyword queries just like those that are processed by Syskill's system. Accordingly, for the reasons stated above, a person of ordinary skill in the art would also have readily recognized that Syskill's limitation of terms in a query to a predetermined number could be easily applied to the search queries of Black, as modified based on Donaldson's teachings and Stibel's teachings.

Moreover, as set forth in the SNQ table presented in Section I of this Request, these teachings, and thus, the Syskill reference, present substantial new questions of patentability with respect to the claims of the '979 patent.

J. Mauldin

U.S. Patent No. 5,748,954 to Mauldin qualifies as prior art under 35 U.S.C. § 102(b) as it was published on May 5, 1998, more than one year prior to the earliest claimed priority date of the '979 patent (i.e., August 13, 2001). The Mauldin reference was not before the Examiner during original prosecution of the '979 patent, and as such, the Mauldin reference presents new teachings relative to items previously considered by the U.S. Patent and Trademark Office.

In particular, the Mauldin reference teaches a system for constructing a catalog of text files, such as Web pages, by searching for the text files on the Internet, downloading text files that are found, and processing the text files to include such information as a significant word list, an excerpt of the downloaded text file, the address (e.g., URL) of the file, the size of the file, and the number of words in

the file in a catalog. (*Appendix I at 1/8-12 and 44-51, 2/52-56*) The significant word list for a text file represents the most important words in the text file for purposes of searching. (*Appendix I at 2/57-63*)

Mauldin teaches that the significant words (e.g., keywords) of a text file can be identified by calculating a term frequency-inverse document frequency measure for every word in the downloaded text file. (*Appendix I at 7/1-4*). The words are ranked by their term frequency-inverse document frequency measure and only those words having the highest term frequency-inverse document frequency measure are kept as the significant words for the text file/document. (*Appendix I at 7/4-12*)

This disclosure teaches items deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims. As such, these teachings are, by definition, not cumulative of items previously considered by the U.S. Patent and Trademark Office. Moreover, as set forth in the SNQ table presented in Section I of this Request, these teachings, and thus, the Mauldin reference, present a substantial new question of patentability with respect to claim 9 of the '979 patent.

V. DETAILED EXPLANATION UNDER 37 C.F.R. § 1.915(B)(1)

Pursuant to 37 C.F.R. § 1.915, a detailed explanation of the pertinence and manner of applying the prior art reference to claims 1-20 of the '979 patent is provided below and in claim charts submitted herewith as Appendices M to Z. Throughout the discussion below, Requestor utilizes italicization for claim limitations.

A. Proposed Rejections

1. Claims 1, 5, 10, 13, 14 and 18 are anticipated by Reader
2. Claims 1, 5, 10, 13, 14 and 18 are rendered obvious by Reader taken in view of Mase
3. Claims 1, 5, 6, 10-14 and 18 are anticipated by Wieser
4. Claims 2, 7, 15 and 19 are rendered obvious by Wieser taken in view of Stibel
5. Claims 3 and 8 are rendered obvious by Wieser taken in view of Stibel and Syskill
6. Claims 1, 2, 5-7, 10-15 and 18-19 are anticipated by Rhodes
7. Claims 4, 16-17, and 20 are rendered obvious by Rhodes taken in view of Stibel
8. Claims 3 and 8 are rendered obvious by Rhodes taken in view of Syskill
9. Claim 9 is rendered obvious by Rhodes taken in view of Black, Syskill, and Mauldin
10. Claims 1, 2, 4-7 and 10-20 are rendered obvious by Black taken in view of HyPursuit
11. Claims 3 and 8 are rendered obvious by Black taken in view of HyPursuit and Syskill
12. Claims 1, 5, 6, 10-14 and 18 are rendered obvious by Black taken in view of Donaldson

13. Claims 2, 4, 7, 15-17, 19 and 20 are rendered obvious by Black taken in view of Donaldson and Stibel

14. Claims 3 and 8 are rendered obvious by Black taken in view of Donaldson, Stibel and Syskill

B. Detailed Explanation of SNQs and Rejections

1. Claims 1, 5, 10, 13, 14 and 18 are Anticipated by Reader under 35 U.S.C. § 102(e)

Requestor respectfully submits that claims 1, 5, 10, 13, 14 and 18 are Anticipated by Reader under 35 U.S.C. § 102(e). A claim chart applying Reader to these claims is submitted herewith as Appendix M.

(a) Reader anticipates independent claim 1

(i) *"A method for automatically generating a query from selected document content, comprising:"*

Reader describes an electronic search system that automatically generates queries from patent documents to find patent-relevant publications on the Internet. Specifically, Reader recognizes that the continued expansion of the Internet has made the Internet a very valuable resource for patent practitioners when attempting to find "prior art" publications to invalidate a patent.

> Patent professionals often search for publications relevant to patents. Searches typically arise in two contexts: when looking for "prior art" publications that might invalidate a patent and when looking for publications that might disclose an infringement of a patent.

(Appendix K at paragraph 0001)

> An ever-increasing number of publications are being published on the Internet, for example, "white papers" published on companies' public websites. Thus, the Internet has become a more and more important resource for patent professionals looking for publications relevant to patents.

(id. at paragraph 0002) Reader notes that practitioners have historically relied on relatively manual Internet search techniques to find patent-relevant publications on the Internet.

> However, patent professionals have for the most part relied on general Internet search techniques, such as applying keywords to general-

> purpose Internet search engines, to discover patent-relevant publications on the Internet.

(id. at paragraph 0002) Reader proposes a system that automates the searching process for such "prior art" publications by enabling a user to select the content of a patent (e.g., the full text of the patent or one or more claims of the patent) and then, using a computer, automatically generates a query from the selected patent content to find relevant publications on the Internet.

> The present invention provides a highly automated search technique for discovering patent-relevant publications on the Internet. The high level of automation may be achieved with the expedient of a search client resident on an end-user station that initiates linked searches for patent data and Internet publication data in a manner transparent to a user. From the user's perspective, a patent-identifying attribute, such as an inventor name, assignee name or patent number, input on an end-user station automatically returns Internet publication data, such as Uniform Resource Locators (URLs) of Web documents. The invention thereby allows a user to find patent-relevant publications on the Internet by merely inputting a patent-identifying attribute. A patent-identifying attribute may be a patent family-identifying attribute, such as an inventor name or assignee name. Or a patent identifying-attribute may be a single patent-identifying attribute, such as a patent number. Or a patent identifying-attribute may be a patent claim-identifying attribute, such as a patent claim number. A basic method for finding patent-relevant documents published on the Internet in accordance with the present invention comprises the steps of: inputting a patent-identifying attribute on an end-user station; identifying patent data from the patent-identifying attribute; identifying Internet publication data from the patent data; and outputting the Internet publication data on the end-user station.

(id. at paragraph 0004) Specifically, the search client 314 of the end user station 310 of Reader's system automatically generates a query for related web documents from the content of one or more selected patent documents (i.e., from the full text of the patent document or from one or more of the claims of the patent document). The content of the patent documents is selected by providing Reader's system with Patent Identifying (PI) attributes (e.g., patent number and optionally claim numbers, inventor names, and assignee names). The Reader system uses the PI attributes to find the corresponding patent language and then automatically abstracts the patent language (e.g., extracts keywords from the text of the selected patent content) to identify Web Document Identifying (WDI) attributes that are then used to form a query for related Web documents. (*Appendix K at paragraph 0015, corresponding to operations 405-440 of Fig. 4*)

(ii) *"defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;"*

Reader contemplates categorizing documents, specifically, company web pages and patents, in accordance with the patent classification scheme setup by the U.S. patent office or an international patent office. As is well-known, the patent classification scheme setup by the US patent office or by an international patent office has different labels for each class. In particular, Reader's system includes a patent server 330 that stores the patent content and associated patent classifications in patent database 332 and stores company website identifiers and associated patent classifications in website database 334.

> Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text. Entries may include full-text patents. Website database 334 has entries stored thereon associating patent classifications with company website identifiers, such as URLs of company home pages.

(id. at paragraph 0014)

> The patent classification may be a U.S. or international patent classification.

(id. at paragraph 0015)

(iii) *"automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;"*

The search client 314 of Reader's end user station 310 abstracts Web Document Identifying (WDI) attributes from patent language corresponding to the patent content selected through specification of the PI attributes. The abstraction of the WDI attributes occurs automatically from the patent content and can be accomplished by reducing the full-text of the patent content to keywords

separated by Boolean operators. Accordingly, Reader's end user station 310 automatically identifies a set of keywords in the selected patent content for searching additional information related to those keywords.

> Abstraction of Web document-identifying attributes from the patent language search result may be accomplished by any of numerous algorithms well known in the art. Abstraction may involve, for example, reduction of a full-text patent claim to keywords separated by Boolean operators, which keywords and operators may be selected taking into account the syntactic and lexico-semantic interdependency of the words (i.e,. context) of the full-text claim.

(Appendix K at paragraph 0013)

> Search client 314 extracts a company website identifier from the CW [Company Website] search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435)

(id. at paragraph 0015, corresponding to operation 435 of FIG. 4).

(iv) *"automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content;"*

The search client 314 of the end user station 310 submits to the patent server 330 a selection of patent content by submitting a PI search query with the PI attributes. Accordingly, the document content that is selected by the user may be the full-text of the patent(s) or the claims of the patent(s) specified by the PI attributes. In response to the PI search query, the patent server 330 returns to the end user station 310 the patent language (i.e., patent text) corresponding to the selected content AND the associated patent classifications. In providing the patent classifications associated with the selected patent content in addition to the patent language of the selected patent content, the patent server 330 is automatically categorizing the selected document content using the US or international patent classification.

> A user of end-user station 310 inputs at least one patent-identifying (PI) attribute on user interface 312 (405). Search client 314 forms a patent-identifying search query using the one or more patent-identifying

attributes (410). In this regard, search client 314 forms a search query targeted, when applied to patent database 332, to retrieve a patent classification/patent language search result that includes pairs of patent classifications and patent language from one or more patents relevant to the one or more patent-identifying attributes. The patent classification may be a U.S. or international patent classification. The patent-identifying search query is transmitted via network interface 316 and network 320 from end-user station 310 to patent server 330. Patent server 330 applies the patent-identifying search query to patent database 332 to generate patent classification/patent language (PC-PL) search result (415). Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425).

(id. at paragraph 0015, corresponding to operations 405-420 of FIG. 4)

(v) *"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 the assigned classification label."*

Reader's system includes an end-user station 310 that automatically formulates a query that inspires a search for Web documents related to keywords in the selected patent content, the search being applied to ONLY Web documents that have been categorized as being in the same classification as that previously automatically assigned to the selected patent content.

Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification attribute (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425). In this regard, end-user station 310 forms a search query targeted, when applied on patent server 330, to retrieve a company website search result that includes one or more company website identifiers, such as URLs of company home pages, relevant to the patent classification attribute. End-user station 310 transmits the CWI search query to patent server 330. Patent server 330 applies the CWI search query to website database 334 to generate company website (CW) search result (430). The CW search result is

> transmitted to end-user station 310. Search client 314 extracts a company website identifier from the CW search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435). Search client 314 passes the company website identifier and WDI attributes to search agent 318 (440). Using the company website identifier and well known DNS addressing, search agent 318 contacts the appropriate one of Web hosts 340 and, using well known "Web crawler" techniques, searches the totality of full-text documents published on the associated company website for Web document language relevant to the WDI attributes (445). Upon completion of the search, search agent 318 generates a Web document (WD) search result including Web document identifiers, such as URLs, of the relevant Web documents (450). Search agent 318 passes the Web document search result to search client 314 (455). Search client 314 extracts Web document identifiers from the Web document search result (460) and outputs the Web document identifiers on user interface 312.

(id. at paragraph 0015, corresponding to operations 420-465 of Fig. 4)

(b) Reader anticipates claim 5

Claim 5 depends from independent claim 1, and further requires *"wherein the organized classification of document content is defined using a hierarchical organization."* The organized classification of document content is a patent office classification organization, which is well-known to be hierarchical (i.e., includes classes and subclasses). As taught by Reader: "[t]he patent classification may be a U.S. or international patent classification." (*Appendix K at paragraph 0015*)

(c) Reader anticipates claim 10

Claim 10 depends from independent claim 1, and further requires *"wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary."* The organized classification of document content is a patent office classification (e.g., a U.S. or international patent office classification), which is well-known to have corresponding terms (i.e., words) that correspond to each class in the organization. As taught by Reader: "[t]he patent classification may be a U.S. or international patent classification." (*Appendix K at paragraph 0015*)

(d) Reader anticipates claim 13

Claim 13 depends from claim 1, and further requires *"automatically identifying the set of entities using a service that recognizes entities of a predefined type."* The search client 314 automatically

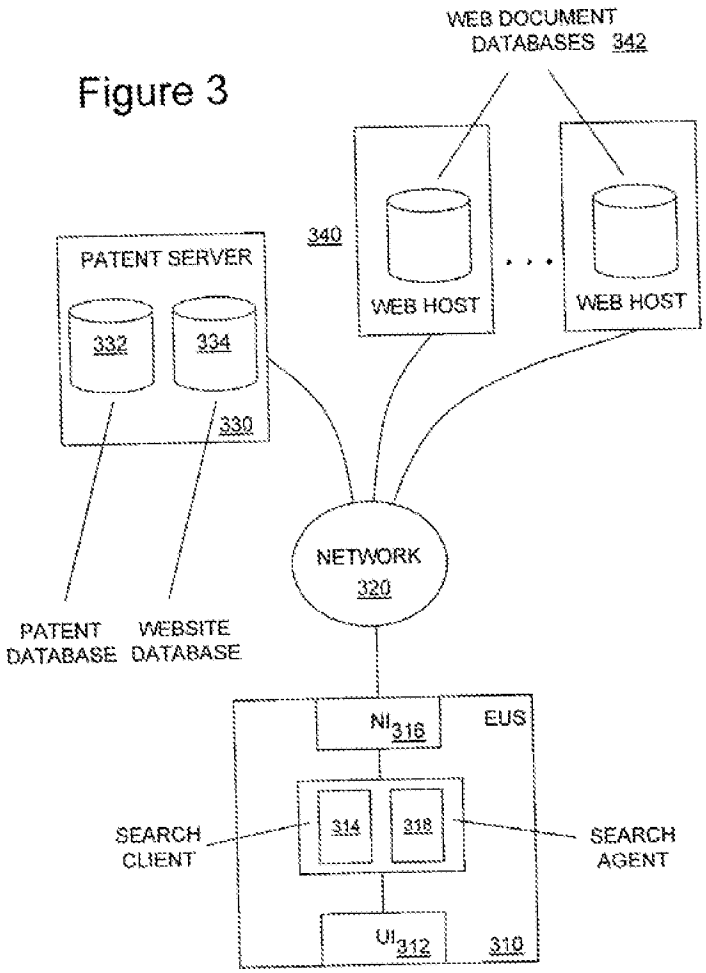
identifies the set of keywords using a service that recognizes words in a document that are keywords. That is, the search client 314 uses a service that recognizes entities of a pre-defined type, i.e., entities that are keywords, as distinguished from other entities (e.g., stop words). *(Appendix K at paragraph 0013, and shown in claim chart at Appendix M)*

(e) Reader anticipates independent claim 14

(i) *"A system for automatically generating a query from selected document content, comprising:"*

As shown in Fig. 3 of Reader (reproduced below), Reader's system includes an end user station (EUS) 310 that communicates with a patent server 330 to automatically generate a query for selected patent content to find Web documents relevant to the selected patent content.

Figure 3



The end-user station 310 may be a personal computer or workstation and the patent server 330 may be a computer server accessible to the workstation via a LAN or WAN 320.

> Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text. Entries may include full-text patents. Website database 334 has entries stored thereon associating patent classifications with company website identifiers, such as URLs of company home pages. In this regard, website database 334 may have entries for various companies associating the home page URLs of such companies with patent classifications in which such companies hold patents. Web hosts 340 are "Web hosting" servers hosting company websites addressable using DNS or IP addressing schemes well known in the art. Resident on Web hosts 340 are respective Web document databases 342 having stored thereon full-text Web documents associated with company websites. Patent server 330 is also addressable by end-user station 310 using DNS or IP addressing schemes well known in the art.

(*Appendix K at paragraph 0014*)

See also the corresponding element of claim 1.

(ii) *"an entity extractor for automatically identifying a set of entities in the selected document content for searching information related thereto using an information retrieval system;"*

See discussion of corresponding element of claim 1, where the end user station 310 (which is a personal computer or a workstation) taught by Reader performs the recited function and, thereby, includes the recited entity extractor. *(Appendix K at paragraph 0014)*

(iii) *"a categorizer for defining an organized classification of document content with each class in the organization of content having associated therewith a classification label; each classification label corresponding to a category of information in the information retrieval system; the categorizer automatically assigning the selected document content a classification label from the organized classification of content;"*

See discussion of corresponding element of claim 1, where Reader's patent server 330 (which is a computer server accessible over a network) is a categorizer that automatically assigns the selected patent content a classification label from the organized patent document classification scheme. *(Appendix K at paragraph 0014)*

(iv) *"a query generator for 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 the assigned classification label."*

See discussion of corresponding element of claim 1, where Reader's patent server 330 (which is a computer server accessible over a network) in combination with Reader's end user station 310 (which is a workstation or personal computer) are the query generator. *(Appendix K at paragraph 0014)*

(f) Reader anticipates independent claim 18

(i) *"An article of manufacture for use in a computer system, comprising: a memory;"*

See discussion of corresponding element of claim 1, where Reader discloses an EUS 310 (which is a workstation or personal computer) and a patent server 330 (which is a computer server accessible over a network) that together automatically generate a query and perform an automatic classification of selected patent content. In performing the various operations involved in automatically generating the query and automatically classifying selected patent content, the EUS 310 and the patent server 330 would necessarily have to store program instructions in a memory for execution. *(Appendix K at paragraph 0014)*

(ii) *"instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:"*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's EUS 310 and in a memory of Reader's patent server 330. *(Appendix K at paragraph 0014)*

(iii) *"defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;"*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's patent server 330. *(Appendix K at paragraph 0014)*

(iv) *"automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;"*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's EUS 310. *(Appendix K at paragraph 0014)*

(v) *"automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content; and"*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's patent server 330. *(Appendix K at paragraph 0014)*

(vi) *"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 the assigned classification label."*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's EUS 310 and in a memory of Reader's patent server 330. *(Appendix K at paragraph 0014)*

2. Claims 1, 5, 10, 13, 14 and 18 are unpatentable over Reader in view of Mase under 35 U.S.C. § 103(a)

To the extent that the Reader reference is considered not to disclose or suggest an automatic categorization of a selected patent document, Requestor respectfully submits that the combination of the Reader and Mase references renders obvious each of claims 1, 5, 10, 13, 14 and 18 under 35 U.S.C. § 103(a). A claim chart applying Reader and Mase to these claims is submitted herewith as Appendix N.

(a) Reader and Mase render obvious independent claim 1

(i) *"A method for automatically generating a query from selected document content, comprising:"*

Reader describes an electronic search system that automatically generates queries from patent documents to find patent-relevant publications on the Internet. Specifically, Reader recognizes that the continued expansion of the Internet has made the Internet a very valuable resource for patent practitioners when attempting to find "prior art" publications to invalidate a patent.

> Patent professionals often search for publications relevant to patents. Searches typically arise in two contexts: when looking for "prior art" publications that might invalidate a patent and when looking for publications that might disclose an infringement of a patent.

(Appendix K at paragraph 0001)

> An ever-increasing number of publications are being published on the Internet, for example, "white papers" published on companies' public websites. Thus, the Internet has become a more and more important resource for patent professionals looking for publications relevant to patents.

(id. at paragraph 0002) Reader notes that practitioners have historically relied on relatively manual Internet search techniques to find patent-relevant publications on the Internet.

> However, patent professionals have for the most part relied on general Internet search techniques, such as applying keywords to general-

> purpose Internet search engines, to discover patent-relevant publications on the Internet.

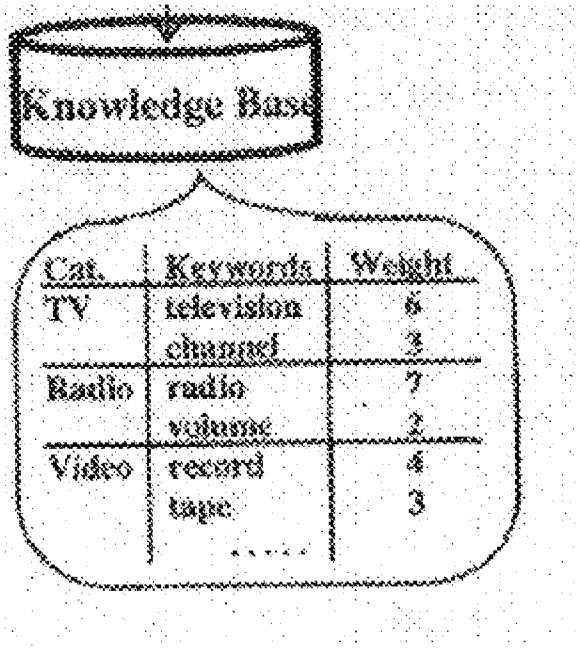
(id. at paragraph 0002) Reader proposes a system that automates the searching process for such "prior art" publications by enabling a user to select the content of a patent (e.g., the full text of the patent or one or more claims of the patent) and then, using a computer, automatically generates a query from the selected patent content to find relevant publications on the Internet.

> The present invention provides a highly automated search technique for discovering patent-relevant publications on the Internet. The high level of automation may be achieved with the expedient of a search client resident on an end-user station that initiates linked searches for patent data and Internet publication data in a manner transparent to a user. From the user's perspective, a patent-identifying attribute, such as an inventor name, assignee name or patent number, input on an end-user station automatically returns Internet publication data, such as Uniform Resource Locators (URLs) of Web documents. The invention thereby allows a user to find patent-relevant publications on the Internet by merely inputting a patent-identifying attribute. A patent-identifying attribute may be a patent family-identifying attribute, such as an inventor name or assignee name. Or a patent identifying-attribute may be a single patent-identifying attribute, such as a patent number. Or a patent identifying-attribute may be a patent claim-identifying attribute, such as a patent claim number. A basic method for finding patent-relevant documents published on the Internet in accordance with the present invention comprises the steps of: inputting a patent-identifying attribute on an end-user station; identifying patent data from the patent-identifying attribute; identifying Internet publication data from the patent data; and outputting the Internet publication data on the end-user station.

(id. at paragraph 0004) Specifically, the search client 314 of the end user station 310 of Reader's system automatically generates a query for related web documents from the content of one or more selected patent documents (i.e., from the full text of the patent document or from one or more of the claims of the patent document). The content of the patent documents is selected for which related web documents are sought by providing Reader's system with Patent Identifying (PI) attributes (e.g., patent number and optionally claim numbers, inventor names, and assignee names). The Reader system uses the PI attributes to find the corresponding patent language and then automatically abstracts the patent language (e.g., extracts keywords from the text of the selected patent content) to identify Web Document Identifying (WDI) attributes that are then used to form a query for related Web documents. (*Appendix K at paragraph 0015, corresponding to operations 405-440 of Fig. 4*)

(ii) *"defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;"*

Mase describes a system called FLUTE that defines an organized classification of document content. Specifically, FLUTE automatically categorizes patent documents (e.g., new patent applications) using a patent classification scheme that is stored in a classification Knowledge Base. As shown in Fig. 1 of Mase (relevant part reproduced below), each class in the scheme has associated with it a label (e.g., "TV", "Radio" and "Video") and corresponds to a category of patent documents in an information retrieval system used by a patent office to find patent documents.



The example knowledge base shown in Fig. 1 of Mase stores an organized classification of document content that includes the category labeled "TV" that has associated with it the keywords "television" and "channel," the category labeled "Radio" that has associated with it the keywords "radio" and "volume," and the category labeled "Video" that has associated with it the keywords "record" and "tape." (*Appendix L at Fig. 1 and page 378*)

Mase's processing of a new patent application results in the new patent application being automatically categorized as being within one of the classes in the organized classification of document

content for subsequent retrieval by an information retrieval system of a patent office for reviewing, searching and citing patent documents.

Specifically, Mase's system defines an organized classification of patent document content stored in a knowledge base.

> This paper describes keywords-based patent categorization using our text classification support tool called FLUTE and discusses a simulation study applied to 154,000 patents. FLUTE automatically generates a classification knowledge base from sample patent texts. Experimental simulation results show that FLUTE is powerful enough to support the patent classification work of indexing experts.

(Appendix L at Abstract)

> (1) *Automatic knowledge base initialization*
> FLUTE is able to obtain word statistics from electronic documents. It first derives keyword candidates for each document. Then, it removes stopwords which are obviously not keywords ("thing", "is", etc.) and common words appearing over every category ("invention and "patent" appear in all patent documents). Next, it identifies keywords for each category by applying weights. Finally, it generates classification rules, which include a certain factor. Since the KB structure is simple, a person can verify and modify the knowledge base.

(*Appendix L at page 378, column 1*)

Mase's system receives a new patent application and automatically categorizes the application into a class in the organized classification of patent document content. The class corresponds to a category of patent documents in an information retrieval system used by a patent office to find patent documents.

> At present most patents are applied electronically. If an intelligent system could read a patent, recognize its purpose, and finally categorize it, the patent management process could be restructured. To classify patents into appropriate categories, the system would have to have a powerful knowledge base, which is difficult to build.
>
> We have been researching automatic patent categorization as a submitted research from Industrial Property Cooperation Center. This paper presents our text classification support tool called FLUTE (Mase, et al. 1996a), the customization of FLUTE for patent categorization and its evaluation using 154,000 patents

(*Appendix L at page 377, column 2*)

> Obtaining patents is important for any organization that needs to maintain intellectual propery rights. In order to manage the review, search, and citation of patents at the patent office efficiently, patent applications should be appropriately indexed as soon as possible.

(*Appendix L at page 377, column 1*)

> (2) *Classification certainty*
> FLUTE presents three kinds of classification results according to values corresponding to the degree of confidence (certainty). If one category's certainty is high, FLUTE presents a unique solution. If the certainties of more than two categories are high, FLUTE offers alternatives.
> Otherwise FLUTE offers no solution and asks an expert to intervene.

(Appendix L at page 378, columns 1 and 2) It would have been obvious to one of ordinary skill in the relevant art to have integrated features taught by Mase into Reader's system to arrive at what is claimed. In particular, Mase teaches a system that may be used by a patent office to automatically categorize received patent applications using a patent classification scheme. Mase indicates that such an automatic categorization is desirable to "manage the review, search, and citation of patents at the patent office efficiently". (*Appendix L at page 377, column 1)*

A person of ordinary skill in the art would have readily recognized that Reader's system also could be used to manage the review of patent applications at a patent office more efficiently. Specifically, a person of ordinary skill in the art would have recognized that Reader's techniques could be advantageously combined with Mase's techniques to achieve a system that would provide an examiner at the patent office with both a patent application that is automatically classified in accordance with Mase's teachings and Web documents automatically determined as potential prior art relevant to the patent application in accordance with Reader's teachings. Such a combined system would make the review of patent applications at the patent office more efficient by decreasing manual indexing efforts in accordance with Mase's express teachings and by decreasing manual Internet searches for prior art for new patent applications in accordance with Reader's express teachings. Moreover, Mase expressly suggests such a combined system by identifying the related work of Yoden: "Yoden developed a patent search system, which supports finding relevant patents to an inputted one. It extracts weighted keywords...it then compares these keywords with those of [other] patents." (*Appendix L at page 382, column 1*) Thus, Mase expressly contemplates combining the disclosed patent

classification system with a patent search system akin to that of Reader (i.e., a patent search system that supports finding documents related to an input patent by extracting keywords from the patent in question).

Furthermore, the combination of Mase's teachings with those of Reader would involve merely combining and/or substituting known prior art elements to yield predictable results. One skilled in the art could have easily combined the known elements as claimed by known methods, with each element in the combination performing the same function as it does separately, and the combination yielding nothing more than predictable results to one skilled in the relevant art. Among other features, the disclosure of the Reader and Mase references teach items deemed missing from prior art during original prosecution, and responsible for Examiner allowance of the claims.

(iii) *"automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;"*

The search client 314 of Reader's end user station 310 abstracts Web Document Identifying (WDI) attributes from patent language corresponding to the patent content selected through specification of the PI attributes. The abstraction of the WDI attributes occurs automatically from the patent content and can be accomplished by reducing the full-text of the patent content to keywords separated by Boolean operators. Accordingly, Reader's end user station 310 automatically identifies a set of keywords in the selected patent content for searching additional information related to those keywords.

> Abstraction of Web document-identifying attributes from the patent language search result may be accomplished by any of numerous algorithms well known in the art. Abstraction may involve, for example, reduction of a full-text patent claim to keywords separated by Boolean operators, which keywords and operators may be selected taking into account the syntactic and lexico-semantic interdependency of the words (i.e,. context) of the full-text claim.

(Appendix K at paragraph 0013)

> Search client 314 extracts a company website identifier from the CW [Company Website] search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435)

(id. at paragraph 0015, corresponding to operation 435 of FIG. 4).

(iv) *"automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a classification label from the organized classification of content;"*

Mase's system (FLUTE) receives a new patent application and automatically categorizes the new patent application using the organized classification of patent document content for assigning the selected document a classification label from the organized classification of document content.

Specifically, Mase asserts that new patent applications "should be appropriately indexed as soon as possible" (i.e., assigned to their appropriate categories/classes in the patent classification scheme as soon as possible) "to manage the review, search and citation of patents at the patent office efficiently." However, classifying the patent applications quickly has proven difficult due to the large number of patent applications received by a patent office and the time intensive nature of the manual classification process currently performed by indexing experts at the patent office.

> Obtaining patents is important for any organization that needs to maintain intellectual property rights. In order to manage the review, search, and citation of patents at the patent office efficiently, patent applications should be appropriately indexed as soon as possible. However, there are problems, as follows:
> - There are approximately 3,000 patent categories. This makes it impossible for any indexing expert to perform categorization work over every technical field.
> - In Japan, there were 370,000 patent applications in 1994. Despite such a large number, it is difficult to increase the number of indexing experts.
> - Most patents include over 5,000 words. In Japan, patent categorization requires experts to read all documents, which is time consuming work.
> - To maintain an applicant's proprietary information, only persons entrusted with confidentiality can be assigned to indexing work. At present, most patents are applied electronically. If an intelligent system could read a patent, recognize its purpose, and finally categorize it, the patent management process could be restructured. To classify patents into appropriate categories, the system would have to have a powerful knowledge base, which is difficult to build.

(Appendix L at page 377, columns 1 and 2) To address this problem, Mase proposes an automated patent classification system that employs a classification tool called "FLUTE."

> We have been researching automatic patent categorization as a submitted research from Industrial Property Cooperation Center. This paper presents our text classification support tool called FLUTE (Mase, et al. 1996a), the customization of FLUTE for patent categorization, and its evaluation using 154,000 patents.

(Appendix L at page 377, column 2) As shown in Fig. 1 of Mase (reproduced below), FLUTE includes a Keywords Extraction Engine that receives new patent documents (e.g., new patent applications) and extracts keywords from the text of the new patent documents. FLUTE also includes a Classification Engine that classifies the new patent documents based on the extracted keywords using the organized classification of patent document content stored in the Knowledge Base.

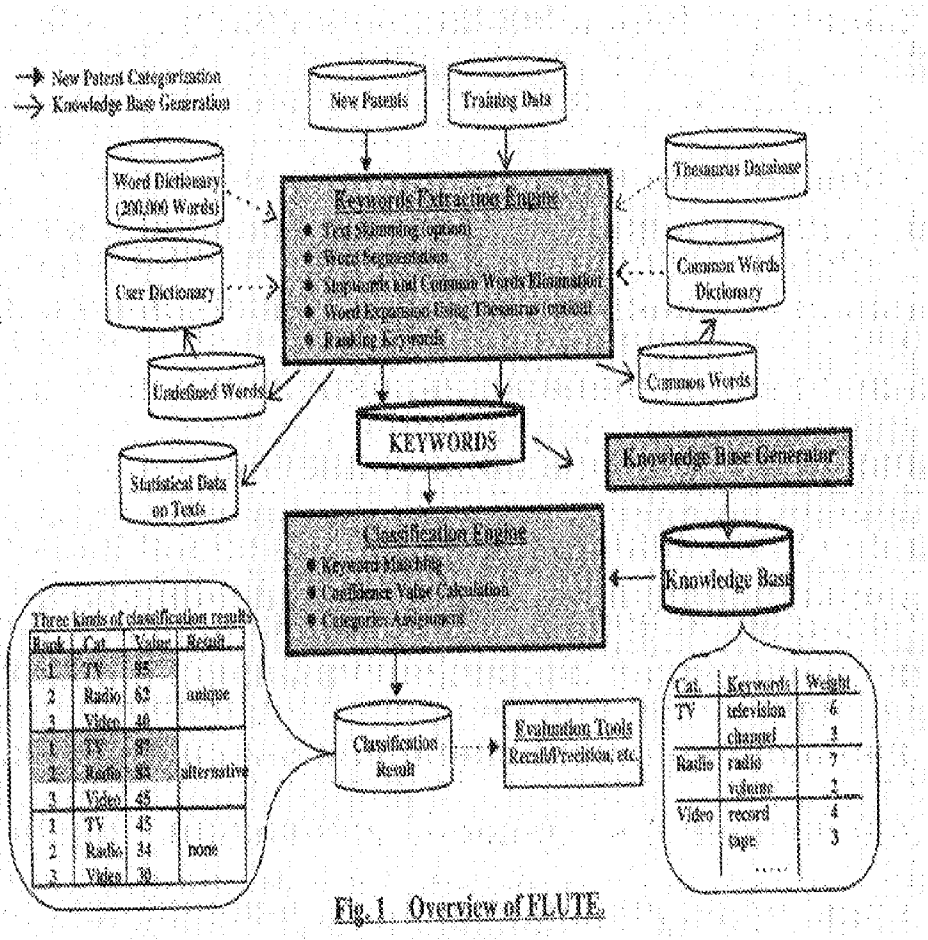


Fig. 1 Overview of FLUTE.

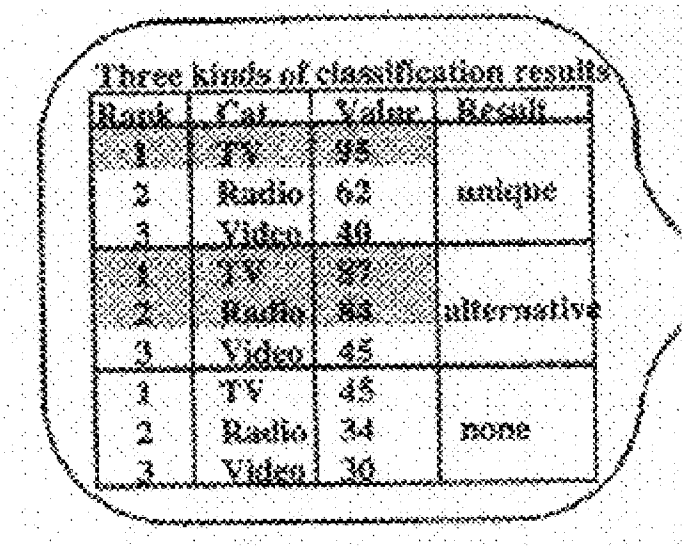
In particular, the Classification Engine automatically determines a category/class for a new patent document from the patent classification scheme stored in the knowledge base by comparing the keywords extracted from the text of the new patent document to the keywords that characterize the classes/categories in the classification scheme.

> The classification engine calculates the similarity between inputted text and each category by comparing the keywords extracted from the inputted text with those stored in the KB.

(Appendix L at page 379, column 1) The Classification Engine produces three kinds of classification results for a new patent document that correspond to different degrees of certainty as to the match between the new patent document and the corresponding classes/categories. If the certainty of only one category/class is high, the classification is deemed unique and the category/class is automatically assigned to the new patent document. If the certainties of two or more categories/classes are high, the classes/categories are automatically assigned as alternatives to the new patent document. If no certainty for any class/category is high, then no class/category is assigned and the new patent document is referred to an indexing expert for classification.

> (2) *Classification certainty*
> FLUTE presents three kinds of classification results according to values corresponding to the degree of confidence (certainty). If one category's certainty is high, FLUTE presents a unique solution. If the certainties of more than two categories are high, FLUTE offers alternatives.
> Otherwise FLUTE offers no solution and asks an expert to intervene.

(Appendix L at page 378, columns 1 and 2) Fig. 1 (relevant part reproduced below) shows an example of automatic classification results produced by FLUTE for three different new patent documents.



The first new patent document has been categorized uniquely as being within the category labeled "TV" and, therefore, has been assigned the classification label "TV." The second new patent document has been categorized as being within the two alternative categories "TV" and "Radio" and, therefore, has been assigned the classification labels "TV" and "Radio." The FLUTE system was unable to automatically categorize the third new patent document and, therefore, the third new patent document is referred to an indexing expert for classification.

In sum, Mase teaches a system that may be used by a patent office to automatically categorize received patent applications using a patent classification scheme. Mase indicates that such an automatic categorization is desirable to "manage the review, search, and citation of patents at the patent office efficiently". (A*ppendix L at page 377, column 1*) A person of ordinary skill in the art would have readily recognized that Reader's system also could be used to manage the review of patents at a patent office more efficiently.

As stated previously, a person of ordinary skill in the art would have recognized that Reader's techniques could be advantageously combined with Mase's techniques to achieve a system that would provide an examiner at the patent office with both a patent application that is automatically classified in accordance with Mase's teachings and Web documents automatically determined as potential prior art relevant to the patent application in accordance with Reader's teachings.

(v) *"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 the assigned classification label."*

Reader's system, when modified based on Mase's teachings as stated above, includes an end-user station 310 that automatically formulates a query that inspires a search for Web documents related to keywords in a selected new patent application received by a patent office, the search being applied to ONLY Web documents that have been categorized as being in the same classification as that previously automatically assigned to the selected new patent application in accordance with Mase's teachings. The newly categorized patent application and the results of the search could then be provided efficiently and in an automated fashion to an examiner that has been assigned by the patent office to review the new patent application.

Specifically, the search client 314 of the end user station 310 of Reader's system receives, from the patent server 330, the classification information generated automatically from the new patent

application in accordance with Mase's teaches. The search client 314 uses the classification information to form a company Website identifying search query that is submitted to the patent server 330. In response to the query, the patent server 330 uses the classification information to retrieve one or more company website identifiers (e.g., home page URLs of web sites) that correspond to the classification information and returns the company website identifiers to the search client 314. The search client 314 extracts the company web site identifiers from the search result received from the patent server 330 and passes the company web site identifiers and the web document identifying attributes abstracted from the patent language of the selected patent content to the search agent 318 as a "search query." The search agent 318 then uses "Web crawler" techniques to search the totality of the full-text of the documents published at the company websites corresponding to the classification information for document language relevant to the WDI attributes. The search agent 318 then passes the Web document search results to the search client 314, which outputs corresponding Web document identifiers on user interface 312.

> Patent server 330 transmits the patent classification/patent language search result to end-user station 310. End-user station 310, particularly search client 314, extracts a patent classification attribute (PC) attribute from the patent classification portion of the PC-PL search result (420) and forms a company website-identifying (CWI) search query using the patent classification attribute (425). In this regard, end-user station 310 forms a search query targeted, when applied on patent server 330, to retrieve a company website search result that includes one or more company website identifiers, such as URLs of company home pages, relevant to the patent classification attribute. End-user station 310 transmits the CWI search query to patent server 330. Patent server 330 applies the CWI search query to website database 334 to generate company website (CW) search result (430). The CW search result is transmitted to end-user station 310. Search client 314 extracts a company website identifier from the CW search result and abstracts Web document-identifying (WDI) attributes from the patent language portion of the PC-PL search result (435). Search client 314 passes the company website identifier and WDI attributes to search agent 318 (440). Using the company website identifier and well known DNS addressing, search agent 318 contacts the appropriate one of Web hosts 340 and, using well known "Web crawler" techniques, searches the totality of full-text documents published on the associated company website for Web document language relevant to the WDI attributes (445). Upon completion of the search, search agent 318 generates a Web document (WD) search result including Web document identifiers, such as URLs, of the relevant Web documents (450). Search agent 318 passes the Web document search result to search client 314 (455).

> Search client 314 extracts Web document identifiers from the Web document search result (460) and outputs the Web document identifiers on user interface 312.

(*Appendix K at paragraph 0015, corresponding to operations 420-465 of Fig. 4*)

(b) Reader and Mase render obvious claim 5

Claim 5 depends from independent claim 1, and further requires *"wherein the organized classification of document content is defined using a hierarchical organization."* The organized classification of document content is a patent office classification organization, which is well-known to be hierarchical (i.e., includes classes and subclasses). As taught by Reader: "[t]he patent classification may be a U.S. or international patent classification." (*Appendix K at paragraph 0015*)

(c) Reader and Mase render obvious claim 10

Claim 10 depends from independent claim 1, and further requires *"wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary."* The organized classification of document content is a patent office classification (e.g., a U.S. or international patent office classification), which is well-known to have corresponding terms (i.e., words) that correspond to each class in the organization. As taught by Reader: "[t]he patent classification may be a U.S. or international patent classification." (*Appendix K at paragraph 0015*)

Moreover, Mase expressly discloses that the classes /categories in the organized classification of document content have characteristic keywords used to characterize the categories for purposes of automatic classification of patent documents. (*Appendix L at page 377, column 2 and page 379, column 1, all of which is in the claim chart at Appendix N*)

(d) Reader and Mase render obvious claim 13

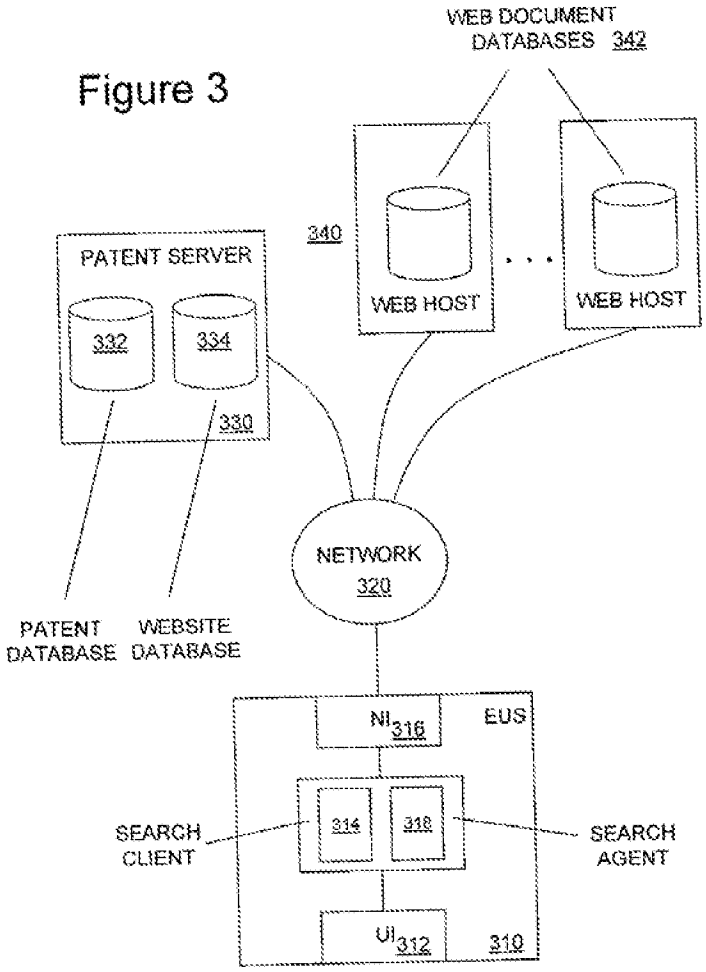
Claim 13 depends from claim 1, and further requires *"automatically identifying the set of entities using a service that recognizes entities of a predefined type."* The search client 314 automatically identifies the set of keywords using a service that recognizes words in a document that are keywords. That is, the search client 314 uses a service that recognizes entities of a pre-defined type, i.e., entities that are keywords, as distinguished from other entities (e.g., stop words). *(Appendix K at paragraph 0013, and shown in claim chart at Appendix N)*

(e) Reader and Mase render obvious independent claim 14

(i) *"A system for automatically generating a query from selected document content, comprising:"*

As shown in Fig. 3 of Reader (reproduced below), Reader's system includes an end user station (EUS) 310 that communicates with a patent server 330 to automatically generate a query for selected patent content to find Web documents relevant to the selected patent content.

Figure 3



The end-user station 310 may be a personal computer or workstation and the patent server 330 may be a computer server accessible to the workstation via a LAN or WAN 320.

> Turning now to FIG. 3, a communication system in which the present invention is operative in accordance with a second embodiment is shown. The communication system includes an end-user station (EUS) 310, such as a personal computer or workstation, having a user

> interface (UI) 312, a processor-implemented search client 314 and search agent 318 and a network interface (NI) 316. Search client 314 and search agent 318 are software applications. End-user station 310 has access to patent server 330 and Web hosts 340 via network 320 that may include local area networks (LANs) and wide area networks (WANs). Patent server 330 has patent database 332 and website database 334 resident thereon. Patent database 332 has entries stored thereon associating patent-identifying attributes, such as inventor names, assignee names and patent numbers, with patent classifications and patent language, such as patent claim text. Entries may include full-text patents. Website database 334 has entries stored thereon associating patent classifications with company website identifiers, such as URLs of company home pages. In this regard, website database 334 may have entries for various companies associating the home page URLs of such companies with patent classifications in which such companies hold patents. Web hosts 340 are "Web hosting" servers hosting company websites addressable using DNS or IP addressing schemes well known in the art. Resident on Web hosts 340 are respective Web document databases 342 having stored thereon full-text Web documents associated with company websites. Patent server 330 is also addressable by end-user station 310 using DNS or IP addressing schemes well known in the art.

(*Appendix K at paragraph 0014*)

Mase describes its FLUTE system being implemented by a Computer Workstation 3050RX/330T to generate the Knowledge Base and perform the classification/categorization simulations.

> We used the Computer Workstation 3050RX/330T to generate the KB and do these simulations. It took approximately 8 seconds for FLUTE to extract keywords from a document and to categorize it

(Appendix L at page 382, column 1) In the system resulting from the combination of Reader's and Mase's teachings contemplated above, a patent server would perform the functions of Reader's patent server 330 and Mase's FLUTE tool for new patent applications. Accordingly, the FLUTE tool aspects of the server would be used to automatically determine the classification for a new patent application and the aspects of the server similar to Reader's patent server 330 would operate in combination with the EUS 310 to automatically perform a search for patent-relevant publications on the Internet to be automatically provided to a patent office examiner when the new patent application is assigned to the examiner for review.

(ii) *"an entity extractor for automatically identifying a set of entities in the selected document content for searching*

information related thereto using an information retrieval system;"

See discussion of corresponding element of claim 1, where the end user station 310 (which is a personal computer or a workstation) taught by Reader performs the recited function and, thereby, includes the recited entity extractor. *(Appendix K at paragraph 0014)*

(iii) *"a categorizer for defining an organized classification of document content with each class in the organization of content having associated therewith a classification label; each classification label corresponding to a category of information in the information retrieval system; the categorizer automatically assigning the selected document content a classification label from the organized classification of content;"*

See discussion of corresponding element of claim 1, where Mase's Classification Engine implemented on Mase's workstation is a categorizer that automatically assigns the selected new patent application a classification label from the organized patent document classification scheme. *(Appendix L at page 382, column 1)*

(iv) *"a query generator for 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 the assigned classification label."*

See discussion of corresponding element of claim 1, where Reader's patent server 330 (which is a computer server accessible over a network) in combination with Reader's end user station 310 (which is a workstation or personal computer) are the query generator. *(Appendix K at paragraph 0014)*

(f) Reader and Mase render obvious independent claim 18

(i) *"An article of manufacture for use in a computer system, comprising: a memory;"*

See discussion of corresponding element of claim 1, where Reader discloses an EUS 310 (which is a workstation or personal computer) and a patent server 330 (which is a computer server accessible over a network) that together automatically generate a query, and Mase discloses a workstation that performs automatic classification of new patent applications. In performing the various operations

involved in automatically generating the query and automatically classifying a new patent application, the EUS 310, the patent server 330 and the Mase workstation would necessarily have to store program instructions in a memory for execution. *(Appendix K at paragraph 0014; and Appendix L at page 382, column 1)*

(ii) *"instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:"*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's EUS 310 and in a memory of a patent server resulting from the combination of the teachings of Reader and Mase, which performs the query generation functions of Reader's patent server 330 and the patent classification functions of Mase's workstation. *(Appendix K at paragraph 0014; and Appendix L at page 382, column 1)*

(iii) *"defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;"*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of a patent server that results from the combination of the teachings of Reader and Mase, which performs the query generation functions of Reader's patent server 330 and the patent classification functions of Mase's workstation. *(Appendix K at paragraph 0014; and Appendix L at page 382, column 1)*

(iv) *"automatically identifying a set of entities in the selected document content for searching information related thereto using the information retrieval system;"*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's EUS 310. *(Appendix K at paragraph 0014)*

(v) *"automatically categorizing the selected document content using the organized classification of document*

content for assigning the selected document content a classification label from the organized classification of content; and"

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of a patent server that results from the combination of the teachings of Reader and Mase, which performs the query generation functions of Reader's patent server 330 and the patent classification functions of Mase's workstation. *(Appendix K at paragraph 0014; and Appendix L at page 382, column 1)*

(vi) *"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 the assigned classification label."*

See discussion of corresponding element of claim 1, where the instructions are stored in a memory of Reader's EUS 310 and in a memory of a patent server resulting from the combination of the teachings of Reader and Mase, which performs the query generation functions of Reader's patent server 330 and the patent classification functions of Mase's workstation. *(Appendix K at paragraph 0014; and Appendix L at page 382, column 1)*

3. Claims 1, 5, 6, 10-14 and 18 are Anticipated by Wieser under 35 U.S.C. § 102(a) and 35 U.S.C. § 102(e)

Requestor respectfully submits that claims 1, 5, 6, 10-14 and 18 are anticipated by Wieser under 35 U.S.C. § 102(a) and 35 U.S.C. § 102(e). A claim chart applying Wieser to these claims is submitted herewith as Appendix O.

(a) Wieser anticipates independent claim 1

(i) *"A method for automatically generating a query from selected document content, comprising:"*

Wieser teaches a system that includes a client 12 that generates a query composed of all or a portion of a document (e.g., a web page) and sends the query to a match server 14. (*Appendix C at p. 5, lines 8-12*) The query is automatically generated from the document content. Specifically:

> Another object of the present invention is to provide a system and method which **automatically** and contextually matches products, advertisements or other content (hereinafter referred to as 'offers') to the content on a web page that a user has selected in real-time"

(id. at page 2, lines 16-19) (emphasis added)

(ii) *"defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;"*

Wieser teaches clusters which function as classes.[8] Additionally, each set of clusters functions as an organized classification of document content. The Wieser system defines an organized classification of document content by grouping document content into different clusters, each cluster having an assigned cluster ID. Wieser further organizes various classes (clusters) as separate spheres in vector-space, yielding classes having associated classification labels (cluster IDs) within an organized classification, with the classification labels corresponding to categories of information in the Wieser system. Specifically:

> According to one embodiment, the present invention provides a novel approach to representing textual documents as high dimensional vectors. Such an approach provides an efficient means of indexing document collections, allowing retrieval of document (querying) based on keywords, **grouping related documents (categorization)**. Additionally, this method supports such contextual queries and document groupings

(id. at page 15, lines 8-11) (emphasis added)

> After the vector generations, the document vector, or feature vector, must be compared with database vectors, or feature vectors. A naïve approach to product matching would be to compare the document vector to the vectors for every product in the database. This process becomes burdensome as the number of products in the database grows. Fortunately, the database vectors are not smoothly distributed throughout the vector space, but rather, tend to "clump" together, leaving vast empty spaces between the clumps, or clusters as they are

[8] As stated previously, Requestor applies the Patent Owner's apparent claim construction in the related litigation without waiver. See *n.* 1.

> commonly known. For any given cluster, there is a sphere that bounds every point in the cluster. For all the various clusters, one could compute the center point and the radius of the sphere which bounds the cluster. Then, when one wishes to find the products that match a given document vector, one need only compare the document to the products in the cluster whose bounding sphere contains the document vector (or the nearest spheres, if no sphere contains the document vector). Thus, the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres.

(id. at page 17, lines 15-32)

> In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. In other words, this narrowing qualification is added to the original metadata query to form an SQL query.

(id. at page 18, lines 16-22)

(iii) *"automatically identifying a set of entities in the selected document content for searching additional information related thereto using the information retrieval system;"*

Wieser identifies keywords within a selected document and uses those keywords for searching to identify related additional information. Specifically, the contextual matching server 300 of the match server 14 generates a query context vector based on keywords within content of a source document, and Wieser uses the vector to search for related additional information. Wieser teaches that "[t]he contextual matching server 300 generates a query context vector." The algorithms used to generate the vectors "characterize documents based on the presence of keywords;" "associate vectors with these keywords;" and "form document vectors by combining the vectors of the keywords present in the document." *(id. at page 14, line 29 to page 15, line 5).*

(iv) *"automatically categorizing the selected document content using the organized classification of document content for assigning the selected document content a*

classification label from the organized classification of content;"

As described above, a cluster corresponds to a class and a set of clusters corresponds to an organized classification of document content when using the Patent Owner's apparent claim construction. Consequently, Wieser automatically categorizes a selected document (e.g., a web page) using the organized classification of document content for assigning the selected document one or more cluster IDs from the organized classification of document content. That is, the selected document is automatically categorized by the match server 14 into one or more clusters of content having corresponding cluster IDs by matching the query context vector of the selected document with the one or more spheres or clusters in vector space. Wieser teaches that "the database vectors are not smoothly distributed throughout the vector space, but rather, tend to 'clump' together, leaving vast empty spaces between the clumps, or clusters as they are commonly known." *(id. at page 17, lines 19-21).* Wieser further describes the bounding spheres: " [f]or any given cluster, there is a sphere that bounds every point in the cluster." *(id. at page 17, lines 22-23).* Wieser further teaches that "the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres." *(id. at page 17, lines 29-31).* Wieser further teaches assigning cluster IDs:

> "In accordance with an embodiment of the present invention, the query context vector is compared to the center vector of each cluster and the clusters with center vectors closest to the query context vector are selected. It is appreciated that these cluster IDs are then used to narrow the scope of products returned by the original metadata query as issued by the client 12. "

(id. at page 18, lines 16-22 (emphasis added))

(v) *"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 the assigned classification label."*

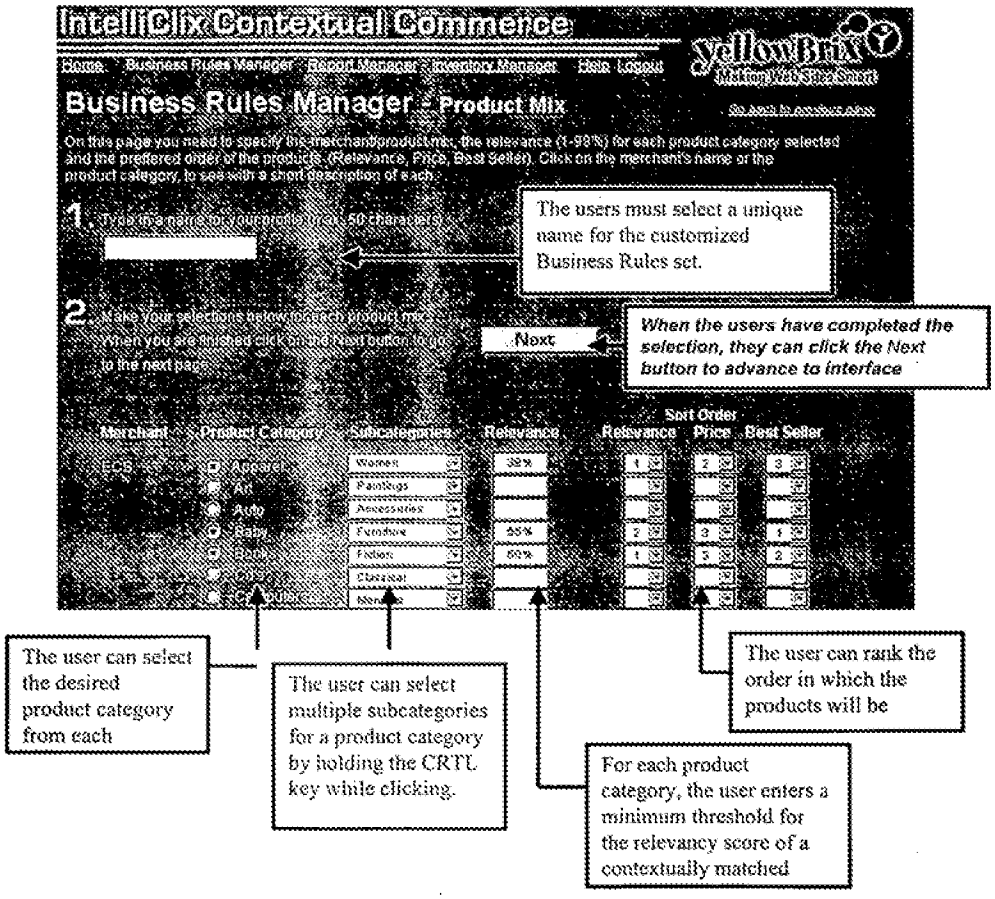
As stated previously, the matching process performed by the match server 14 automatically forms an SQL query that restricts the search of the inventory of product offers to those product offers

having product vectors within only the clusters having cluster IDs assigned to the selected document content. This restriction of the search to only those product offers having product vectors in the clusters having the assigned cluster IDs results in a less burdensome search, particularly as the inventory of offers grows. As described above, Weiser teaches that "the computation is reduced from comparing the document vector to all product vectors to simply comparing the document vector to the center vectors for the spheres, followed by comparison to the product vectors for the products in the matching spheres." *(id. at page 17, lines 29-31).* Wieser further teaches that "[t]he contextual matching server 300 then compares the query context vector to pre-determined item context vectors to narrow the search to focus on products that are most likely relevant to the selected text." *(id. at page 18, lines 14-16)*

(b) Wieser anticipates claim 5

Claim 5 depends from independent claim 1, and further requires *"wherein the organized classification of document content is defined using a hierarchical organization."* Wieser, in allowing site owners to specifically select categories and relevance scores for each category (*Appendix C at Fig. 15, reproduced below),* contemplates mapping categories and subcategories (e.g., Books->Fiction Books) to different cluster IDs for contextual matching. In doing so, if under a broadest reasonable construction standard a set of clusters is considered an organized classification of document content, Wieser defines the organized classification of content (e.g., the clusters in vector space) using a hierarchical organization.

Fig. 15



As shown in the example GUI displayed in Fig. 15 of Wieser, a user has selected the product category "Apparel" and the product subcategory "Women" for their web page/site and has indicated that a minimum relevance threshold of 38% is required for a product offer within this category/subcategory to be included in the web page/site. In setting a relevance threshold of 38% for contextual matching of web page/site content to product offers in the "Apparel" category and the "Women" subcategory, Wieser contemplates mapping a cluster ID or cluster IDs to a hierarchical organization that includes the category "Apparel" and the subcategory "Women" for similarity/relevance calculations.

(c) Wieser anticipates claim 6

Claim 6 depends from independent claim 1, and further requires *"using a text categorizer to assign the classification label assigned from the organized classification of content."* As described

above, the contextual matching server 300 of the match server 14 analyzes the text of the document (to extract keywords) to form a query context vector and then assigns cluster IDs to the query context vector by matching the query context vector to nearest clusters. Accordingly, if, under a broadest reasonable construction standard, a cluster is considered a class and a set of clusters is considered an organized classification of document content, Wieser uses a text categorizer (e.g., the contextual matching server 300) to assign a classification label (e.g., a cluster ID) from the organized classification of content (e.g., the document vectors in vector-space organized in clusters).

(d) Wieser anticipates claim 10

Claim 10 depends from independent claim 1, and further requires *"wherein each class in the organized classification of document content has associated therewith a characteristic vocabulary."* The system of Wieser, in allowing site owners to specifically select categories and relevance scores for each category (see Fig. 15, reproduced below), contemplates associating terms that describe categories (e.g., Fiction Books, Women's Apparel) to different cluster IDs. Accordingly, Wieser contemplates that each cluster ID will have associated with it one or more terms that describe the cluster or clusters corresponding to the cluster ID.

As shown in the example GUI displayed in Fig. 15 of Wieser (reproduced above), a user has selected the product category "Apparel" and the product subcategory "Women" for their web page/site and has indicated that a minimum relevance threshold of 38% is required for a product offer within this category/subcategory to be included in the web page/site. In setting a relevance threshold of 38% for contextual matching of web page/site content to product offers in the "Apparel" category and the "Women" subcategory, Wieser contemplates mapping a cluster ID or cluster IDs to the category "Apparel" and to the subcategory "Women" for similarity/relevance calculations. Accordingly, these cluster IDs will have associated with them the terms "Apparel" and "Women" that describe their contents.

(e) Wieser anticipates claim 11

Claim 11 depends from claim 10, and further requires *"ranking results from the query performed at the information retrieval system in accordance with one of the assigned classification label and the characteristic vocabulary."* The system of Wieser allows a web site owner to select a sort order of the product offers to be included in the web site owner's web page based on the candidate category or categories manually selected by the web site owner for the web page. Specifically, as shown in the GUI

of Fig. 15 of Wieser (reproduced above), the ranking of the product offerings to be added to the web page differs based on which category is ultimately assigned to the web page. That is, if women's apparel is the category ultimately assigned to the web page through the contextual matching process, the sort order will be first based on relevance, then price, and lastly best seller. In contrast, if Baby furniture is the category ultimately assigned to the web page through the contextual matching process, the sort order will be first based on best seller, then relevance, and lastly price.

Accordingly, since Wieser contemplates that cluster IDs are mapped to the product categories of Fig. 15, it can be said that the order or ranking of product offerings returned in response to the contextual query is "in accordance with" the assigned classification label (e.g., the assigned cluster ID). *(Appendix C at page 22, lines 16-23, and shown in claim chart at Appendix O*)

(f) Wieser anticipates claim 12

Claim 12 depends from claim 11, and further requires *"using the method in a system for enriching selected content of a document with personalities that identify enrichment themes."* The system of Wieser uses a matching process to incorporate offers in an existing web page. (*Appendix C at Figs. 2A and 2B* -which show a browser 22B/22C showing a Web page in which is inserted matched return data 30, 36 corresponding to offers).

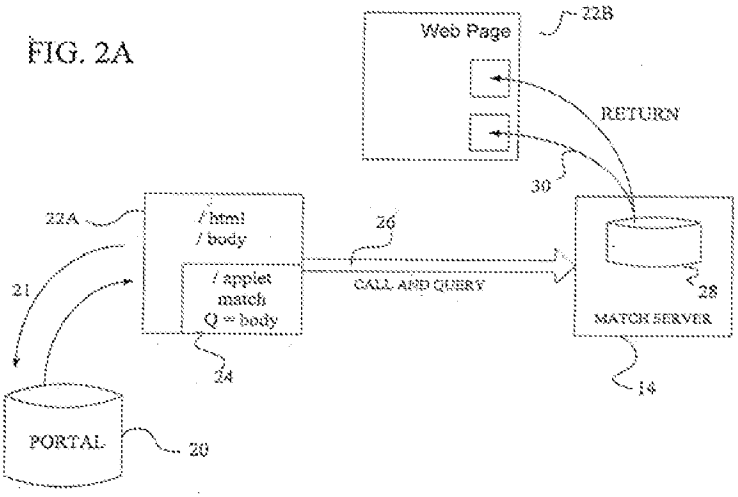
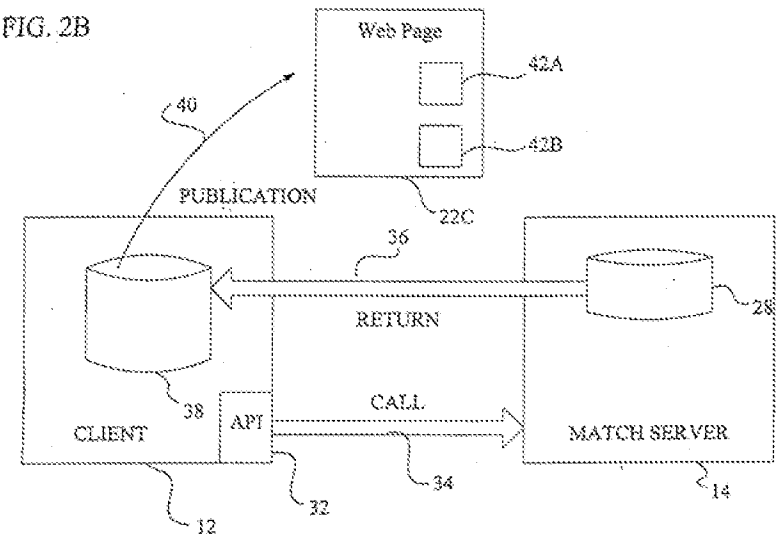


FIG. 2B



Accordingly, Wieser's system enriches documents by creating a reference between a region of the document (i.e., represented by the squares in the web page shown in Figs. 2A and 2B) and a text segment and/or images corresponding to an offer. The enriching is in accordance with personalities that identify enrichment themes in that the offers added to the web page can be offers for products that are contextually matched (i.e., based on all or a portion of the document content) or offers for products that are related to entities extracted from the document (i.e., based on one or more names of persons, companies, organizations, places, products or the likes). (*Appendix C at page 5, line 22 to page 6, line 15; page 8, lines 7-15; page 9, lines 9-16*)

Requestor notes that the '979 patent states that to "enrich" a document means merely to "annotate a document in accordance with a predefined personality." (*Appendix A at col. 6, lines 63 and 64*)

(g) Wieser anticipates claim 13

Claim 13 depends from claim 1, and further requires *"automatically identifying the set of entities using a service that recognizes entities of a predefined type."* The contextual matching server 300, taught by Wieser, automatically identifies the words in the selected document using a service that recognizes entities of a predefined type – e.g., words in the selected text document that are not stop words. *(Appendix C at page 15, lines 4-5 and 11-20, and shown in claim chart at Appendix O)*

(h) Wieser anticipates independent claim 14

(i) *"A system for automatically generating a query from selected document content, comprising:"*

Wieser describes a system "that is readily implemented by presently available communication apparatus and electronic components. The invention finds ready application in virtually all commercial communications and/or computer networks including but not limited to world wide web (Internet), intranet, local area network (LAN), wide area network (WAN), wireless network and wired cable transmission systems." (*Appendix C at page 4, line 26 to page 5, line 1*) As stated above with respect to claim 1, the client 12 of Wieser's system generates a query composed of all or a portion of a document (e.g., a web page) and sends the query to a match server 14. The client 12 and the match server 14 are readily implemented by available electronic components.

(ii) *"an entity extractor for automatically identifying a set of entities in the selected document content for searching information related thereto using an information retrieval system;"*

As described above with reference to claim 1, the match server 14 of Wieser includes electronic components that perform this function.

(iii) *"a categorizer for defining an organized classification of document content with each class in the organization of content having associated therewith a classification label; each classification label corresponding to a category of information in the information retrieval system; the categorizer automatically assigning the selected document content a classification label from the organized classification of content;"*

As described above with reference to claim 1, the match server 14 of Wieser includes electronic components that perform this function.

(iv) *"a query generator for 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 the assigned classification label."*

As described above with reference to claim 1, the match server 14 of Wieser includes electronic components that perform this function.

(i) Wieser anticipates independent claim 18

(i) *"An article of manufacture for use in a computer system, comprising: a memory;"*

Wieser describes a system "that is readily implemented by presently available communication apparatus and electronic components. The invention finds ready application in virtually all commercial communications and/or computer networks including but not limited to world wide web (Internet), intranet, local area network (LAN), wide area network (WAN), wireless network and wired cable transmission systems." *(Appendix C at page 4, line 26 to page 5, line 1)*

Accordingly, Wieser describes that the match server 14 is readily implemented by electronic components known in the year 1999. Such components would include computers having a processor and a memory for storing the programming instructions executed by the processor to perform various functions. Wieser, therefore, contemplates match server 14 being a computer having a processor and a memory for storing programming instructions to be executed by the processor to perform the various functions described by Wieser as being performed by the match server 14.

(ii) *"instructions stored in the memory for operating a method for automatically generating a query from selected document content, comprising:"*

See discussion of the corresponding element of claim 1 (above), where the instructions are stored in a memory of the match server 14.

(iii) *"defining an organized classification of document content with each class in the organized classification of document content having associated therewith a classification label; each classification label corresponding to a category of information in an information retrieval system;"*

See discussion of the corresponding element of claim 1, where the instructions are stored in a memory of the match server 14.