

**THE UNITED STATES DISTRICT COURT  
FOR THE DISTRICT OF DELAWARE**

CHINOOK LICENSING DE, LLC, )  
a Delaware Limited Liability Company, )

Plaintiff )

vs. )

RozMed LLC, Iron Dome LLC, John J. Yim & )  
Associates LLC, Steven S. Yu, and John J. )  
Yim, )

Defendants )

C.A. No.

JURY TRIAL DEMANDED

**COMPLAINT**

This is an action to curtail and remedy the improper, fraudulent and unlawful conduct by RozMed LLC, Iron Dome LLC, John J. Yim & Associates LLC, Steven S. Yu, and John J. Yim (collectively, “Defendants”). Plaintiff Chinook Licensing DE, LLC (“Chinook”), by its attorneys, brings this action for relief against Defendants for tortious interference with Chinook’s business relations.

**PARTIES**

1. Plaintiff Chinook Licensing DE, LLC is a Delaware limited liability company with a place of business at 320 Wilmette Avenue, Glenview, Illinois 60025.

2. On information and belief, Defendant RozMed LLC is a limited liability company organized and existing under the laws of the State of Virginia with its principal place of business at 9810 Cresence Way, Fairfax, Virginia 22032. RozMed’s members include at least Steven S. Yu, a citizen of the state of Maryland. RozMed LLC can be served via its registered agent, Hungju Yu, at 9810 Cresence Way, Fairfax, Virginia 22032.

E-Watch, Inc  
Exhibit 2012  
Petitioner - Iron Dome LLC  
Patent Owner - E-Watch Inc  
IPR2014-00439

3. On information and belief, Defendant Iron Dome LLC is a limited liability company organized and existing under the laws of the State of Virginia, a citizen of the state of Virginia with its principal place of business at 501 Watkins Pond Blvd, Rockville, Maryland 20850. Iron Dome LLC's members include at least Steven S. Yu, a citizen of the state of Maryland. Iron Dome LLC is a wholly –owned subsidiary of RozMed LLC. Iron Dome can be served via its registered agent, RozMed LLC, at 9810 Cresence Way, Fairfax, Virginia 22032.

4. On information and belief, Defendant John J. Yim & Associates LLC is a limited liability company organized and existing under the laws of the State of Virginia with its principal place of business at 7600 Leesburg Pike, East Building, Suite 470, Falls Church, Virginia 22043. John J. Yim & Associates LLC's members include at least John J. Yim, a citizen of the state of Virginia. John J. Yim & Associates can be served via its registered agent, John J. Yim, at 7600 Leesburg Pike, East Building, Suite 470, Falls Church, Virginia 22043.

5. On information and belief, Defendant Steven S. Yu is an individual residing in Rockville, Maryland and a citizen of the State of Maryland. Steven Yu is a managing member of RozMed LLC and principal of Iron Dome LLC.

6. On information and belief, Defendant John J. Yim is an individual residing in Falls Church, Virginia. John Yim is the managing partner of John J. Yim & Associates LLC.

#### **JURISDICTION AND VENUE**

7. This Court has diversity jurisdiction pursuant to 28 U.S.C. § 1332. Complete diversity exists between Plaintiff – a citizen of Delaware - and Defendants – citizens of either Virginia or Maryland - and the amount in controversy is in excess of \$75,000.

8. This Court has personal jurisdiction over all Defendants pursuant to the Delaware Long-Arm Statute, 10 DEL. CODE. ANN. Tit. 3, § 3104, by virtue of Defendants' actions bringing about this cause of action, as alleged herein, and causing injury to Chinook.

9. Venue is proper in this district under 28 U.S.C. § 1391 because among other reasons, Defendants are subject to personal jurisdiction in this District and because of Defendants' actions within this District giving rise to this cause of action.

#### **FACTUAL BACKGROUND**

10. Chinook is in the business of licensing patents that it owns and defending its patent rights against wrongful infringers of those rights.

11. Chinook owns United States Patent No. 7,047,482 (the "'482 patent").

12. Chinook has asserted the '482 patent against several companies for patent infringement in the United States District Court for the District of Delaware. These matters are pending before Judge Stark.

13. Several of the companies against whom Chinook has asserted the '482 patent have settled their respective patent disputes by obtaining a license from Chinook to continue their use of the '482 patent.

14. Litigation is ongoing with several remaining defendants in pending patent infringement actions; however, Chinook continues to work towards resolving its disputes with those remaining companies.

15. Stephen B. Brauerman, Esquire is Chinook's Delaware counsel handling the litigation of the '482 patent in Delaware. His office address is 222 Delaware Avenue, Suite 900, Wilmington, DE 19801.

16. On March 26, 2014 Defendant John J. Yim sent a letter to Mr. Braerman in Delaware threatening to file, on behalf of his client, a petition for *Inter Partes* Review seeking to invalidate Chinook's '482 patent unless Chinook immediately granted three (3) retroactive and transferable licenses to the '482 patent. A true and accurate copy of the March 26, 2014 letter from Mr. Yim to Mr. Braerman is attached hereto as Exhibit A.

17. The March 26, 2014 letter was typed on the letterhead of Defendant John J. Yim & Associates and signed by Defendant John J. Yim. (*Id.*)

18. The March 26, 2014 letter states: “[w]e are attorneys for Iron Dome LLC (www.irondome.com).” (*Id.*)

19. Enclosed with the March 26, 2014 letter was a draft patent license agreement regarding the '482 patent and a draft petition for *Inter Partes* Review (“IPR”) against the '482 patent. True and accurate copies of the draft patent license agreement and draft IPR petition are attached hereto as Exhibit B and Exhibit C, respectively.

20. An IPR is a trial proceeding conducted before the Patent Trial and Appeal Board whereby a third party may seek a review the patentability of one or more claims in a patent. In effect, an IPR is a vehicle by which a third party can seek to invalidate an issued patent.

21. The draft license agreement enclosed with the March 26, 2014 letter is the mechanism by which Defendant Iron Dome LLC proposed to obtain three (3) transferable licenses to the '482 patent. (*See* Exhibit B.)

22. The enclosed draft patent license agreement indicates that Defendant Steven Yu is the Managing Member of Iron Dome LLC. (*Id.* at 5.)

23. The draft license agreement states that “the Parties wish to resolve their patent dispute and avoid the attendant risks, fees, costs, and expenses that are associated with litigation and other patent related proceedings.” (*Id.* at 1.)

24. Prior to Chinook’s local counsel’s receipt of this letter and enclosures, neither Chinook nor any of its representatives or affiliates had contact with any of the Defendants.

25. Defendants knew of and specifically referenced the litigation that Chinook has pending against other parties.

26. In their letter, Defendants state: “we request a rapid resolution of this dispute.” The referenced “dispute” is Defendants’ threatened draft petition for IPR enclosed with the March 26, 2014.

27. The draft license agreement does not propose to purchase licensing rights or include any provision wherein Defendant Iron Dome would remit monetary compensation to Chinook in exchange for the requested licenses.

28. Paragraph 3(a) of the draft license states that Defendant Iron Dome LLC may transfer the transferable licenses to the parties in Chinook’s pending patent infringement actions. (*Id.* at 2.)

29. Paragraph 5 of the draft license states:

**5. Admission of Patent Validity:** Iron Dome admits that the Asserted Patent is valid and enforceable, and as such, will not challenge or participate in any challenge to the validity and enforceability of the Asserted Patent in any kind of legal proceeding.

(*Id.* at 3.)

30. The draft petition for IPR included with the March 26, 2014 letter takes the position that the claims of the ’482 patent are obvious and therefore invalid and not enforceable.

(*See* Exhibit C at v.)

31. The assertion in Paragraph 5 of the draft license agreement and the basis for Defendants' petition for IPR are diametrically opposed.

32. Chinook refused Defendants' coercive offer to license the '482 patent.

33. On April 22, 2014, a petition seeking *Inter Partes* Review of the '482 patent was filed with the United States Patent and Trademark Office by Defendant John J. Yim on Defendant Iron Dome LLC's behalf. Defendant Steven S. Yu is designated as back-up counsel for the petition.

34. On information and belief, in light of the foregoing facts, Defendants' filing of the petition for *Inter Partes* Review of the '482 patent was not in good faith and not for a proper purpose.

35. Specifically, for example, Defendant Iron Dome's press release regarding its petition for *Inter Partes* Review, included the following quote from Defendant Steven Yu: "No one should have to surrender to these lawsuits exploiting defective patents." A true and correct copy of the press release is attached hereto as Exhibit D.

36. The press release, however, failed to mention that Defendants previously sought to enter into a license agreement for the same '482 patent. Despite Iron Dome's declaration that the "defective" patents were being exploited, Defendants are apparently willing explicitly to assert that the patent was valid and enforceable if only they could obtain rights to license unidentified third parties to the patent in question, presumably for their own financial benefit.

37. On information and belief, it is likely that the remaining defendants in the ongoing Delaware patent infringement actions will find out about the pending petition for IPR, which will greatly reduce the chances that Chinook will be able to amicably resolve the infringement actions without further litigation.

38. The malicious actions of Defendants in attempting to coerce a license from Chinook and then filing an unfounded legal IPR proceeding will cost Chinook both time and money and interfere with Chinook's ability to license its patent.

39. Defendants sent their injurious communication from their office in Virginia, using letterhead from their counsel with a Virginia address, where the majority of Defendants are located.

40. Defendants knew that they were communicating with a Delaware LLC (Chinook) and contacted Chinook's Delaware counsel in Wilmington, Delaware.

41. Defendants' actions will cause Chinook to incur damages exceeding \$75,000.

**TORTIOUS INTERFERENCE WITH BUSINESS RELATIONS**  
**(VIRGINIA COMMON LAW)**

42. Chinook repeats, realleges and incorporates by reference the allegations set forth in paragraphs 1–41 as if fully set forth herein.

43. Chinook has existing business relationships with licensees of the '482 patent.

44. Chinook has a reasonable expectation of potential business relationships with alleged patent infringers who have not yet resolved their disputes with Chinook and have not yet obtained a license to the '482 patent.

45. On information and belief, Defendant had knowledge of Chinook's existing business relationships as well as other potential relationships with others regarding Chinook's licensing of the '482 patent. (*E.g.*, Exhibit A.)

46. Defendants knowingly, intentionally, wrongfully, and maliciously interfered with and continue to interfere with Chinook's existing and potential business relationships.

47. The IPR proceeding is unfounded litigation. The only purpose for the litigation is to coerce Chinook into capitulating and tendering a license to Defendant Iron Dome, LLC.

48. Defendants' actions have harmed Chinook's business relationships with current licensees and potential licensees. Chinook's loss of these advantageous business relations resulted directly from Defendants' improper and unlawful actions.

49. Chinook will suffer substantial damages, in an amount to be determined at trial, as a result of Defendants' improper and unlawful actions.

**DEMAND FOR JURY TRIAL**

50. Chinook demands a trial by jury of any and all causes of action.



**PRAYER FOR RELIEF**

WHEREFORE, Chinook respectfully prays for judgment:

- A. Ordering Defendants to withdraw its petition for *Inter Partes* Review of the '482 patent;
- B. Enjoining Defendants from taking any further action or threatening to take further action against the validity of the '482 patent;
- C. Enjoining Defendants from soliciting, entering into, or enforcing any agreements with potential licensees with regard to the '482 patent or otherwise impair Chinook's ability to license the '482 patent;
- D. Awarding Chinook monetary damages; and
- E. Granting such other and further relief as this Court deems just and proper.

Dated: May 12, 2014

Bayard, P.A.

/s/ Stephen B. Braerman  
Richard D. Kirk (rk0922)  
Stephen B. Braerman (sb4952)  
Vanessa R. Tiradentes (vt5398)  
Sara E. Bussiere (sb5725)  
222 Delaware Avenue, Suite 900  
Wilmington, DE 19801  
(302) 655-5000  
rkirk@bayardlaw.com  
sbraerman@bayardlaw.com  
vtiradentes@bayardlaw.com  
sbussiere@bayardlaw.com

Attorneys for Plaintiff Chinook Licensing DE,  
LLC

# **EXHIBIT A**

## JOHN J. YIM & ASSOCIATES, LLC

Tysons Corner  
7600 Leesburg Pike  
East Building, Suite 470  
Falls Church, VA 22043

Tel. 703.749.0500  
Fax. 202.379.1723

John J. Yim\*  
Managing Partner  
E-MAIL: [jyim@yimassociates.com](mailto:jyim@yimassociates.com)

\*Admitted in CA, DC, MO, VA &  
United States Patent and Trademark Office

March 26, 2014

Via Federal Express  
Stephen B. Brauerman  
Bayard, P.A.  
222 Delaware Avenue, Suite 900  
Wilmington, DE 19801

Re: *Inter Partes* Review of U.S. Patent No. 7,047,482  
Chinook Licensing DE, LLC  
**SETTLEMENT PURPOSES ONLY**

Dear Counsel:

We are attorneys for Iron Dome LLC ([www.irondome.com](http://www.irondome.com)).

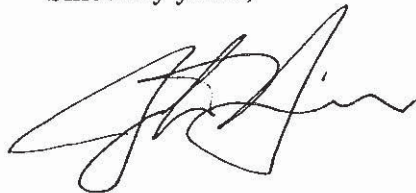
This letter addresses the invalidity of the patent asserted by Chinook Licensing DE, LLC, against Match.com, Inc., Scribd, Inc., StumbleUpon, Inc., Facebook, Inc., Hulu, LLC, LinkedIn Corporation, Project Rover, Inc., Zoosk, Inc., and Pandora Media, Inc., in civil actions recently filed in the U.S. District Court for the District of Delaware. Attached is a fully prepared, but not yet filed, petition for *Inter Partes* Review (IPR) against the asserted patent.

Although the validity of the asserted patent is questioned, we wish to acquire retroactive and fully transferable licenses to the asserted patent. After reviewing and considering the merits of the enclosed draft IPR petition, please contact me so that we can resolve this matter.

With the understanding that you are fully acquainted with the new IPR proceedings, we request a rapid resolution of this dispute.

Please contact us no later than two weeks of receipt of this letter. We enclose a license agreement for your review.

Sincerely yours,

A handwritten signature in black ink, appearing to read "John J. Yim". The signature is fluid and cursive, with the first name "John" being the most prominent part.

John J. Yim

Enclosures

# **EXHIBIT B**

## Patent License Agreement

This patent license agreement ('Agreement') is dated \_\_\_\_\_ ('Effective Date') and is between:

IRON DOME LLC, a Virginia limited liability company,  
*and*  
CHINOOK LICENSING DE LLC, a Delaware limited liability company.

Chinook Licensing DE, LLC is the owner of U.S. Patent No. 7,047,482 issued on May 16, 2006 ('Asserted Patent'). The patent owner has brought civil actions for patent infringement against various parties for operating websites that make personalized recommendations to users. Iron Dome LLC wishes to acquire transferable licenses to the Asserted Patent.

The Parties wish to resolve their patent dispute and avoid the attendant risks, fees, costs, and expenses that are associated with litigation and other patent-related proceedings. Therefore, the Parties agree as follows:

### 1. Definitions

(a) 'Chinook' means Chinook Licensing DE, LLC, the owner of the Asserted Patent, and all of its Affiliates.

(b) 'Iron Dome' means Iron Dome LLC and all of its Affiliates.

(c) 'Party' means Chinook or Iron Dome. 'Parties' means both Chinook and Iron Dome.

(d) 'Third Party' means any party that is neither Chinook nor Iron Dome.

(e) 'Affiliate' means any present or future entity, corporation, company, association, partnership, joint venture, organization or other entity that directly or indirectly controls, is controlled by, or is under common control with a given entity. For purposes of this definition, 'control' means (i) in the case of a corporation, the direct or indirect ownership of 50% or more of the shares of stock entitled to vote for the election of directors (or of persons performing similar functions); or (ii) in the case of any other type of legal entity, the direct or indirect ownership of 50% or more of the equity interests, or status as a general partner in any partnership or joint venture, or any other arrangement whereby a party controls or has the right to control the Board of Directors or equivalent governing body of a corporation or other entity.

(f) 'Infringement Actions' means those legal actions that Chinook has brought asserting infringement of the Asserted Patent against various defendants in any forum, including any actions brought in the U.S. International Trade Commission, and including the

following in the U.S. District Court for the District of Delaware on or about December 20, 2013 and January 20, 2014:

- |     |               |  |
|-----|---------------|--|
| (1) | 1:13-cv-02077 | Chinook Licensing DE, LLC v. Match.com, Inc. |
| (2) | 1:13-cv-02078 | ... v. Scribd, Inc.                          |
| (3) | 1:13-cv-02079 | ... v. StumbleUpon, Inc.                     |
| (4) | 1:14-cv-00073 | ... v. Facebook, Inc.                        |
| (5) | 1:14-cv-00074 | ... v. Hulu, LLC                             |
| (6) | 1:14-cv-00075 | ... v. LinkedIn Corporation                  |
| (7) | 1:14-cv-00076 | ... v. Project Rover, Inc.                   |
| (8) | 1:14-cv-00077 | ... v. Zoosk, Inc.                           |
| (9) | 1:14-cv-00105 | ... v. Pandora Media, Inc.                   |

**2. Grant of Patent Licenses:** Chinook grants to Iron Dome three (3) separate retroactive, royalty-free, non-exclusive licenses for the Asserted Patent (each a ‘Transferable License’), as well as for any and all United States patents now and in the future owned, controlled, assigned, or licensed to Chinook that are necessary for operating a website that is covered by the claims of the Asserted Patent.

### 3. Transferability of Patent Licenses

(a) **Transferability:** Iron Dome is permitted to separately transfer each of the Transferable Licenses to separate Third Parties and its Affiliates. For avoidance of doubt, this is intended to mean that the first Transferable License is transferable to one Third Party and its Affiliates, the second Transferable License is transferable to another Third Party and its Affiliates, and so on. Chinook understands that Iron Dome may transfer these Transferable Licenses to defendants in the Infringement Actions.

(b) **Notification:** Iron Dome shall notify Chinook of any transfer of a Transferable License in writing (including the identity of the Third Party transferee) within five business days after such transfer.

(c) **Release:** Upon the transfer of a Transferable License to a Third Party who is a defendant in any of the Infringement Actions, Chinook shall release such Third Party from:

(i) all past and present claims, allegations, damages, obligations, liabilities or expenses of any kind or nature relating to the subject matter of the relevant Infringement Action;

(ii) all claims that were or could have been asserted in the relevant Infringement Action, and

(iii) all claims based on or arising out of the alleged infringement of the Asserted Patent.

**(d) Dismissal:** After Chinook receives written confirmation by a Third Party that they are the transferee of a Transferable License, Chinook shall dismiss its Infringement Action against such Third Party within six (6) business days.

**(e) Covenant Not-To-Sue:** Chinook will not assert any claim, or commence or join in any legal, administrative or other proceeding under the Asserted Patent against Iron Dome or any Third Party transferees, or any of its customers, suppliers, importers, manufacturers, or distributors.

**(f) Non-Assertion Runs with Patent:** Chinook will impose this covenant not-to-sue on any Third Party to whom Chinook may assign the Asserted Patent.

**(g) Single Transfer Only:** Each of the Transferable Licenses is transferable only once. Once Iron Dome transfers a Transferable License to a Third Party, that Third Party may not subsequently transfer the Transferable License to another Third Party. Iron Dome will inform of and impose this single-transfer limitation upon any Third Party transferee.

#### 4. Enforcement

**(a)** Any Third Party who is a transferee of the Transferable License shall have standing and the right to enforce this Agreement (including the provisions for Release and Dismissal set forth in Sections 3(c) and (d) of this Agreement) against Chinook, without requiring the joining of Iron Dome.

**(b)** Chinook will not delay its duties of Release and Dismissal set forth above in Sections 3(c) and (d) of this Agreement. Chinook will be responsible for all expenses (including attorney fees) incurred by Iron Dome and/or Third Parties relating to the enforcement of this Agreement due to any such delay.

**5. Admission of Patent Validity:** Iron Dome admits that the Asserted Patent is valid and enforceable, and as such, will not challenge or participate in any challenge to the validity and enforceability of the Asserted Patent in any kind of legal proceeding.

**6. Confidentiality:** The Parties shall treat this Agreement as confidential and shall not disclose the existence, contents, terms, or conditions of this Agreement to any Third Party without the prior written consent of the other Party, except as necessary in the following conditions:

- (a) as required by any court or other governmental body;
- (b) as otherwise required by law;
- (c) as otherwise may be required by applicable securities and other law and regulation, including the regulations of the U.S. Securities and Exchange Commission;



Patent License Agreement

(d) to legal counsel, accountants, and other financial advisors of the Parties, subject to obligations of confidentiality;

(e) to the extent necessary for the enforcement of this Agreement or rights under this Agreement;

(f) to banks, investors, and other financing sources, subject to a non-disclosure agreement respecting confidentiality customary to the corresponding prospective transaction;

(g) in connection with an actual or prospective merger, acquisition, or other transaction with a Third Party, subject to a non-disclosure agreement respecting confidentiality customary to such prospective transaction;

(h) to prospective transferees of the Transferable Licenses, including those defendants in the Infringement Actions.

**7. Ownership:** Chinook represents that it is the sole owner of the Asserted Patent and has the right to grant the licenses and covenants in this Agreement related thereto.

**8. Representations:** Each Party represents to the other Party, as of the Effective Date, as follows:

(a) that it has all requisite corporate power and authority to enter into this Agreement and to perform its obligations hereunder and to grant the licenses, releases, promises, covenants, and other rights contained herein;

(b) that all acts required to be taken by it to authorize the execution and delivery and performance of this Agreement, and the consummation of the transactions contemplated herein have been duly and properly taken, and no other corporate proceedings on its part are necessary to authorize such execution, delivery, and performance;

(c) that this Agreement has been duly executed and delivered by it and constitutes a legal, valid, and binding obligation of it, enforceable against it in accordance with its terms.

**9. Entire Agreement:** This Agreement sets forth all the rights and obligations between the Parties.

**10. Severability:** If any provision of this Agreement or the application of any such provision to any person or circumstance is declared judicially or by arbitration to be invalid, unenforceable, or void, such decision will not invalidate or void the remainder of this Agreement. And this Agreement is to be deemed amended by modifying such provision to the extent necessary to render it valid, legal, and enforceable while preserving as much as possible its intent or, if such modification is not possible, by replacing it with another provision that is legal and enforceable and that achieves similar objectives.

Patent License Agreement

**11. Choice of Law & Venue:** The laws of the state of New York, without reference to its conflict of laws principles, will govern this Agreement. The exclusive venue for any action brought by Chinook against Iron Dome regarding the construction, validity, enforceability, performance, or otherwise regarding a challenge to this Agreement will be the state courts of the Commonwealth of Virginia sitting in Fairfax County. Otherwise, the exclusive venue for any actions among the Parties and Third Party transferees of the Transferable Licenses under this Agreement will be the U.S. District Court for the District of Delaware and wherever other venue to which any of the Infringement Actions may be transferred.

**12. Notice:** All notices relating to this Agreement shall be given in writing and will be delivered through one or more of the following means: (1) in-person, (2) by certified mail with prepaid postage and return receipt, or (3) by a commercial overnight courier that guarantees next day delivery and provides a receipt. Notices are to be addressed as follows:

If to Iron Dome: Steven Yu, M.D.  
Iron Dome LLC  
PO Box 10034  
Gaithersburg, MD 20898

If to Chinook: \_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

The Parties sign this Agreement on the Effective Date given above:

**Chinook Licensing DE LLC**

By: \_\_\_\_\_

Name: \_\_\_\_\_

Title: \_\_\_\_\_

**Iron Dome LLC**

By: \_\_\_\_\_

Name: Steven Yu

Title: Managing Member

# **EXHIBIT C**

United States Patent & Trademark Office  
Patent Trial & Appeal Board

---

IRON DOME LLC  
Petitioner

v.

CHINOOK LICENSING DE LLC  
Patent Owner

---

Petition for Inter Partes Review  
of

**Patent No. 7,047,482 (to Gary Odom)**

Titled: *Automatic directory supplementation*

Issue date: May 16, 2006

---

For Paralegal:

Number of Claims Challenged = 19  
Power of Attorney enclosed  
Fee paid online by credit card

Contact: John Yim  
Phone: 703.749.0500  
Email: [jyim@yimassociates.com](mailto:jyim@yimassociates.com)

Table of Contents

**I. Introductory Matters ..... - 1 -**

    A. Relief Requested..... - 1 -

    B. Grounds for Standing..... - 1 -

    C. Mandatory Notices..... - 1 -

**II. Prior Art References.....- 3 -**

    A. The claims have an effective filing date of February 28, 2001 ..... - 3 -

    B. List of Prior Art..... - 3 -

**III. Technical Background & Claim Construction..... - 4 -**

    A. Technical Background of the Challenged Patent..... - 4 -

    B. Claim Construction..... - 5 -

**IV. Grounds for Challenge.....- 9 -**

    A. Chen - primary prior art reference ..... - 9 -

    B. Lieberman - second prior art reference..... - 9 -

**V. Claim Analysis ..... - 10 -**

    Independent Claim 1..... - 10 -

    Claim 2..... - 15 -

    Claim 3..... - 15 -

    Claim 4..... - 16 -

    Claim 5..... - 16 -

    Claim 6..... - 17 -

    Claim 7..... - 17 -

    Claim 9..... - 18 -

    Claim 10..... - 18 -

    Independent Claim 11 ..... - 19 -

    Claim 12..... - 23 -

    Claim 13..... - 23 -

    Claim 14..... - 24 -

US 7,047,482

Claim 15.....	- 24 -
Independent Claim 16 .....	- 24 -
Claim 17.....	- 28 -
Claim 18.....	- 29 -
Claim 19.....	- 29 -
Claim 20.....	- 29 -

Exhibit List

- Exh. 1001 U.S. Patent No. 7,047,482 (‘challenged patent’)
- Exh. 1002 Liren Chen & Katia Sycara, “WebMate: A Personal Agent for Browsing and Searching” in *Proceedings of the Second International Conference on Autonomous Agents*. Sponsored by ACM SIGART in Minneapolis/St. Paul, MN; May 9-13, 1998. Selected pages: Table of Contents, pp. 132-139 (‘Chen’)
- Exh. 1003 Henry Lieberman, “Letizia: An Agent That Assists Web Browsing” in *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*. Sponsored by IJCAI in Montréal, Québec, Canada; August 20-25, 1995. Selected pages: Table of Contents, pp. 924-929 (‘Lieberman’)
- Exh. 1004 “BRIEF FOR APPELLANT” dated Feb. 18, 2005 in the prosecution history of the challenged patent (‘Appeal Brief’)

Citation Form Used

Reference to supporting documents indicated by “@\_\_\_\_\_”.

Citations to U.S. Patents are shown as [column number : line numbers].

Citations to line-numbered documents are shown as [page number : line numbers].

Claim terms are distinguished from other text by “underlining.”

**Issue Presented**

The challenged patent was recently asserted in patent infringement lawsuits against Facebook, Scribd, Hulu, Pandora, Match.com, and others for operating websites that observe a user's selections and make personalized recommendations for other similar selections (e.g. "other music you may like"). A search of the prior art reveals that the claims of the challenged patent are obvious in view of prior web browsing agents that explore the web and recommend webpages of interest to the user.



I. Introductory Matters

IRON DOME LLC (Petitioner) petitions for *Inter Partes* Review (IPR) of U.S. Patent No. 7,047,482 (challenged patent @Exh. 1001), which is owned by CHINOOK LICENSING DE LLC.

**A. Relief Requested**

Petitioner requests cancellation of claims 1-7 and 9-20 (total of 19 claims) of the challenged patent for obviousness under 35 U.S.C. § 103.

**B. Grounds for Standing**

Petitioner certifies that the challenged patent is available for IPR and that Petitioner is not barred or estopped from requesting an IPR challenging the patent claims on the grounds identified in this petition.

**C. Mandatory Notices**

Real Parties-in-Interest: (1) IRON DOME LLC, a Virginia limited liability company, which is a wholly-owned subsidiary of ROZMED LLC, a Virginia limited liability company; and (2) Steven S. Yu, M.D., an individual residing in Rockville, Maryland and the managing member of ROZMED LLC.

Individual Steven S. Yu, M.D. declares that there are no other parties that are funding this IPR, nor participating in any manner in this IPR; and further that this statement is being made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Related Matters: The challenged patent has been asserted by the patent owner in litigation against several defendants alleging infringement by websites that observe a subscriber's selections and recommend other similar selections to the subscriber. On or about December 20, 2013 and January 20, 2014, the patent owner Chinook Licensing DE LLC filed civil actions 1:13-cv-02077 through 02079, 1:14-cv-00073 through 00077, and 1:14-cv-00105 in the U.S. District Court for Delaware.

Individual Steven S. Yu, M.D. declares that Petitioner is not a party to any of these civil actions, nor has Petitioner been given or taken any direct financial interest relating to the outcome of these civil actions; and further that this statement is being made with the knowledge that willful false statements and the like so made are punishable by fine or imprisonment, or both, under Section 1001 of Title 18 of the United States Code.

Electronic Service: Petitioner consents to service by email at:

[jjim@yimassociates.com](mailto:jjim@yimassociates.com).

Lead Counsel	Back-Up Counsel
John J. Yim (Reg. No. 47,197) JOHN J. YIM & ASSOCIATES LLC 7600 Leesburg Pike East Building, Suite 470 Tysons Corner, VA 22043 Tel: 703.749.0500 Fax: 202.379.1723 Email: <a href="mailto:jjim@yimassociates.com">jjim@yimassociates.com</a>	Steven S. Yu (Reg. No. 58,776) IRON DOME LLC PO Box 10034 Gaithersburg, MD 20898 Tel: 202.262.0426 Email: <a href="mailto:syu@irondome.com">syu@irondome.com</a>

## II. Prior Art References

### A. The claims have an effective filing date of February 28, 2001

The challenged patent was granted from application Serial No. 09/796,235 filed on February 28, 2001. This application does not claim priority to any prior-filed applications. Accordingly, the earliest possible effective filing date for the claims of the challenged patent is **February 28, 2001**.<sup>1</sup>

### B. List of Prior Art

The prior art publications referenced herein are as follows.

1. Liren Chen & Katia Sycara, "WebMate: A Personal Agent for Browsing and Searching" in *Proceedings of the Second International Conference on Autonomous Agents*. Sponsored by ACM SIGART in Minneapolis/St. Paul, MN; May 9-13, 1998. Selected pages: Table of Contents, pp. 132-139 ('Chen'; Exh. 1002)<sup>2</sup>
2. Henry Lieberman, "Letizia: A Personal Agent for Web Browsing" in *Proceedings of the Fourteenth International Conference on Artificial Intelligence*. Sponsored by IJCAI in Montreal, Quebec, Canada; August 20-25, 1995. Selected pages: Table of Contents, pp. 1003-1004 ('Lieberman'; Exh. 1003)<sup>3</sup>

<sup>1</sup> We reserve the right to dispute whether the challenged claims should legitimately have the benefit of this filing date (i.e. lack of written support).

<sup>2</sup> Chen is an article published in a book distributed to participants at a symposium proceeding in May 1998. The copy held by the Library of Congress in Washington, D.C. is date stamped October 18, 1999. Therefore, Chen is prior art under 35 U.S.C § 102(b).

<sup>3</sup> Lieberman is an article published in a book distributed to participants at a symposium proceeding in August 1995. The copy held by the Georgetown University Libraries in Washington, D.C. is date stamped March 25, 1998. Therefore, Lieberman is prior art under 35 U.S.C § 102(b).

Neither of the above publications were cited in the original prosecution of the challenged patent.

### III. Technical Background & Claim Construction

#### A. Technical Background of the Challenged Patent

The challenged patent (Exh. 1001) describes a computer software program that assists a user in browsing the Internet by making recommendations for webpages that might interest the user (sometimes referred to as a browsing agent). To make these recommendations, the software examines a set of items that the user has selected (e.g. the user's bookmark folder of "favorite" websites) as indicative of the user's field of interest.

FIG. 2 of the challenged patent shows a directory 3 of links or documents (e.g. webpages) that the user has selected and categorized according to a topic as designated by the directory title 5. @3:14-18. The browsing agent applies linguistic analysis techniques to the textual content of the selected items and extracts words that appear to be relevant to the user's field of interest (e.g. by analysis of word frequency, word placement, syntax, etc.). @3:33-36. Extracted keywords are rated or ranked according to such factors as location in the document, prominence, and frequency of appearance. @4:22-33.

This keyword extraction and ranking process is performed on all the items in directory 3 and those that best represent the content of the directory 3 as a whole are

selected as directory keywords 88. @4:63-5:17. Using the directory keywords 88, the browsing agent searches the network for other items having textual content similar to those in directory 3. @5:18-23. The user can set a breadth threshold to widen or narrow the scope of the search. @4:54-62. As shown in FIG. 6, the original directory 3K (the user's interest in the musical group King Crimson) is supplemented with the search results in a new box beneath (directory supplementation 6K). The new links 1F are displayed along with a ranking 66 of their relevance. @5:64-67.

## B. Claim Construction

In the context of an *inter partes* review, claim terms must be given their broadest reasonable interpretation in view of the specification.

### 1. “autonomously” (see claims 10 and 11)

This term relates to how the software operates without direct user participation. This term was added to the claims by amendment during the prosecution of the challenged patent, but it does not appear anywhere in the specification of the challenged patent. However, the challenged patent explains the “automatic” operation of the software agent as follows:

Directory supplementation 6 may be enabled 10 by default, by software-determined protocol, or by user determination. Automatically supplementing a directory 6 refers to adding links 1 or documents 2 to a directory 3 without a user having to search 12 or manually add links 1 to that directory 3.

@4:35-40. Thus, the term “autonomously” is synonymous with “automatically” and means that the directory supplementation process occurs “without a user having to search [] or manually add links [] to that directory [].”

2. “without contemporaneous user input” (and variations thereof)

This is another term relating to how the software operates without direct user participation. This term was added to the claims by amendment during the prosecution of the challenged patent, but it does not appear anywhere in the specification of the challenged patent. We therefore turn to the applicant’s statements during the prosecution of the challenged patent to understand the meaning of this term. The applicant gave the following explanation in his Appeal Brief:

So, Examiner ... tacitly concurred with appellant, that, in context, the two limitations applicable to the meaning of ‘without user input’ comprise:

1. no user input of search parameters;
2. no user input of search locations.

That is exactly what appellant had explained in his 08/27/2004 reply to the first office action rejection.

@Appeal Brief 14 top (Exh. 1004). Thus, according to the applicant’s statements in the prosecution history, the term “without contemporaneous user input” means that the software works with “no user input of search parameters” and “no user input of search locations.”

3. “searching as a background operation” (and variations thereof)

This term relates to how the software runs as a background process relative to other processes that the computer may be running. This term was added to the claims by amendment during the prosecution of the challenged patent, but it does not appear anywhere in the specification of the challenged patent. We therefore turn to the applicant’s statements during the prosecution of the challenged patent to understand the meaning of this term. The applicant gave the following explanation in his Appeal Brief:

[The software process in my patent application] is fairly characterized as lazy because time is not of the essence. A user doesn’t initiate search: the process works in the background, without arousing expectation of quick results.

@Appeal Brief 2 ¶2<sup>nd</sup> (Exh. 1004). Thus, according to the applicant’s statements in the prosecution history, the term “searching as a background operation” means that the searching operation occurs in the background, simultaneously but with a lower priority to other operations that the computer may be performing.

4. “precondition” (claims 10 and 11)

This term relates to the selections made by the user before the software begins its autonomous operation. This term was added to the claims by amendment during the prosecution of the challenged patent, but it does not appear anywhere in the specification of the challenged patent. We therefore turn to the applicant’s statements

during the prosecution of the challenged patent to understand the meaning of this term. The applicant gave the following summary of his invention in his Appeal Brief:

[My patent application] describes an autonomous search mechanism, solving the problem of finding similar documents to ones already known without any user effort whatsoever. The only *precondition* to initiating the claimed process is user placement of one or more documents in a file system directory as reference material for guiding the search.

...

As an exemplary use-case scenario, a user browses the web, saving topically-related document links in the same web-favorites folder. Once this *precondition* is met, the claimed invention software kicks in: deriving keywords from the saved documents, thus discerning the topic of interest, then searching for other related documents, resulting in supplementing the directory with newly-found documents - hence the title of [my patent application]: ‘automatic directory supplementation’.

@Appeal Brief 2 top (italics added; Exh. 1004). Further clarification of the term “precondition” is given by the manner in which the applicant distinguished his invention over the prior art:

By contrast, the claimed invention relies solely upon documents in a directory, without relying upon user input. Yes, [in my invention] a user must first put the documents in the directory, but that is a *precondition*; user input is not required for the claimed process to work, unlike [the prior art].

@Appeal Brief 14 ¶6<sup>th</sup> (italics added).

Thus, according to the applicant’s statements in the prosecution history, the



term “precondition” means the selection of items that the user has identified as being of interest and encompasses at least “user placement of one or more documents in a file system directory as reference material for guiding the search” such as “saving topically-related document links in the same web-favorites folder.” In the words of the applicant, when the user has performed this step, “this precondition is met.”

#### IV. Grounds for Challenge

Petitioner requests cancellation of claims 1-7 and 9-20 of the challenged patent. Claims 1-7 and 9-19 are obvious over Chen alone. Claim 20 is obvious over Chen in view of Lieberman. Claim 8 is not of commercial interest to Petitioner and is therefore omitted from this IPR.

##### A. Chen - primary prior art reference

Chen (Exh. 1002) is an article published in 1998 and describes the WebMate software, which is a web browsing agent that automatically learns the user’s interest and finds webpages that match the user’s interest. @132 L Abstract. WebMate compiles a “personal newspaper” of the recommended webpages. @Id.

##### B. Lieberman - second prior art reference

Lieberman (Exh. 1003) is an article published in 1995 and describes the Letizia software, which is a “user interface agent that assists a user browsing the World Wide Web.” @924 L Abstract. The software agent infers the user’s interest from browsing behavior and explores other items that may interest the user. @Id. Chen specifically

cites to Lieberman as “Related work.” @Chen 138 R, #17 (Exh. 1002). Thus, it is appropriate to combine Chen with Lieberman for an obviousness analysis.

## V. Claim Analysis

The challenged patent has three independent claims, which are claims 1, 11, and 16, all being for a “computer-implemented method,” i.e. computer software. In general, all three independent claims are directed to a software agent that automatically (“without contemporaneous user input”) finds items (e.g. links or webpages) that match the user’s interest and adds them to a directory. A well-known example would be a web browsing agent that analyzes a user’s bookmark folder of “favorite” links, searches the Internet for links with similar content, and recommends other links that might be of interest to the user.

### Independent Claim 1

- a) (preamble) “A computer-implemented method for augmenting a directory without contemporaneous user input comprising:”

Chen (Exh. 1002) describes the WebMate software agent, which automatically learns a user’s interest by extracting keywords from documents selected by the user and “automatically spiding news sources” (“spiding” means crawling the web, as a spider would) on the Internet for webpages having similar content. @Abstract. As explained above, the term “without contemporaneous user input” means that the software works with “no user input of search parameters” and “no user input of

search locations.” @IPR 6. More detailed explanation about this and other features of WebMate are given below.

- b) (claim 1) “accessing at least a first document via a first directory without contemporaneous user selection of said first document, said first document comprising at least in part topical textual content”

This claim term is essentially stating that the software agent automatically examines the text in the documents selected by the user (e.g. bookmarked links in a “favorites” folder). In the challenged patent, an example of a “first directory” is given in FIG. 6. @Exh. 1001 p. 6. Here, the directory 3K is titled “King Crimson” and contains links related to this particular musical group. The directory 3K (upper portion) corresponds to the “first directory.” As explained above, the software agent analyzes the textual content of the linked pages in this directory for keywords that appear to represent the user’s interest. Regarding the claim term “without contemporaneous user selection,” the links in directory 3K were previously selected by the user, and the software agent automatically accesses the linked pages in directory 3K.

WebMate works in a similar manner by “learning user interests incrementally and with continuous update and automatically providing documents (e.g. a personalized newspaper) that match the user interests.” @132 R ¶1bott. WebMate develops the user profile automatically and “unobtrusively” by keyword extraction from documents that the user selects by marking them as “I like it.” @132 L Abstract, 134 L mid. The “user can provide multiple pages as similarity/relevance guidance for

the search.” @132 L Abstract. Thus, WebMate accesses the user-selected documents “without contemporaneous user selection,” which is understood to mean “no user input of search parameters” and “no user input of search locations.” @IPR 6.

“In order to save on storage space,” WebMate “doesn’t keep any of the previous positive example documents.” @133 R ¶top. However, storage space is not a concern for higher capacity machines and it would be obvious to keep the set of “I like it” documents selected by the user as a collection of links instead of discarding them. This collection of “I like it” documents selected by the user would constitute a “first directory” recited in the claim. To be explained in detail below, these user-selected documents have “topical textual content.”

- c) (claim 1) “deriving at least one keyword indicative of at least one topical content from said first document”

Once the user has selected the items of interest, WebMate automatically parses the selected web documents and extracts words contained therein, including those in the title and headers. @133 R – 134 L. Using the extracted words, WebMate then constructs a word vector for the document. @133 R ¶2<sup>nd</sup>. The word vector is an aggregate representation of the content as a series of word elements, each weighted according to its frequency of occurrence in the document. @Id. The combined word vectors become the user profile. @Id.

d) (claim 1) “searching as a background operation a plurality of documents in storage in at least one computer without contemporaneous user input of a search location, such that said search comprises searching for documents related by said at least one keyword to said first document, thereby accessing a second document”

This claim term is essentially stating that the search for similar documents on the Internet or network occurs as a background process on the computer. With the user’s interest profile created, WebMate can search the entire web for pages relevant to the user’s interest, or monitor only a selection of websites designated by the user. @134 §3.2. If the user has selected a particular list of websites (URLs) for monitoring, WebMate can search these websites to find relevant webpages based on their similarity to the word vector in the user’s profile. @134 L ¶10. If the user does not provide a list of websites to monitor, then WebMate can simply construct a search query “using the top several words in the current [user] profile and sends it to search engines (e.g. Altavista, Yahoo).” @134 R ¶2<sup>nd</sup>.

When WebMate is searching selected websites for pages that match the word vector in the user’s profile, or searching the entire web using the top words in the user’s profile, it is “searching for documents related by said at least one keyword to said first document.” The webpages being examined by WebMate constitute a “second document” which are “in storage in at least one computer,” i.e. in the computers running the websites.

WebMate can perform its search “in the middle of the night when the network traffic is low” such that “[i]n the morning, the user can read the recommended personal newspaper.” @134 R top. WebMate can also perform its search contemporaneously “[i]f the result is needed immediately.” @Id. R ¶ 2<sup>nd</sup>. As “Related work,” Chen also points to the Letizia web browsing agent which “can recommend nearby pages by doing lookahead search.” @138 R top. Thus, Chen discloses three different timeframes in which searching can be performed: (1) overnight during machine idle time, (2) immediately, and (3) as a “lookahead search” while the user is browsing the Internet. All three searching methods are performed “without contemporaneous user input of a search location.” Moreover, from these teachings about the different timeframes for performing the search, it would be obvious to have WebMate perform “searching as a background operation” in the same manner as Letizia’s “lookahead search.”

- e) (claim 1) “determining relevance of said second document to said at least one keyword”

WebMate decides whether to recommend the pages it has found to the user on the basis of whether the page content “similarity is greater than some threshold” as compared against the user’s profile (which is represented by word vectors). @134 L ¶1bott. (see also R ¶2<sup>nd</sup>).

f) (claim 1) “adding a reference to said second document in a results directory.”

If the user has selected websites for monitoring, WebMate can perform the search overnight and in the morning, provide the user with a “personal newspaper” listing the recommended pages. @134 L ¶11. If WebMate performs a web-wide search, WebMate analyzes the search results and presents those meeting the required threshold for similarity “in descending order of relevance.” @134 R ¶2<sup>nd</sup>. In both of the aforementioned situations, WebMate works by “adding a reference to said second document [i.e. the recommended webpage] in a results directory.”

### Claim 2

Claim 2 depends from claim 1 and specifies that “at least part of said storage is on a different computer than the computer storing said first directory.” WebMate performs a search of the World Wide Web for documents that are relevant to the user’s interest. The World Wide Web is a network of many different computers throughout the world. For example, Chen describes an experiment with WebMate in which 14 news sites covering technology news were monitored by WebMate for news articles matching the user’s interest. @134 R ¶3<sup>rd</sup>. The 14 news sites on the Internet selected for monitoring are on a “different computer” than the user’s computer. @Id.

### Claim 3

Claim 3 depends from claim 1 and adds the step of “deriving a plurality of keywords.” As explained above, WebMate parses through the documents selected by

the user (marked as “I like it”) and extracts words from the documents to generate a weighted word vector representing the user’s field of interest.

#### Claim 4

Claim 4 depends from claim 3 and adds the step of “ranking at least two of said plurality of keywords.” As explained above for claim 1(c), WebMate generates a word vector for the words in the document, and each word in the vector is weighted according to its frequency of occurrence. @IPR 12. In addition, WebMate can operate by constructing a search query “using the top several words in the current [user] profile and sends it to popular search engines (e.g. Altavista, Yahoo).” @134 R ¶2<sup>nd</sup>. In Chen’s experimental demonstration of WebMate, the “top 5 words” were used to construct a search query for the Lycos search engine. @138 L ¶3<sup>rd</sup>. Both of the aforementioned functions, weighting of the words in the word vector or selection of the “top several words,” constitutes “ranking at least two of said plurality of keywords.”

#### Claim 5

Claim 5 depends from claim 1 and adds the step of “accessing a plurality of documents in said first directory.” WebMate performs the parsing routine “[f]or each positive example (i.e. an HTML documents [sic] that the user has marked ‘I like It’).” @133 R ¶5<sup>th</sup>. Thus, if there are multiple web documents marked by the user as “I like



it,” then WebMate will perform the parsing routine on each of the multiple web documents.

### Claim 6

Claim 6 depends from claim 1 and adds the step of “signifying the relevance of said second document to documents in the first directory when displaying said results directory.” In the challenged patent, FIG. 6 shows the results directory 6K with the new links 1F displayed along with a ranking 66 of their relevance. @Exhibit 1001 p. 6 (see also 5:64-67).

In WebMate, the “personal newspaper” provided to the user “sorts all the recommended pages in decreasing order of similarity.” @134 L ¶1. Alternatively, if instead WebMate performs a search query, WebMate lists the recommended pages in “descending order of relevance.” @Id. R ¶2<sup>nd</sup>. Moreover, Chen shows the search results returned by the Lycos search engine in one of the experimental demonstrations of WebMate. @138 L ¶3<sup>rd</sup>. Regarding the Lycos search results shown, the “content of links marked with ‘\*’ are similar to the content of the page given as the ‘relevant’ feedback.” @Id. Thus, WebMate performs the step of “signifying the relevance of said second document ... when displaying said results directory.”

### Claim 7

Claim 7 depends from claim 1 and adds the step of “comparing the relevance of said second document to a preset threshold.” WebMate recommends a webpage to

the user if its “similarity is greater than some threshold.” @134 L ¶11 bott. (see also R ¶12<sup>nd</sup>).

### Claim 9

Claim 9 depends from claim 1 and adds the step of “displaying said results directory.” As explained above for claim 1(f), WebMate can provide the user with a “personal newspaper” listing the recommended pages or present the search results “in descending order of relevance.” @IPR 15. This constitutes “displaying said results directory.”

### Claim 10

Claim 10 depends from claim 1 and adds the step of “recognizing a precondition for autonomously augmenting said results directory, prior to accessing said first document.” As explained above, the term “autonomously” is synonymous with “automatically.” @IPR 5. The term “precondition” means the selection of items that the user has identified as being of interest and encompasses at least “user placement of one or more documents in a file system directory as reference material for guiding the search” such as “saving topically-related document links in the same web-favorites folder.” @IPR 7.

As explained above, although WebMate does not keep the documents marked “I like it” to save on storage space, it would be obvious to do so if storage space is not a concern. @Chen 133 R ¶11 top (Exh. 1002), IPR 12. This collection of “I like it”

documents selected by the user would constitute a “precondition.” WebMate analyzes the documents selected by the user (i.e. the “precondition”) and automatically finds other webpages for recommending to the user.

### Independent Claim 11

- a) (preamble) “A computer-implemented method for augmenting a directory comprising:”

Chen (Exh. 1002) describes the WebMate software agent, which automatically learns a user’s interest by extracting keywords from documents selected by the user and “automatically spiding news sources” (“spiding” means crawling the web, as a spider would) on the Internet for webpages having similar content. @Abstract.

- b) (claim 11) “autonomously initiating operation based upon a stored precondition”
- c) (claim 11) “accessing at least a first document without contemporaneous user selection, wherein said first document comprises at least in part topical textual content”

Claim steps (b) and (c) are taken together because step (b) of autonomously initiating operation of the software causes step (c) of accessing the documents “without contemporaneous user selection.” As explained above, the term “autonomously” is synonymous with “automatically.” @IPR 5. The term “precondition” means the selection of items that the user has identified as being of interest and encompasses at least “user placement of one or more documents in a file system directory as reference material for guiding the search” such as “saving topically-related document links in the same web-favorites folder.” @IPR 7.

WebMate works by “learning user interests incrementally and with continuous update and automatically providing documents (e.g. a personalized newspaper) that match the user interests.” @132 R ¶bott. WebMate develops the user profile automatically and “unobtrusively” by keyword extraction from documents that the user selects by marking them as “I like it.” @132 L Abstract, 134 L mid. The “user can provide multiple pages as similarity/relevance guidance for the search.” @132 L Abstract. Thus, WebMate access the user-selected documents “without contemporaneous user selection,” which is understood to mean “no user input of search parameters” and “no user input of search locations.” @IPR 6.

“In order to save on storage space,” WebMate “doesn’t keep any of the previous positive example documents.” @133 R ¶top. However, storage space is not a concern for higher capacity machines and it would be obvious to keep the set of “I like it” documents selected by the user as a collection of links instead of discarding them. This collection of “I like it” documents selected by the user would constitute a “precondition” recited in the claim. To be explained in detail below, these user-selected documents have “topical textual content.”

- d) (claim 11) “deriving at least one keyword indicative of at least one topical content within said first document”

Once the user has selected the items of interest, WebMate automatically parses the selected web documents and extracts words contained therein, including those in the title and headers. @133 R – 134 L. Using the extracted words, WebMate then

constructs a word vector for the document. @133 R ¶2<sup>nd</sup>. The word vector is an aggregate representation of the content as a series of word elements, each weighted according to its frequency of occurrence in the document. @Id. The combined word vectors become the user profile. @Id.

- e) (claim 11) “as a background operation, searching in storage in at least one computer for documents related by said at least one keyword to said first document”

With the user’s interest profile created, WebMate can search the entire web for pages relevant to the user’s interest, or monitor only a selection of websites designated by the user. @134 §3.2. If the user has selected a particular list of websites (URLs) for monitoring, WebMate can search these websites to find relevant webpages based on their similarity to the word vector in the user’s profile. @134 L ¶bott. If the user does not provide a list of websites to monitor, then WebMate can simply construct a search query “using the top several words in the current [user] profile and sends it to search engines (e.g. Altavista, Yahoo).” @134 R ¶2<sup>nd</sup>.

When WebMate is searching selected websites for pages that match the word vector in the user’s profile, or searching the entire web using the top words in the user’s profile, it is “searching ... for documents related by said at least one keyword to said first document.” These documents being searched by WebMate are “in storage in at least one computer.”

WebMate can perform its search “in the middle of the night when the network traffic is low” such that “[i]n the morning, the user can read the recommended personal newspaper.” @134 R top. WebMate can also perform its search contemporaneously “[i]f the result is needed immediately.” @Id. R ¶ 2<sup>nd</sup>. As “Related work,” Chen also points to the Letizia web browsing agent which “can recommend nearby pages by doing lookahead search.” @138 R top. Thus, Chen discloses three different timeframes in which searching can be performed: (1) overnight during machine idle time, (2) immediately, and (3) as a “lookahead search” while the user is browsing the Internet. From these teachings about the different timeframes for performing the search, it would be obvious to have WebMate perform searching “as a background operation” in the same manner as Letizia’s “lookahead search.”

f) (claim 11) “wherein at least some of said searched documents are independent and not organized in relation to one another”

Chen describes an experiment with WebMate in which 14 news sites covering technology news were monitored by WebMate for news articles matching the user’s interest. @134 R ¶ 3<sup>rd</sup>. The webpages on these 14 different news sites are “independent” and the webpages at one news site are “not organized in relation” to webpages at another news site.

g) (claim 11) “determining relevance of a search-accessed second document to said at least one keyword”

WebMate decides whether to recommend the pages it has found to the user on the basis of whether the page content “similarity is greater than some threshold” as

compared against the user's profile (which is represented by word vectors). @134 L ¶bott. (see also R ¶2<sup>nd</sup>).

h) (claim 11) “adding a reference to said second document in a results directory.”

If the user has selected websites for monitoring, WebMate can perform the search overnight and in the morning, provide the user with a “personal newspaper” listing the recommended pages. @134 L ¶bott. If WebMate performs a web-wide search, WebMate analyzes the search results and presents those meeting the required threshold for similarity “in descending order of relevance.” @134 R ¶2<sup>nd</sup>. In both of the aforementioned situations, WebMate works by “adding a reference to said second document [i.e. the recommended webpage] in a results directory.”

### Claim 12

Claim 12 depends from claim 11 and specifies that the “storage is on a plurality of computers connected to at least one network.” WebMate performs a search of the World Wide Web for documents that are relevant to the user's interest. The World Wide Web is a network of many different computers throughout the world. For example, in Chen's experiments with WebMate, the 14 news sites on the Internet selected for monitoring are on a “plurality of computers” in the World Wide Web network. @134 R ¶3<sup>rd</sup>.

### Claim 13

Claim 13 depends from claim 11 and adds the steps of “deriving a plurality of

keywords” and “determining relevance of said second document to said plurality of keywords.” This claim term is essentially reciting the plural form (“plurality of keywords”) of main claim 11, steps (d) and (g). @1PR 20. As already explained, WebMate can extract *multiple* keywords from the user-selected “I like it” documents and compare the searched documents against these multiple keywords.

#### Claim 14

Claim 14 depends from claim 11 and adds the step of “comparing the relevance of said second document to a preset threshold.” WebMate recommends a webpage to the user if its “similarity is greater than some threshold.” @134 L ¶1 bott. (see also R ¶2<sup>nd</sup>).

#### Claim 15

Claim 15 depends from claim 11 and adds the step of “conditionally adding said reference to said second document depending upon whether said reference to said second document already exists in said results directory.” This claim term is essentially stating that the software agent will not add a duplicate entry into the results directory. This is simply a common sense feature and there is nothing inventive about it.

#### Independent Claim 16

- a) (preamble) “A computer-implemented method for augmenting a directory comprising:”

Chen (Exh. 1002) describes the WebMate software agent, which automatically



learns a user's interest by extracting keywords from documents selected by the user and "automatically spiding news sources" ("spiding" means crawling the web, as a spider would) on the Internet for webpages having similar content. @Abstract.

b) (claim 16) "accessing a plurality of grouped documents without contemporaneous user selection initiating said access"

WebMate works by "learning user interests incrementally and with continuous update and automatically providing documents (e.g. a personalized newspaper) that match the user interests." @132 R ¶bott. WebMate develops the user profile automatically and "unobtrusively" by keyword extraction from documents that the user selects by marking them as "I like it." @132 L Abstract, 134 L mid. The "user can provide multiple pages as similarity/relevance guidance for the search." @132 L Abstract. Thus, WebMate accesses the user-selected documents "without contemporaneous user selection," which is understood to mean "no user input of search parameters" and "no user input of search locations." @IPR 6.

"In order to save on storage space," WebMate "doesn't keep any of the previous positive example documents." @133 R ¶top. However, storage space is not a concern for higher capacity machines and it would be obvious to keep the set of "I like it" documents selected by the user as links instead of discarding them. This collection of "I like it" documents selected by the user would constitute a "plurality of grouped documents" recited in the claim.

c) (claim 16) “deriving a plurality of keywords indicative of an aggregate content of said grouped documents”

As already explained above for claim 1(c), WebMate parses the user-selected webpages and generates a weighted word vector using the words contained therein. @IPR 12. The word vector elements are updated with words extracted from additional new documents selected by the user. @134 L ¶top. The final word vector represents the “aggregate content” of the group of documents selected by the user.

d) (claim 16) “prioritizing a relative relevance of said keywords; storing said plurality of keywords with regard to said relevance”

In WebMate, “[e]ach dimension of the vector space represents a word and its weight.” @133 R ¶2<sup>nd</sup>. Each word is weighted according to its frequency of occurrence. @Id. Thus, this assignment of weights to each of the words in the word vector is “prioritizing a relative relevance of said keywords.” The words (or dimensional variables representing the words) and their assigned weighting values are stored in the word vector.

e) (claim 16) “searching as a background operation storage in at least one computer for documents related to said plurality of stored keywords”

This claim term is essentially stating that the search for similar documents on the Internet or network occurs as a background process on the computer. With the user’s interest profile created, WebMate can search the entire web for pages relevant to the user’s interest, or monitor only a selection of websites designated by the user. @134 §3.2. If the user has selected a particular list of websites (URLs) for monitoring,

WebMate can search these websites to find relevant webpages based on their similarity to the word vector in the user's profile. @134 L ¶11 bott. If the user does not provide a list of websites to monitor, then WebMate can simply construct a search query "using the top several words in the current [user] profile and sends it to search engines (e.g. Altavista, Yahoo)." @134 R ¶2<sup>nd</sup>.

When WebMate is searching selected websites for pages that match the word vector in the user's profile, or searching the entire web using the top words in the user's profile, it is "searching ... for documents related to said plurality of stored keywords." The webpages being examined by WebMate are in "storage in at least one computer," i.e. on the computers running the websites.

WebMate can perform its search "in the middle of the night when the network traffic is low" such that "[i]n the morning, the user can read the recommended personal newspaper." @134 R top. WebMate can also perform its search contemporaneously "[i]f the result is needed immediately." @Id. R ¶ 2<sup>nd</sup>. As "Related work," Chen also points to the Letizia web browsing agent which "can recommend nearby pages by doing lookahead search." @138 R top. Thus, Chen discloses three different timeframes in which searching can be performed: (1) overnight during machine idle time, (2) immediately, and (3) as a "lookahead search" while the user is browsing the Internet. From these teachings about the different timeframes for

performing the search, it would be obvious to have WebMate perform “searching as a background operation” in the same manner as Letizia’s “lookahead search.”

f) (claim 16) “determining relevance of a found second document to said plurality of stored keywords”

WebMate decides whether to recommend the pages it has found to the user on the basis of whether the page content “similarity is greater than some threshold” as compared against the user’s profile (which is a word vector). @134 L ¶1 bott. (see also R ¶2<sup>nd</sup>).

g) (claim 16) “conditionally adding a reference to said second document in a results directory.”

If the user has selected websites for monitoring, WebMate can perform the search overnight and in the morning, provide the user with a “personal newspaper” listing the recommended pages. @134 L ¶1 bott. If WebMate performs a web-wide search, WebMate analyzes the search results and presents those meeting the required threshold for similarity “in descending order of relevance.” @134 R ¶2<sup>nd</sup>. In both of the aforementioned situations, WebMate works by “conditionally adding a reference to said second document [i.e. the recommended webpage] in a results directory.”

### Claim 17

Claim 17 depends from claim 16 and adds the step of “comparing the relevance of said second document to a preset threshold.” WebMate decides to recommend a found page to a user if the “similarity is greater than some threshold.” @134 L ¶1 bott. (see also R ¶2<sup>nd</sup>).

### Claim 18

Claim 18 depends from claim 16 and specifies that the “storage is on a plurality of computers connected to at least one network.” WebMate performs a search of the World Wide Web for documents that are relevant to the user’s interest. The World Wide Web is a network of many different computers throughout the world. For example, in Chen’s experiments with WebMate, the 14 news sites on the Internet selected for monitoring are on a “plurality of computers” in the World Wide Web network. @134 R ¶3<sup>rd</sup>.

### Claim 19

Claim 19 depends from claim 16 and adds the step of “adding a duplicate reference in said results directory is avoided.” This claim term is essentially stating that the software agent will not add a duplicate entry into the results directory. This is simply a common sense feature and there is nothing inventive about it.

### Claim 20

Claim 20 depends from claim 16 and adds the step of “adding a reference that was previously deleted from said results directory is avoided.” When a user deletes a reference, that would indicate that the user is no longer interested in that reference. Lieberman (Exh. 1003) suggests how the web browsing agent should handle this. As explained above, Chen specifically cites to Lieberman as “Related work.” @Chen 138

R, #17 (Exh. 1002). Therefore, someone reading Chen would consider the teachings of Lieberman to be relevant.

Lieberman describes the Letizia software agent, which explores the web for other items that may interest the user. @924 L Abstract (Exh. 1003). Lieberman explains how the user's interest can be inferred from browsing behavior. @925 R §4. Lieberman also considers a user's *disinterest* in a particular item, stating:

If the user returns immediately without having either saved the target document, or followed further links, an indication of *disinterest* can be assumed. Letizia saves the user considerable time that would be wasted exploring those "dead-end" links.

@925 R ¶3<sup>rd</sup> (*italics added*). Lieberman further adds:

Indications of interest probably ought to have a factor of decaying over time so that the agent does not get clogged with searching for interests that may indeed have fallen from the user's attention. Some actions may have been highly dependent upon the local context, and *should be forgotten* unless they are reinforced by more recent action.

@928 L ¶2<sup>nd</sup> (*italics added*).

Thus, Lieberman suggests that the browsing agent should carefully observe the user's behavior for indications of interests as well as *disinterests*. Certainly, if the user expressly shows disinterest in the reference by deleting it ("was previously deleted from said results directory"), then there is no reason to waste the user's time by adding what Lieberman calls "dead-end" links. As Lieberman suggests, interests that

have “fallen from the user’s attention ... should be forgotten,” and accordingly, there is strong motivation to program WebMate to avoid adding a “reference that was previously deleted from said results directory.”

### Conclusion

For the foregoing reasons, the challenged patent’s method of observing a user’s selections and recommending other similar selections to the user is obvious over prior web browsing agents that explore the web and make personalized recommendations for other webpages of interest.

JOHN J. YIM & ASSOCIATES LLC

/John J. Yim/

John J. Yim

Reg. No. 47,197

Lead Counsel for Petitioner

Listing of the Claims in Patent No. US 7,047,482

1. A computer-implemented method for augmenting a directory without contemporaneous user input comprising:
  - accessing at least a first document via a first directory without contemporaneous user selection of said first document, said first document comprising at least in part topical textual content;
  - deriving at least one keyword indicative of at least one topical content from said first document;
  - searching as a background operation a plurality of documents in storage in at least one computer without contemporaneous user input of a search location, such that said search comprises searching for documents related by said at least one keyword to said first document, thereby accessing a second document;
  - determining relevance of said second document to said at least one keyword;and  
adding a reference to said second document in a results directory.
2. The method according to claim 1, wherein at least part of said storage is on a different computer than the computer storing said first directory.
3. The method according to claim 1, further comprising deriving a plurality of keywords.
4. The method according to claim 3, further comprising ranking at least two of said plurality of keywords.
5. The method according to claim 1, further comprising accessing a plurality of documents in said first directory.



6. The method according to claim 1, with the additional step of signifying the relevance of said second document to documents in the first directory when displaying said results directory.
7. The method according to claim 1, with the additional step of comparing the relevance of said second document to a preset threshold.
8. The method according to claim 1, wherein said results directory is said first directory.
9. The method according to claim 1, with the additional step of displaying said results directory.
10. The method according to claim 1, further comprising recognizing a precondition for autonomously augmenting said results directory, prior to accessing said first document.
11. A computer-implemented method for augmenting a directory comprising:
  - autonomously initiating operation based upon a stored precondition;
  - accessing at least a first document without contemporaneous user selection, wherein said first document comprises at least in part topical textual content;
  - deriving at least one keyword indicative of at least one topical content within said first document;
  - as a background operation, searching in storage in at least one computer for documents related by said at least one keyword to said first document, wherein at least some of said searched documents are independent and not organized in relation to one another;
  - determining relevance of a search-accessed second document to said at least one keyword; and
  - adding a reference to said second document in a results directory.

12. The method according to claim 11, wherein said storage is on a plurality of computers connected to at least one network.
13. The method according to claim 11, further comprising:
  - deriving a plurality of keywords; and
  - determining relevance of said second document to said plurality of keywords.
14. The method according to claim 11, further comprising comparing the relevance of said second document to a preset threshold.
15. The method according to claim 11, further comprising conditionally adding said reference to said second document depending upon whether said reference to said second document already exists in said results directory.
16. A computer-implemented method for augmenting a directory comprising:
  - accessing a plurality of grouped documents without contemporaneous user selection initiating said access;
  - deriving a plurality of keywords indicative of an aggregate content of said grouped documents;
  - prioritizing a relative relevance of said keywords;
  - storing said plurality of keywords with regard to said relevance;
  - searching as a background operation storage in at least one computer for documents related to said plurality of stored keywords;
  - determining relevance of a found second document to said plurality of stored keywords;
  - conditionally adding a reference to said second document in a results directory.

17. The method according to claim 16, with the additional step of comparing the relevance of said second document to a preset threshold.
18. The method according to claim 16, wherein said storage is on a plurality of computers connected to at least one network.
19. The method according to claim 16, wherein adding a duplicate reference in said results directory is avoided.
20. The method according to claim 16, wherein adding a reference that was previously deleted from said results directory is avoided.

## CERTIFICATE OF SERVICE

The undersigned hereby certifies that the foregoing Petition for *Inter Partes* Review of Patent No. 7,047,482 and associated Exhibits 1001-1004 were served on \_\_\_\_\_, via FEDEX courier to the following addresses:

Eugenio Torres-Oyola  
FERRAIUOLI LLC  
221 Ponce De Leon Ave., 5th Floor  
San Juan, Puerto Rico 00917  
*Patent owner's correspondence address  
of record*

Stephen B. Brauerman  
BAYARD, P.A.  
222 Delaware Avenue, Suite 900  
Wilmington, DE 19899  
*Additional address likely to effect service*

JOHN J. YIM & ASSOCIATES LLC

/John J. Yim/  
John J. Yim  
Reg. No. 47,197  
Lead Counsel for Petitioner

# Exh. 1001



US007047482B1

(12) **United States Patent**  
**Odom**

(10) **Patent No.:** US 7,047,482 B1  
(45) **Date of Patent:** May 16, 2006

(54) **AUTOMATIC DIRECTORY SUPPLEMENTATION**

(76) **Inventor:** Gary Odom, 15505 SW. Bulrush La., Tigard, OR (US) 97223

(\* ) **Notice:** Subject to any disclaimer, the term of this patent is extended or adjusted under 35 U.S.C. 154(b) by 623 days.

(21) **Appl. No.:** 09/796,235

(22) **Filed:** Feb. 28, 2001

(51) **Int. Cl.**  
G06F 7/76 (2006.01)  
G06F 17/21 (2006.01)

(52) **U.S. Cl.** ..... 715/500; 707/5; 707/3

(58) **Field of Classification Search** ..... 715/501.1, 715/500; 707/3, 5; 709/202  
See application file for complete search history.

(56) **References Cited**

**U.S. PATENT DOCUMENTS**

5,535,382 A	7/1996	Ogawa	
5,598,557 A	1/1997	Doner	
5,680,511 A	10/1997	Baker	
5,893,092 A *	4/1999	Driscoll	707/5
6,100,890 A *	8/2000	Bates et al.	715/826
6,122,647 A	9/2000	Horowitz	707/513
6,138,157 A	10/2000	Welter	709/224
6,182,133 B1	1/2001	Horvitz	709/223
6,184,886 B1	2/2001	Bates	345/357
6,480,853 B1 *	11/2002	Jain	707/5
6,493,702 B1 *	12/2002	Adar et al.	707/3

6,584,468 B1	6/2003	Gabriel	
2001/0039563 A1*	11/2001	Tian	709/202
2003/0195877 A1*	10/2003	Ford et al.	707/3

**OTHER PUBLICATIONS**

Maglio, Paul et al., Communications of the ACM, vol. 43, No. 8; "Intermediaries Personalize Information Streams" (Aug. 2000).\*

Lieberman, Henry et al., Communications of the ACM, vol. 44, No. 8; "Exploring the Web with Reconnaissance Agents" (Aug. 2001).\*

Balabanovic, Marko et al., Communications of the ACM, vol. 30, No. 3; "Content-Based, Collaborative Recommendation" Mar. 1997).\*

Geisler, Gary et al., JCDL '01 in Roanoke, Virginia; "Developing Recommendation Services for a Digital Library with Uncertain and Changing Data" (Jun. 24-28, 2001).\*

Grasso, Antonietta et al., GROUP 99 in Phoenix, Arizona; "Augmenting Recommender Systems by Embedding Interfaces into Practices" (1999).\*

Voss, Angi et al., GROUP 99 in Phoenix, Arizona; "Concept Indexing" (1999).\*

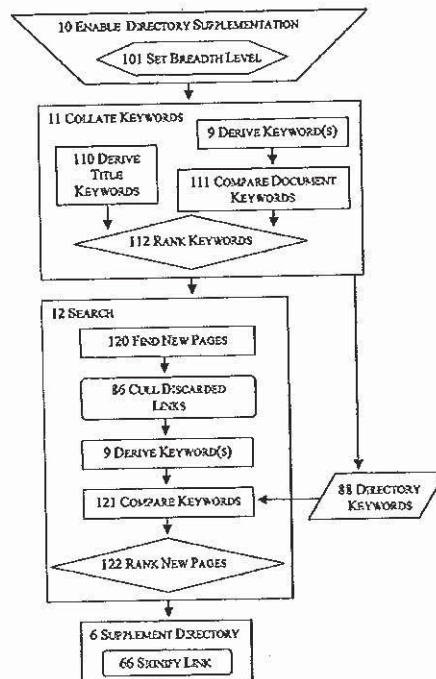
\* cited by examiner

Primary Examiner—Doug Hutton

(57) **ABSTRACT**

The present invention is computer software that automatically finds, saves, and displays links to documents topically related to document links residing in a directory without a user having to search.

20 Claims, 5 Drawing Sheets



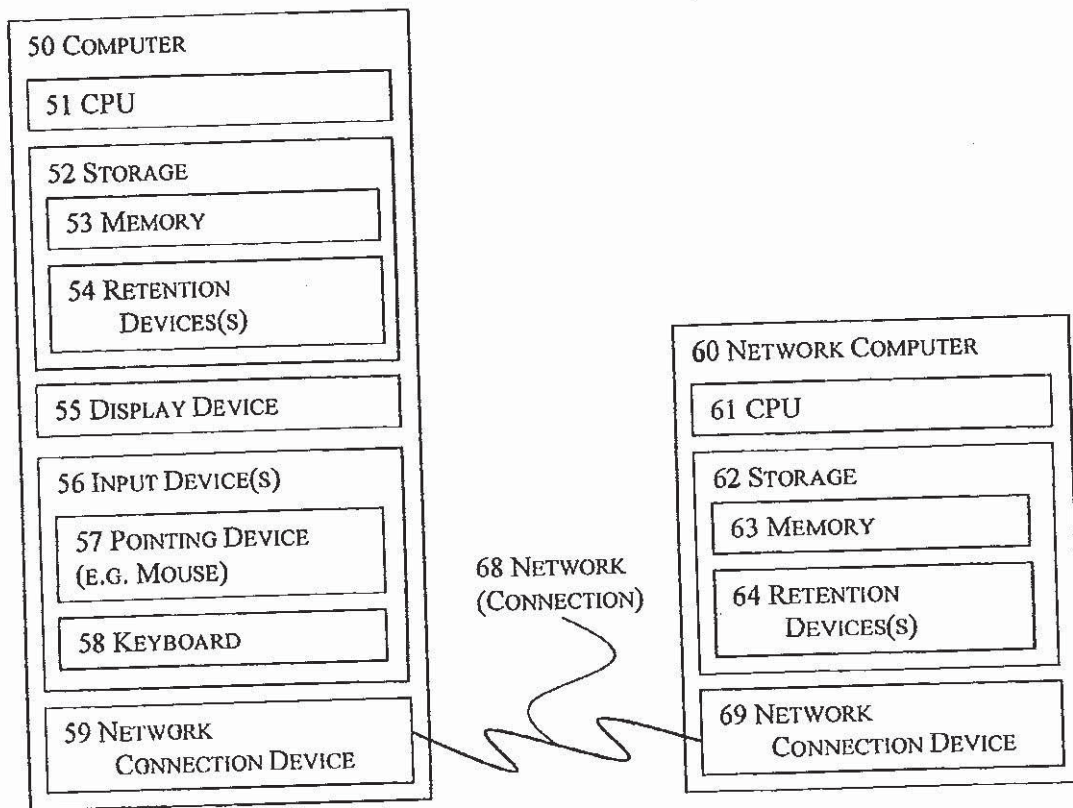


FIGURE 1

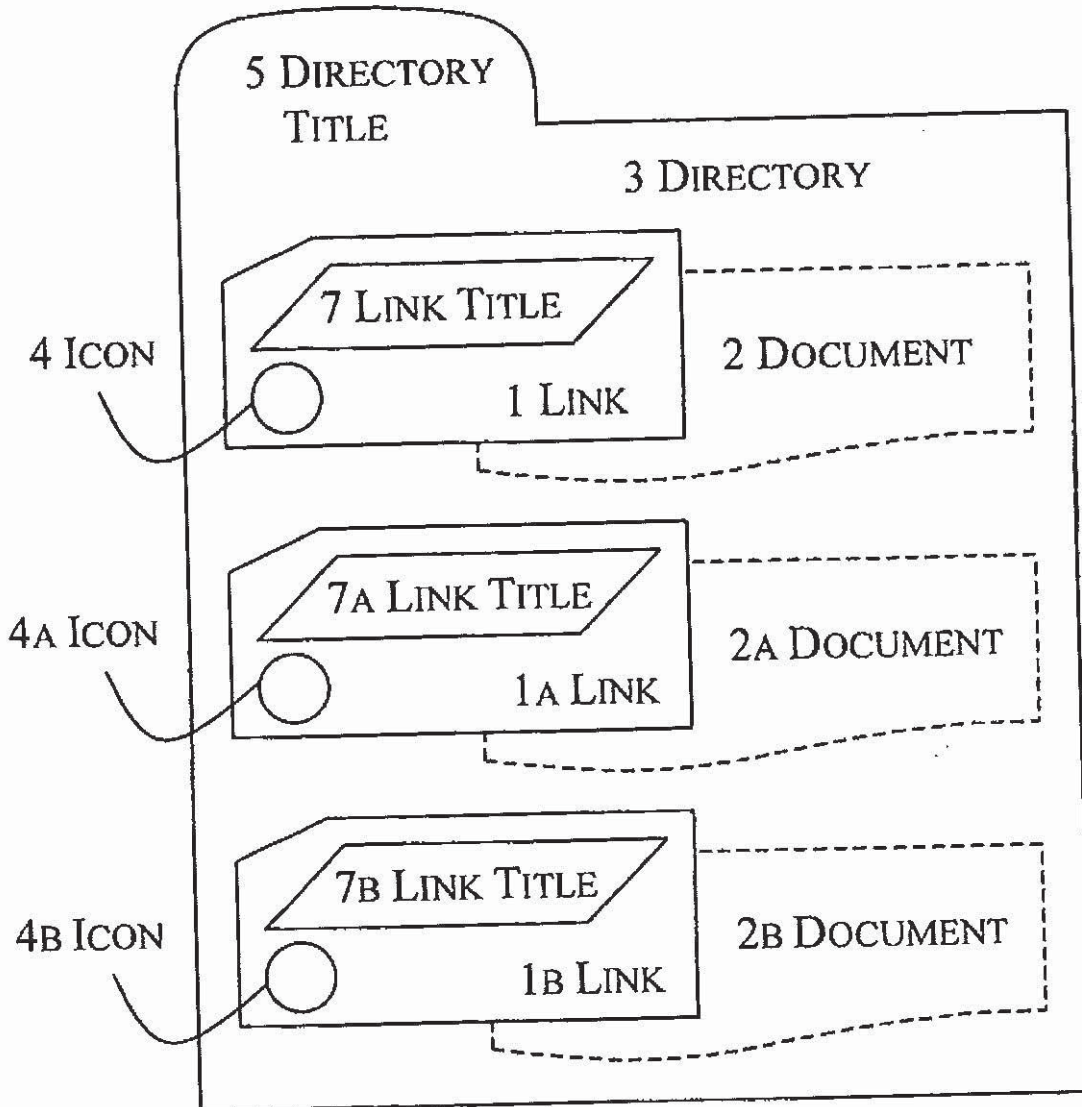


FIGURE 2



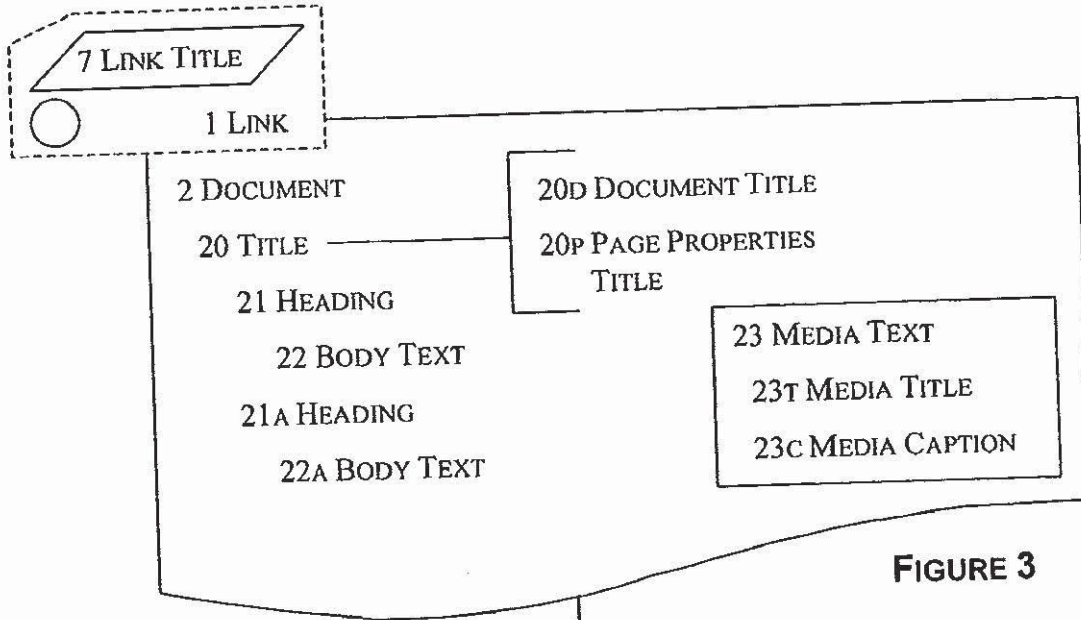


FIGURE 3

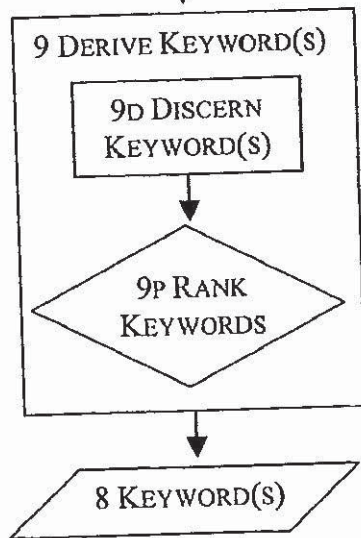


FIGURE 4

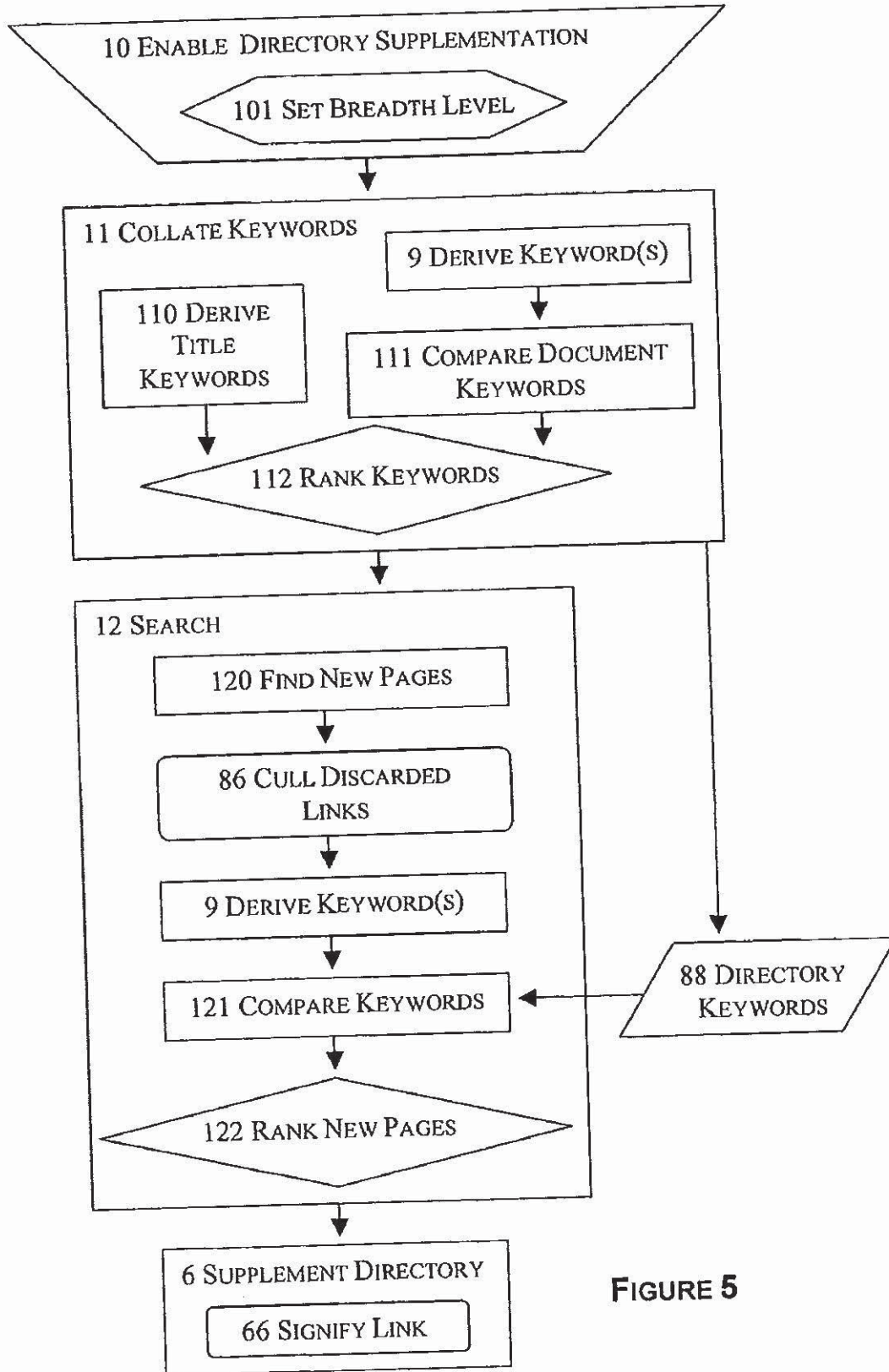


FIGURE 5

U.S. Patent

May 16, 2006

Sheet 5 of 5

US 7,047,482 B1

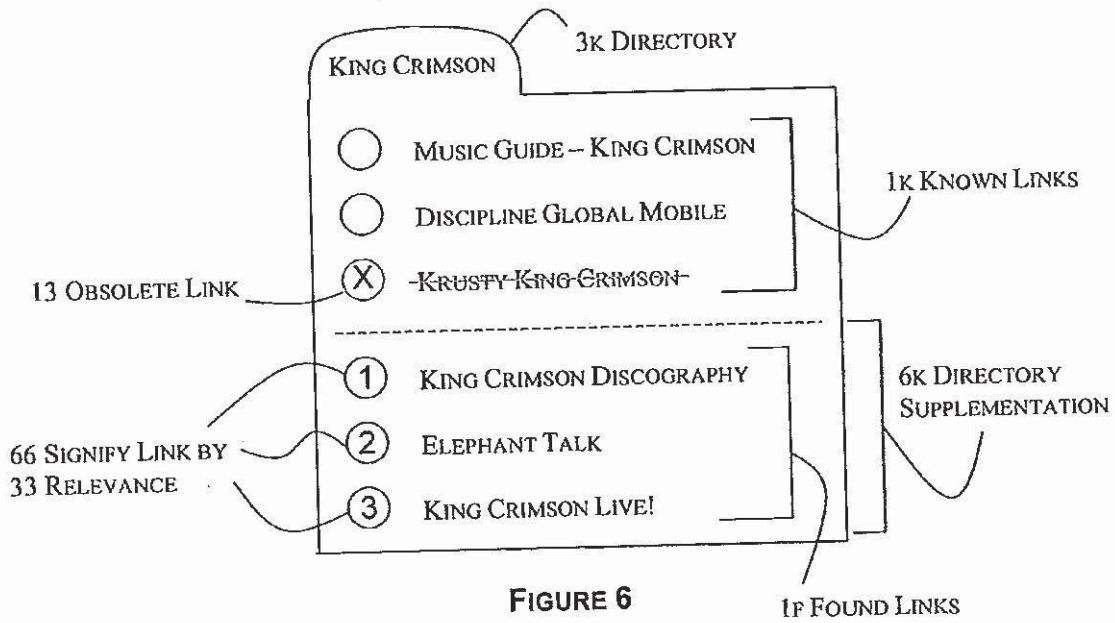


FIGURE 6

US 7,047,482 B1

1  
AUTOMATIC DIRECTORY  
SUPPLEMENTATION

TECHNICAL FIELD

The present invention relates generally to information retrieval systems, and more particularly, to automatically finding and displaying related document links without user-initiated searching.

BACKGROUND OF THE INVENTION

The Internet has become the world's information retrieval system. One of the distinguishing features of Internet (and intranet) documents is the use of embedded document links. Such a link is a portion of a source document that links to a target document: another document, or a different section of the same document. The other document may be on any computer system on a network supporting the appropriate communication protocols. Selecting a link navigates from the source document to the target document.

A web site is a collection of linked documents accessible through the World Wide Web, a part of the Internet. Such documents are commonly called web pages. Typically a web site has a "home page" that is the entry document into the site. The World Wide Web is commonly referred to as "the web".

Web pages commonly use a description language such as HTML (hypertext markup language) or XML (extensible markup language) to embed links and provide document formatting.

A link on a web page is by convention expressed as a uniform resource locator (URL). A link is often associated with a word or phrase in a source document, hence the common nomenclature: hypertext link. But a link may also be associated with images, or controls such as buttons, menus, and the like.

A web browser is a program for displaying web pages. Examples of popular web browsers include Microsoft Internet Explorer and Netscape Navigator.

Web browsers allow users to create and maintain directories of web page links. Such directories are commonly represented as folders or, sometimes, tabs.

New web pages or web sites are commonly found by links in known documents, or by keyword search. Users typically topically group links to related documents in self-titled directories, the directory title being the common topic of links within it.

Web sites are often extensive enough (so many pages) that a site typically offers a search facility for the site; commercial web sites almost always offer site search. Search refers to inquiry based upon one or more keywords (search terms). Search engines that search a multitude of sites abound on the web. A good search engine provides a commercial advantage. Some search engines, and some commercial products, such as Copernic® from Copernic Technologies, tap into multiple search engines to conglomerate searches.

Based upon keywords, quality search engines glean the most probably related pages using a confluence of linguistic analysis methods. Word location analysis is based upon the assumption that the topic of a document is specified in the title, headings, or the early paragraphs of text. Word frequency analysis counts the number of times search terms appear in a document. Syntactic analysis processes the grammatical structure of a document, serving to indicate nouns and verbs. Semantic analysis interprets the contextual meaning of words by examining word relationships. Mor-

2

phological analysis reduces verbs and nouns to their base form, providing a basis for direct word matching. At least one commercial product, LinguistX® from Luxight Software, provides advanced natural language text analysis.

In spite of software sophistication, as every experienced web user knows, user-initiated keyword search can be vexing; searches commonly return a plethora of pages, many unrelated to the desired topic. Search for 'watch', for example, thinking time pieces, and you'll likely end up with a bushel of pages about voyeurism. Careful application of search terms yields more relevant links, but the process and results are problematic: beyond searching for "this 'and' that", search Boolean logic is not exactly intuitive; different search engines have different syntaxes for search Boolean logic, and different ways to apply it, making that bit of business even less amenable; a bit of search pruning still leaves an abundance of junk, while a search result leaving out the chaff probably leaves out a good bit of wheat too.

The technology of document linking, search, and software-based linguistic analysis are well established. Recent advances enhance utility in locating desired information. For example, the subject of U.S. Pat. No. 6,122,647 is dynamically linguistically analyzing the text of a user-selected portion of a target document and generating new links to related documents. The subject of U.S. Pat. No. 6,184,886 is allowing a user to generate and maintain a list of prioritized bookmarks (links) that allow later access to selected sites (documents). The subject of U.S. Pat. No. 6,182,133 is pre-fetching pages for later viewing, thus saving a user time retrieving documents.

SUMMARY OF THE INVENTION

The present invention automatically finds, saves, and displays links to documents topically related to a set of documents without a user having to search or specify search terms. An incidental aspect of the invention is automatically signifying links by their status.

BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a block diagram of computers suitable for practicing the invention.

FIG. 2 depicts a directory of links.

FIG. 3 depicts a document.

FIG. 4 depicts the process to derive keywords from a document.

FIG. 5 depicts the directory supplementation process.

FIG. 6 depicts an example of directory supplementation.

DETAILED DESCRIPTION OF THE  
INVENTION

FIG. 1 is a block diagram of a computer 50 connected to a network computer 60 through a network 68. A computer 50 comprises at least a CPU 51; storage 52, which comprises memory 53 and optionally one or more devices with retention medium(s) 54 such as hard disks, diskettes, compact disks, or tape; an optional display device 55; and optionally one or more input devices 56, examples of which include but are not exclusive to, a keyboard 58, and/or one or more pointing devices 57, such as a mouse. A computer 50 also optionally includes a device for connection to a network 59. A network computer 60 comprises at least a CPU 51; storage 52, which comprises memory 53 and optionally one or more devices with retention medium(s) 64 such as hard disks, diskettes, compact disks, or tape; and a device for commec-

## US 7,047,482 B1

3

tion to a network 59. In one embodiment, a computer 50 is a client to a network computer 60 that is a server. A client-server environment is a setup whereupon one or more clients 50 are connected to one or more servers 60 through a network 68. A client 50 in a client-server environment primarily receives data. A server 60 primarily transmits data to be received by one or more clients 50. A peer-to-peer network is a setup whereupon one or more computers 50 are connected to one another 60 with or without a server on the network 68. A computer 50 in a peer-to-peer environment shares data with other computers 60. A network 68 may be any means by which one or more computers 50 are connected to one or more other computers 60 for data transfer.

As depicted in FIG. 2, a directory 3, if not empty, comprises a set of documents 2, or a set of links 1 to documents 2, or a combination of documents 2 and links 1. A link 1 is a reference to a document 2. A user-determined directory title 5 may provide concise topic indication.

FIG. 3 depicts a document 2 to which a link 1 may refer, and document 2 components. A document 2 comprises at least a passage of text 22, and may optionally include one or more titles 20, section headings 21, or adjunctive text such as media titles 23r or captions 23c. A document 2 may comprise other components besides text, such as media objects. A media object is a non-text software entity, examples of which include a picture, video, or sound. Text related to a media object is media text 23.

FIG. 4 depicts keyword derivation 9. A keyword 8 is one or more words used as an indication of the contents of a document. A keyword 8 may be a combination of words: for example, the Grateful Dead are significantly different than being either grateful or dead.

Various linguistic analysis methods may be applied to documents 2 for keyword 8 derivation: lexical, word frequency, word placement, syntactic, semantic, or morphological. Such methods are known to those skilled in the art.

Automatically displaying a link 1 refers to displaying a link 1 of a found document 2 without a user having to manually add a link 1 to a directory 3.

Signifying a link 66 refers to visibly indicating the current status of a link 1. Examples of visible indication include color coding or other visible distinction of link 1 text, such as a font style; or striking icon 4: either the usual icon 4 color coded, or icons 4 indicating status. Examples of status include a newly found link 1, a level of relevance for a newly discovered link 1, or an obsolete link 13.

Attempting to retrieve a document 2 from a link 1 sometimes reveals that the link 1 is no longer valid: the document 2 is gone, having been moved or removed. In this instance, the link 1 should be signified 66 as obsolete 13 if its document 2 has certainly been removed, or, if a link 1 to a moved document 2 can be ascertained, the stored link 1 should be updated to reflect the new document's 2 location. Pages 2 or sites that have moved often temporarily leave a notice behind telling where the site or page 2 has moved to. In such an instance, software linguistic analysis of the notification can glean the new link 1.

Document 2 inaccessibility does not necessarily mean link obsolescence 13: other possible causes exist, such as, for example, temporary server problems at the document's 2 home site. A link 1 should be signified 66 obsolete 13 only if document 2 removal can be verified: inaccessibility over a prolonged period of time would be indicative. For example, by keeping track of attempted access times, link obsolescence 13 may be concluded given document 2 inaccessibility at different times of the day for over a period of a week or so. Sometimes, document 2 removal is noted on

4

a web page 2. In such an instance, software linguistic analysis of the notification can determine document 2 removal.

Titles 20, including document title 20b, and associated page properties title for web pages 20p, media object titles 23r, and headings (section titles) 21 are prime fodder for keywords. For a document 2 with a link 1, the link title 7 should also be considered for keyword derivation 9. Titles may be considered highly indicative of document topics/keywords 8. Likewise document headings 21, which can be identified by location, possibly font formatting, and isolation from body text 22; headings 21 in HTML documents are most always distinguished by font formatting, hence, easily identified.

Body text 22 may provide the bulk of information upon which keywords 8 are derived 9. A common technique is to highly regard the first paragraph of body text 22 (and the body text 22 immediately following headings 21) for keyword derivation 9, as the topic of a document 2 or section is typically revealed in the first paragraph (academically known as the "topic paragraph").

Once a document 2 has been analyzed and keywords discerned 9d, document 2 keywords 8 can be rated or ranked 9p. Factors esteeming a keyword 8 include the following: prominence and frequency primarily in titles 20 and secondarily in headings 21; prominence and frequency in topic paragraphs and media text 23. Otherwise, word frequency may be a primary keyword 8 indicator. A suggested method to rank keywords 9p is to use a point system to weigh relative prominence and frequency, where, for example, prominence may comprise two-thirds of a keyword's 8 score and frequency one-third. Keyword 8 relevancy rating schemes 9p are known to those skilled in the art.

FIG. 5 depicts the directory supplementation 6 process. Directory supplementation 6 must be enabled 10. Directory supplementation 6 may be enabled 10 by default, by software-determined protocol, or by user determination. Automatically supplementing a directory 6 refers to adding links 1 or documents 2 to a directory 3 without a user having to search 12 or manually add links 1 to that directory 3.

Optionally, a breadth threshold level may be set 101. A breadth threshold level is intended as user-determined setting that possibly adjusts the number and potential relevance of accepted documents 2. Greater breadth casts a wider net: more links 1 or documents 2 are retained, and vice versa. If a user desires closely related documents 2 as a product of directory supplementation 6, set a low breadth level 101.

A relation threshold level would the mirror image equivalent to a breadth threshold level 101: a higher setting would be indication to limit directory supplementation 6 to closely related documents 2, and vice versa. Level indication 101 may be ordinal or numeric, such as percentage.

In an embodiment where breadth level setting 101 is employed, the setting 101 may be applied before and/or after search 12. A search 12 may use a broader setting 101 than the user specified. If then directory supplementation 6 presents sparse results, a user may want to adjust to a broader setting 101: if broader documents 2 have already been retrieved, the outcome of a broadened search may appear to the user immediately (with presentation of additional links 1).

Documents 2 in a directory 3 are analyzed 9 for keywords 8. Derived keywords 8 and attendant data may be stored to avoid repetition of the process 9. Attendant keyword data 8 may include keyword 8 rating data, such as keyword frequency and prominence in a document 2.

## US 7,047,482 B1

5

Though titles are necessarily terse, that very terseness makes directory 5 and link titles 7 an esteemed source of keywords 8. If directories 3 are hierarchical, topical information regarding a nested (lower level) directory 3 may be gleaned 110 by looking up the directory title 5 hierarchy. 5 Title-derived 110 keywords 8 may be given the highest regard.

The final step in keyword collation 11 is ranking 112 the gleaned sets of keywords 8 from directory 3 documents 2 by cumulating and collating keywords 11. This is, in essence, 10 a way of comparing documents via their derived keywords 8. If a document's 2 keywords 8 vary markedly from other documents 2 in its directory 3, that document's 2 keywords 8 may be disregarded. The outcome is a set of directory keywords 88 which may be retained, along with attendant data 15 or intermediate results, to avoid unnecessary repetition of the directory keyword collation process 11.

A Boolean logic search 12 for relevant documents 2 throughout all or part of a computer's or network storage (52, 62) proceeds based upon directory keywords 88. Candidate documents 2 may be found using cursory search 120 20 techniques, as winnowing may occur after documents 2 are found.

Once candidate documents 2 are found 120, links 1 to pages 2 or sites previously eliminated from the target 25 directory 3 may be culled 86. The obvious implication is that to perform this function, previously deleted links 1 from a directory 3 must be remembered (though no longer displayed). Culling discarded links 86, though optional, is 30 highly recommended, as not doing so degrades utility: making a user discard the same links 3 repeatedly would annoy the user.

Candidate document 2 keywords 8 are derived 9, then compared 121 to directory keywords 88. Unlike keyword 35 collation 11, where keywords 8 may be incorporated (albeit on a prioritized basis), candidate document keyword comparison 121 to directory keywords 88 is a critical fitness evaluation which provides the basis for ranking candidate documents 122 for directory supplementation 6. A variety of 40 methods for rating found documents 122 for relevance 33 to target keywords 88 are known to those skilled in the art.

Links 1 to pages 2 on the same site may be collated into a single link 1. This may be done after analyzing the pages 2 to determine the page 2 most closely related 33 to the 45 desired information. As a result, the selected link 1 for supplementation 6 may be the site's home page 2, the top-most page 2 for that topical aspect of the site, or the particular page 2 with the most relevant information. A standout page 2 should not be hidden: in the instance of a 50 fairly relevant site with a spot-on page 2, the smart choice is to use both.

Finally, in the preferred embodiment, the target directory 3 is supplemented 6 with links 1, concomitant to breadth 55 level setting 101 if employed. Optionally, visibly signify links 66 to indicate relevance 33. In an alternate embodiment, the target directory 3 is supplemented 6 with newly found documents 2 in a manner similar to the preferred embodiment.

FIG. 6 depicts an example directory 3k of links relating to the musical group King Crimson. The top section of the 60 directory 3k shows existing links 1k. During the process of checking known linked documents 2 to derive 9 keywords 8, the "Krusty King Crimson" link is found obsolete 13, and visibly signified as such. The bottom section of the directory 3k illustrates directory supplementation 6k. In the depicted 65 example, three newly discovered links 1f are displayed, along with indication 66 of their respective relevance 33. If

6

a user had specified via breadth level setting 101 only displaying links 1 level 2 or better, the "King Crimson Live!" link 1f would not be displayed.

The invention claimed is:

1. A computer-implemented method for augmenting a directory without contemporaneous user input comprising: accessing at least a first document via a first directory without contemporaneous user selection of said first document, said first document comprising at least in part topical textual content; deriving at least one keyword indicative of at least one topical content from said first document; searching as a background operation a plurality of documents in storage in at least one computer without contemporaneous user input of a search location, such that said search comprises searching for documents related by said at least one keyword to said first document, thereby accessing a second document; determining relevance of said second document to said at least one keyword; and adding a reference to said second document in a results directory.
2. The method according to claim 1, wherein at least part of said storage is on a different computer than the computer storing said first directory.
3. The method according to claim 1, further comprising deriving a plurality of keywords.
4. The method according to claim 3, further comprising ranking at least two of said plurality of keywords.
5. The method according to claim 1, further comprising accessing a plurality of documents in said first directory.
6. The method according to claim 1, with the additional step of signifying the relevance of said second document to documents in the first directory when displaying said results directory.
7. The method according to claim 1, with the additional step of comparing the relevance of said second document to a preset threshold.
8. The method according to claim 1, wherein said results directory is said first directory.
9. The method according to claim 1, with the additional step of displaying said results directory.
10. The method according to claim 1, further comprising recognizing a precondition for autonomously augmenting said results directory, prior to accessing said first document.
11. A computer-implemented method for augmenting a directory comprising: autonomously initiating operation based upon a stored precondition; accessing at least a first document without contemporaneous user selection, wherein said first document comprises at least in part topical textual content; deriving at least one keyword indicative of at least one topical content within said first document; as a background operation, searching in storage in at least one computer for documents related by said at least one keyword to said first document, wherein at least some of said searched documents are independent and not organized in relation to one another; determining relevance of a search-accessed second document to said at least one keyword; and adding a reference to said second document in a results directory.
12. The method according to claim 11, wherein said storage is on a plurality of computers connected to at least one network.

US 7,047,482 B1

7

13. The method according to claim 11, further comprising:

deriving a plurality of keywords; and  
determining relevance of said second document to said plurality of keywords.

14. The method according to claim 11, further comprising comparing the relevance of said second document to a preset threshold.

15. The method according to claim 11, further comprising conditionally adding said reference to said second document depending upon whether said reference to said second document already exists in said results directory.

16. A computer-implemented method for augmenting a directory comprising:

accessing a plurality of grouped documents without con-  
temporaneous user selection initiating said access;  
deriving a plurality of keywords indicative of an aggregate content of said grouped documents;  
prioritizing a relative relevance of said keywords;  
storing said plurality of keywords with regard to said  
relevance;

8

searching as a background operation storage in at least one computer for documents related to said plurality of stored keywords;

determining relevance of a found second document to said plurality of stored keywords;

conditionally adding a reference to said second document in a results directory.

17. The method according to claim 16, with the additional step of comparing the relevance of said second document to a preset threshold.

18. The method according to claim 16, wherein said storage is on a plurality of computers connected to at least one network.

19. The method according to claim 16, wherein adding a duplicate reference in said results directory is avoided.

20. The method according to claim 16, wherein adding a reference that was previously deleted from said results directory is avoided.

\* \* \* \* \*