UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

UIPATH, INC.,

Petitioner,

v.

RULE 14 LLC.,

Patent Owner.

Case No. IPR2024-01306
U.S. Patent No. 11,048,712

DECLARATION OF DR. JAIDEEP SRIVASTAVA, PH.D.

TABLE OF CONTENTS

I. INTRODUCTION 1

II. EXPERIENCE AND QUALIFICATIONS 1

III. MATERIALS CONSIDERED 4

IV. LEGAL STANDARDS 6

A. Written Description Support 6

B. Anticipation 6

C. Obviousness 7

V. LEVEL OF ORDINARY SKILL IN THE ART 9

VI. THE SPECIFICATION AND CLAIMS OF THE '712 PATENT 10

A. Specification Of '712 Patent 10

B. Claims Of The '712 Patent 11

VII. THE '712 PATENT CLAIMS ARE NOT SUPPORTED BY THE '313 APPLICATION OR THE PROVISIONAL APPLICATION 13

A. The Independent Claims Of The '712 Patent Correspond To The New Subject Matter Added At The Time The '712 Patent Was Filed 14

B. The Subject Matter Of The Independent Claims Of The '712 Patent Was Not Disclosed In Provisional Patent Appl. 61/679,024 15

C. The Subject Matter Of The Independent Claims Of The '712 Patent Was Not Disclosed In Patent Appl. 13/928,313 16

VIII. THE PROSECUTION HISTORY OF THE '712 PATENT 18

IX. CLAIM CONSTRUCTION 19

X. OVERVIEW OF PRIOR ART 19

A. The “Collection Selection” Problem 19

B. U.S. Patent No. 6,018,733 - Kirsch 23

C. U.S. Patent Publication No. 2006/0106793 - Liang 24

D. U.S. Patent No. 5,721,902 (Schultz) 29

E. U.S. Patent Publ. No. 2014/0040301 (the '301 Publication).....31

F. U.S. Patent No. 7,765,228 (Eichstaedt).....32

G. U.S. Patent No. 7,844,603 (Lucovsky).....32

H. WO Publication No. 01/65336 (Mellin).....32

I. U.S. Patent No. No. 7,043,482 (Vinsonneau) 32

J. U.S. Patent No. 6,243,713 (Nelson) 32

XI. SUMMARY OF GROUNDS FOR UNPATENTABILITY 33

XII. ALL OF THE CHALLENGED CLAIMS ARE OBVIOUS.....35

A. Ground 1 - Claims 1, 2, 4-8, 10, 11, 13-17 And 19-20 Are Obvious Over Kirsch, Schultz, Eichstaedt And Lucovsky.....35

1. Claim 1 35

2. Claims 10 and 19.....43

3. Claims 2, 11 and 20.....43

4. Claims 4 and 13.....46

5. Claims 5 and 14.....47

6. Claims 6 and 15.....48

7. Claims 7 and 16.....49

8. Claims 8 and 17.....50

B. Ground 2 - Claims 1, 4-8, 10, 13-17 And 19 Are Obvious Over Kirsch, Schultz, Eichsteadt And Mellin 50

1. Claim 1 50

2. Claims 4-8, 10, 13-17 And 19.....51

C. Ground 3 - Claims 2, 11 and 20 Are Obvious Over Kirsch, Schultz, Eichsteadt And Mellin.....52

D. Ground 4 - Claims 3 And 12 Are Obvious Over Kirsch, Schultz, Eichsteadt And Lucovsky, In View Of Liang 54

E. Ground 5 - Claims 9 and 18 Obvious Over Kirsch, Schultz, Eichsteadt And Lucovsky, And Further In View Of Vinsonneau 55

F. Ground 6 - Claims 9 and 18 Obvious Over Kirsch, Schultz, Eichsteadt And Mellin, And Further In View Of Vinsonneau.....56

G. Ground 7 - Claims 1-4, 7-8, 10-13, 16-17 And 19-20 Are Obvious Over Liang, Nelson And Lucovsky.....57

1. Claim 1.....57

2. Claims 10 and 19.....66

3. Claims 2, 11 and 20.....67

4. Claims 3 and 12.....68

5. Claims 4 and 13.....68

6. Claims 7 and 16.....68

7. Claims 8 and 17.....69

H. Ground 8 - Claims 1-8, 10-17 and 19-20 Are Obvious Over Liang, Nelson, Kirsch, Schultz and Lucovsky 69

1. Claims 1, 10 and 19.....69

2. Claims 2, 11 and 20.....72

3. Claims 3 and 12.....72

4. Claims 4 and 13.....72

5. Claims 5, 14 and 20.....73

6. Claims 6 and 15.....73

7. Claims 7 and 16.....74

8. Claims 8 and 17.....74

I. Ground 9 - Claims 1, 4, 7, 8, 10, 13, 16, 17 and 19 Are Obvious Over Liang, Nelson and Mellin.....74

1. Claim 1.....74

2. Claims 4, 7, 8, 10, 13, 16, 17 and 19 75

J. Ground 10 - Claims 2, 3, 5, 6, 11, 12, 14, 15 And 20 Are Obvious Over Liang, Nelson and Mellin In Combination With Kirsch And Schultz 76

1. Claims 2, 11 and 20.....76

2. Claims 3, 5, 6, 12, 14 and 15. 76

K. Ground 11 - Claims 9 and 18 Are Obvious Over Liang, Nelson, Lucovsky And Vinsonneau 77

L. Ground 12 - Claims 9 and 18 Are Obvious Over Liang, Nelson, Lucovsky, Kirsch And Schultz, In View Of Vinsonneau 77

M. Ground 13 - Claims 1-20 Are Obvious Over The '301 Publication And Schultz 78

1. Claim 1 78

2. Claims 10 and 19.....82

3. Claims 2, 11 and 20.....82

4. Claims 3 and 12.....82

5. Claims 4 and 13.....82

6. Claims 5 and 14.....83

7. Claims 6 and 15.....83

8. Claims 7 and 16.....83

9. Claims 8 and 17.....83

10. Claims 9 and 18.....84

XIII. DECLARATION 85

I. INTRODUCTION

1. My name is Dr. Jaideep Srivastava. I have been retained as an expert witness on behalf of Petitioner to provide expert opinions on the patentability of U.S. Patent No. 11,048,712 (the "'712 patent") (Ex. 1001).

2. I am being compensated at my normal rate, plus reimbursement for expenses, for my analysis. My compensation does not depend on the content of my opinions or the outcome of this proceeding.

II. EXPERIENCE AND QUALIFICATIONS

3. My qualifications are set forth in my curriculum vitae, a copy of which is submitted as Ex. 1005.

4. I received my Bachelor's degree from the Indian Institute of Technology (IIT), India (1983), and my Master's (1985) and PhD (1988) degrees from the University of California – Berkeley. All of my degrees are in Computer Science.

5. I am currently a Professor of Computer Science and Engineering at the University of Minnesota, a position I have held since 1988.

6. I have more than thirty five (35) years of technical expertise in the field of computer science and engineering, including experience with databases, optimization techniques, applications of graphs and database system design and architecture.

7. I have taught a range of database courses at the University of Minnesota, including core database courses like *Principles of Database Systems*, *Database System Implementation*, *Introduction to Database Research*, and specialized ones like *Introduction to Social Network Analysis* – which have strong database components.

8. I have supervised forty-six (46) doctoral dissertations and seventy (70) Masters theses/projects in the field of Computer Science & Engineering. Sixteen (16) of my PhD advisees themselves are professors in various universities across the world, while many of the rest are in some of the top technology companies. I have also mentored a number of university undergraduates and high schools students, many of which have gone on to careers in the STEM disciplines.

9. I have authored or co-authored over four hundred sixty (460) papers in peer-reviewed journals and conferences in the field of Computer Science & Engineering.

10. I am a co-inventor on 6 United States patents in the field of Computer Science & Engineering.

11. My research in the field of Computer Science & Engineering has been funded by various federal agencies including the National Science Foundation (NSF), the National Aeronautics and Space Administration (NASA), the Defense Advanced Research Projects Agency (DARPA), the Intelligence Advanced

Research Projects Agency (IARPA), the U.S. Army, and the US Air Force – via multiple, peer-reviewed grants. My research has also been supported by the State of Minnesota and private organizations.

12. I have worked with a number of commercial database systems while serving as a Data Mining Architect at Amazon from December 1999 through June 2000, and the Director of Data Analytics at Yodlee from July 2000 through June 2001. In my role as Director of Data Analytics at Yodlee, Inc., I led the creation of an analytics system from scratch, and was responsible for various technical decisions related to it.

13. I have advised a number of companies on database technologies, including Apertus, Honeywell, Eaton and United Technologies, among others.

14. I was an advisor to the State of Minnesota from 2006 through 2008 and the Government of India from 2011 through the present, in the use of database and other information technologies for more effective governance.

15. In 2010, I co-founded an analytics company, and was responsible for overseeing the end-to-end development of the first version of the system. I was responsible for the architecture and other technical aspects of the system, and leading the team that did the programming and development. A SQL database was at the core of the system, and I was personally responsible for the design, architecture, and implementation decisions around it.

16. I am a Fellow of the Institute of Electrical and Electronics Engineers (IEEE), the world's largest professional society of Electrical Engineering, Communications and Computing professionals. I have been an IEEE Distinguished Visitor, and received the PAKDD's Lifetime Achievement Award for my contributions to the field of data mining. I have served on various journal editorial boards and conference organizing committees – many times chairing the technical committees. I have been invited to deliver over fifty (50) keynote speeches, tutorials and other invited talks in national and international conferences.

17. I am an active researcher, with eight (8) current PhD and four (4) current MS students, working in the areas of Social Media Analytics and Health Informatics.

18. My analyses set forth in this Declaration are informed by my experience in the field of computer science, including computer architecture, operating systems, distributed systems, database systems, data mining and machine learning. Based on my above-described experience, I believe that I am considered to be an expert in these fields.

III. MATERIALS CONSIDERED

19. In preparing this Declaration, I have reviewed the '712 patent and file history. I also have considered the accompanying Petition, and the following documents cited in this Declaration:

Ex. 1001	U.S. Patent No. 11,048,712
Ex. 1002	Provisional Application No. 61/679,024, filed on August 2, 2012
Ex. 1003	File History of Patent Appl. 13/928,313, filed June 26, 2013
Ex. 1006	File History of Patent Appl. 16/228,218, filed December 20, 2018
Ex. 1007	J. Callan et al., *Searching Distributed Collections With Inference Networks*, Presented at the 18th Annual International Conference on Research and Development in Information Retrieval, 1995
Ex. 1008	Lu et al., *Measures in Collection Ranking Evaluation*, Technical Report TR-96-39, Computer Science Department, University of Massachusetts, 1996
Ex. 1009	French et al., *Comparing the Performance of Database Selection Algorithms*, Proceedings of the 22nd annual international ACM SIGIR conference on research and development in information retrieval
Ex. 1010	D'Souza et al., *Collection Selection for Managed Distributed Document Databases*, Information Processing & Management, Volume 40, Issue 3 (May, 2004)
Ex. 1011	U.S. Patent No. 6,018,733 (Kirsch et al.)
Ex. 1012	U.S. Patent Publication No. 2006/0106793 (Liang)
Ex. 1013	U.S. Patent No. 5,721,902 (Schultz)
Ex. 1014	U.S. Patent Publ. No. 2014/0040301 ('301 Publication)
Ex. 1015	U.S. Patent No. 7,765,228 to (Eichstaedt)
Ex. 1016	U.S. Patent No. 7,844,603 (Lucovsky)
Ex. 1017	WO Publication No. 01/65336 (Mellin)
Ex. 1018	U.S. Patent No. 7,043,482 (Vinsonneau)
Ex. 1019	U.S. Patent No. 6,243,713 (Nelson)
Ex. 1020	Infringement Contentions of Patent Owner in *Rule 14, LLC v. UiPath, Inc*., Case No. 2:23-cv-627 (E.D. Tex.)

In formulating my opinions, I have further relied upon my extensive experience.

IV. LEGAL STANDARDS

A. Written Description Support

20. I understand that a claim is not entitled to the benefit of an earlier filed application unless the subject matter of the claim is disclosed in the earlier filed application in a manner that reasonably conveys to those skilled in the art that the inventor had possession of and actually invented the claimed subject matter. I understand that adequate written description requires that one skilled in the art, reading the original disclosure, be able to immediately discern all limitations in the claim.

21. I understand that the test for adequate written description support requires an objective inquiry into the four corners of the specification from the perspective of a person of ordinary skill in the art. I understand that the specification need not recite the claimed invention *in haec verba*. I also understand, however, that a description that merely renders the invention obvious does not satisfy the written description requirement. In other words, entitlement to a filing date does not extend to subject matter which is not disclosed but would be obvious over what is expressly disclosed.

B. Anticipation

22. It is my understanding that the claims of a patent are anticipated by a prior art reference if each and every element of the claim is found either explicitly or inherently in a single prior art reference or system. I understand that inherency

requires a showing that the missing descriptive matter in the claim is necessarily or implicitly present in the allegedly anticipating reference, and that it would have been so recognized by a person of ordinary skill in the art at the time (POSITA). In addition, I understand that an enabling disclosure is one that allows a POSITA to make the invention without undue experimentation. Although anticipation typically involves the analysis of a single prior art reference, I understand that additional references may be used to show that the primary reference has an enabling disclosure, to explain the meaning of a term used in the primary reference, and/or to show that a characteristic is inherent in the primary reference.

C. Obviousness

23. I understand that a claim is unpatentable if the differences between the claimed subject matter and the prior art are such that the subject matter as a whole would have been obvious to a POSITA at the time the claimed subject matter was made. I understand that a patent claim may be obvious to a POSITA in view of the prior art teachings of a single reference or a combination of references. I understand that a claim is obvious based on a combination of references where a POSITA would have been motivated to make such a combination, and would have had a reasonable expectation of success in arriving at the claimed subject matter.

24. To assess obviousness, I understand that I am to consider the scope and content of the prior art, the differences between the prior art and the claim, the level of ordinary skill in the art, and any secondary considerations to the extent they exist.

25. It is also my understanding that there are principles that may be used as further guidance in obviousness analysis (especially when considering combinations of references), which include considering whether:

- the claimed subject matter is simply a combination of prior art elements according to known methods to yield predictable results;
- the claimed subject matter is a simple substitution of one known element for another to obtain predictable results;
- the claimed subject matter uses known techniques to improve similar devices or methods in the same way;
- the claimed subject matter applies a known technique to a known device or method to yield predictable results;
- the claimed subject matter would have been "obvious to try" choosing from a finite number of identified, predictable solutions, with a reasonable expectation of success;
- there is known work in one field of endeavor that may prompt variations of it for use in either the same field or a different one based

on design incentives or other market forces if the variations would have been predictable to a POSITA;

- there existed at the time of conception and reduction to practice a known problem for which there was an obvious solution encompassed by the patent's claims; and
- there is some teaching, suggestion, or motivation in the prior art that would have led a POSITA to modify the prior art reference or to combine prior art reference teachings to arrive at the claimed subject matter.

26. I understand that there are secondary considerations that may tend to show that a claimed invention would not have been obvious to a POSITA. These include, for example, commercial success, long-felt but unsolved needs, failure of others, and unexpected results. I further understand that there must be a relationship, or nexus, between any secondary considerations and the novelty of the claimed invention.

V. LEVEL OF ORDINARY SKILL IN THE ART

27. I am informed and understand that various factors can be considered in determining the skill and experience of a POSITA. I am informed and understand that those factors include: (1) the educational level of the inventors; (2) the type of problems encountered in the art; (3) prior art solutions to those problems; (4) the

rapidity with which innovations are made; (5) sophistication of the technology; and (6) education level of active workers in the field.

28. A POSITA as of December 4, 2015 would have had a working knowledge of techniques for querying and monitoring data sources as pertinent to the '712 Patent. A POSITA would have had a bachelor's degree in electrical engineering, computer science, or equivalent training, and approximately two years of work experience in software development with a basic knowledge of database searching techniques. My definition of the levels of education and experience of a POSITA would be the same as of August 2, 2012 (the filing date of Provisional Application No. 61/679,024 (Ex. 1002)) and June 26, 2013 (the filing date of Application No. 13/928,313.)

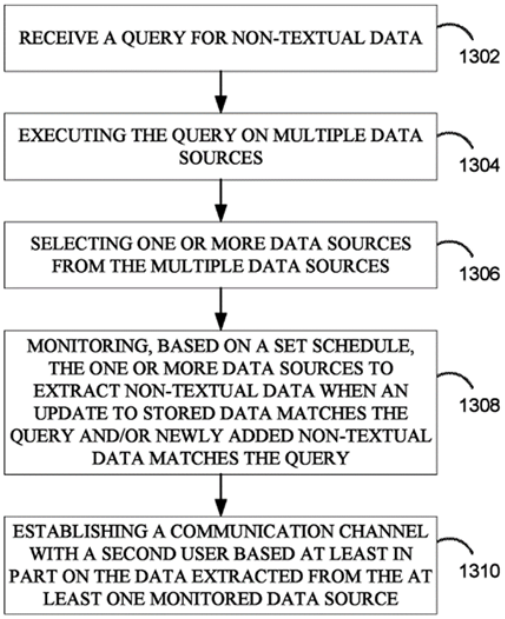
VI. THE SPECIFICATION AND CLAIMS OF THE '712 PATENT

A. Specification Of '712 Patent

29. The specification of the '712 Patent describes "a data mining system" in which "a query" for "non-textual data" is "generated by a user." Ex. 1001, 4:39-47. The "non-textual data" could be an image, video or audio. *Id*., 36-37. The query is "expanded" by, for example, supplementing the query terms with synonyms. *Id*., 4:51-57. Next, multiple "data sources"[1] are searched using the query and "the results

[1] The specification provides a variety of examples of different "data sources" that can be searched including "social media, web feeds, customer feedback, internal company databases, and proprietary data silos," "publicly available data, such as

of the query are reviewed." *Id*., 4:66-5:3; 5:27. "If the results are satisfactory, specific data sources may be selected to be monitored" periodically for updates. *Id*., 5:27-29. Monitoring results are communicated to users by "establishing a communication channel," which means sending an "email" or "short message service (SMS)" to users. *Id*., 5:41-43; 14:56-59. These steps are shown in Figure 13 of the '712 Patent, set forth below:



B. Claims Of The '712 Patent

30. Claim 1 of the '712 Patent closely tracks Figure 13 (above). Claim 1 recites:

---

data available via Wikipedia or the CIA fact book," "[s]ources that are indexed by search engines," "a blog," and "classified databases, data behind a firewall, or data that is available via a paid subscription, such as Lexis." Ex. 1001, 4:12-14, 5:12 and 5:24-26.

> 1. A method of analyzing data, comprising:
>
> receiving, from a first user, a query for non-textual data;
>
> expanding search terms of the query;
>
> executing the query on a plurality of data sources;
>
> selecting at least one data source from the plurality of data sources, the at least one data source being selected when results of the query are greater than or equal to an accuracy thresholds;
>
> monitoring, based on a set schedule, the at least one data source to extract non-textual data from the at least one data source when at least an update to stored data matches the query, newly added non-textual data matches the query, or a combination thereof; and
>
> establishing a communication channel with a second user based at least in part on the data extracted from the at least one monitored data source.

31. The "selecting" step in Claim 1 recites that data source(s) are selected "when results of the query are greater than or equal to *an accuracy thresholds*." (emphasis added). While the specification of the '712 Patent does not define "an accuracy thresholds," the specification equates the *relevancy* of the search results with the *accuracy* of the search results:

> ... As previously discussed, at the results module 502, *the user selects relevant results* for processing (e.g., calibration). *That is*, the user may judge the results *for accuracy* based on the intent of the query. *Id*., 12:25-28 (emphasis added).

32. Similarly, in its infringement contentions in parallel district court litigation (the "Litigation,") the Patent Owner has equated “an accuracy thresholds” with relevancy. Specifically, the Patent Owner's infringement contentions state:

> ... UiPath Communication Mining selects at least one data source from the plurality of data sources, the at least one data source being selected when results of the query are greater than or equal to *an accuracy thresholds*, including *by selecting a data source with information that is considered "most relevant" to the query*. Ex. 1020, at 35 (emphasis added).

Thus, based on both the specification and Patent Owner's infringement contentions, it is clear that in the context of the '712 Patent, a "relevancy" threshold is an example of "an accuracy thresholds."

VII. THE '712 PATENT CLAIMS ARE NOT SUPPORTED BY THE '313 APPLICATION OR THE PROVISIONAL APPLICATION

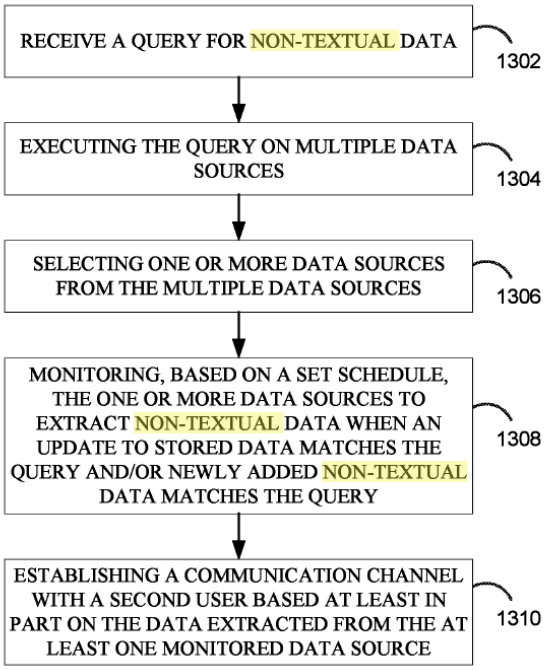
33. The '712 Patent was filed as Application No. 16/228,218 on December 20, 2018 as a continuation-in-part of Application No. 14/960,316, filed on December 4, 2015 (the '316 Application), which is a continuation-in-part of Application No. 13/928,313, filed June 26, 2013 (the '313 Application") (now the '977 Patent), which in turn claimed priority to Provisional Application No. 61/679,024, filed on August 2, 2012 (Ex. 1002) (the "Provisional Application.")

34. Neither the Provisional Application nor the '313 Application provide written description support for the Claims of the '712 Patent. Therefore, the effective filing date of the '712 Patent is no earlier than December 4, 2015 -- the filing date of Application No. 14/960,316.[2]

[2] I express no opinion as to whether the challenged claims are supported by the '316 Application. However, for purposes of this Petition, it is unnecessary to determine

A. The Independent Claims Of The '712 Patent Correspond To The New Subject Matter Added At The Time The '712 Patent Was Filed

35. All of the independent claims of the '712 Patent recite searching and monitoring for "*non-textual* data," as disclosed in Figure 13 of the '712 Patent, reproduced below (highlighting added):



36. Figure 13 and its accompanying description were added to the specification when the '712 Patent was filed on December 20, 2018, and thus *neither Figure 13 nor its accompanying description were included in any of the earlier*

---

whether the Claims of the '712 Patent are supported by the '316 Application because the latest prior art reference relied on herein is US 2014/0040301, published *February 4, 2014*. Thus, whether the Claims of the '712 Patent have a priority date of *December 15, 2015* (based on the '316 Application) or *December 20, 2018* (based on the '218 Application) is immaterial to whether any of the references relied on in this IPR are prior art.

applications to which priority is claimed. The description of Figure 13 (added when the '712 Patent was filed on December 20, 2018) closely tracks the language of Claim 1 (and the other independent claims) of the '712 Patent.

B. The Subject Matter Of The Independent Claims Of The '712 Patent Was Not Disclosed In Provisional Patent Appl. 61/679,024

37. The Provisional Application (Ex. 1002) was a one-page document, which appears to be an advertisement, directed to loss prevention, which is depicted below:

Ex. 1002, at 14.

38. Nothing in the Provisional Application discloses the "non-textual data" limitations in the claims of the '712 Patent.

C. The Subject Matter Of The Independent Claims Of The '712 Patent Was Not Disclosed In Patent Appl. 13/928,313

39. The '313 Application -- which matured in the '977 Patent -- makes no mention of "non-textual data." Ex. 1003, at 233-280. While the '313 Application discusses the use of a "sentry" that identifies *a phrase (i.e., text) that happens to be included in an image* posted on a data source (*id*., at 255, 279 ([00106] and Figure 11)), that disclosure fails to provide written description support for the independent claims of the '712 Patent for at least four reasons:

40. First, the independent claims of the '712 Patent recite that a query for *non-textual* data is received from the first user. Nothing in the discussion of Figure 11 in the '313 Application states that the first user is *querying for non-textu*al *data*. Rather, the description of Figure 11 teaches that the *text embedded within the image* is being queried. *Id*., at [00106] ("[t]he sentry may identify a *phrase from* a message and/or *image posted on a data source*, such as a message board and/or a social media site. The post may include an image 1104 and text 1106. ... Once the post 1102 has been identified, *the text 1106 may be analyzed* by a data analysis module.") Simply put, querying for *text* embedded in an image is not querying for "*non-textual* data."

The '313 Application does not state that any "*non-textual*" aspects of the image are queried.

41. Second, there is no discussion in connection with Figure 11 of the '313 Application that one or more *data sources are selected* based on results of a *"query for non-textual data*," as required by the independent claims of the '712 Patent. In describing Figure 11, the '313 Application states only that "the results may be *analyzed* by a data analysis module or a sentiment module." *Id.* A generic reference to "*analyzing results*" lacks sufficient specificity to provide written description support for *selecting data sources* based on results of a query for non-textual data.

42. Third, the '313 Application fails to connect the concept of "an *accuracy thresholds*" to "*a query for non-textual data*." In fact, there is no disclosure in the '313 Application as to how a POSITA would go about determining "an accuracy thresholds" in the case of a query for "*non-textual data*."

43. Fourth, the '313 Application fails to describe *extracting non-textual data* from a data source. In connection with Figure 11, the '313 Patent states only that *textual* information identifying the user of the post is extracted. *See*, Ex. 1003, at 255 [00106] ("the text 1106 may be analyzed ..."), and at 257 [00112] ("the information that is extracted may include the username, actual name, age, location, and/or any other information related to the user that posted the data."). The '313

Application does not disclose that *non-textual* aspects of the image are either *extracted or provided to a second user*, as claimed.

44. For at least these reasons, the independent claims of the '712 Patent are not supported by the '313 Application. Since neither the Provisional Application nor the '313 Application provide written description support for the independent claims of the '712 Patent, the effective filing date of the '712 Patent is no earlier than December 4, 2015 -- the filing date of Application No. 14/960,316.

VIII. THE PROSECUTION HISTORY OF THE '712 PATENT

45. During the prosecution history of the '712 Patent, the effective filing date of the claims was not discussed.

46. The prosecution history of the '712 Patent was limited to one Office Action consisting of double patenting rejections over, *inter alia*, the '977 Patent. As the Examiner recognized, the claims of the '712 Patent were not patentably distinct from those of the '977 Patent, because searching non-textual data was well known in the prior art. *See*, Office Action dated October 6, 2020, at 3-4 (Ex. 1006, at 91) ("US Patent Application 20120158739 by Ah-Pine et. al. ... discloses "the non-textual data comprises at least one of an image, audio, a video, or combination thereof" ... It would have been obvious to a person of an ordinary skill in the art at the time the invention was made to have to have applied image, audio, video, or combination thereof for non-textual data based on the disclosure Ah-Pine for the purpose of

querying and retrieving those objects.") The Patent Owner acquiesced in the double-patenting rejection by submitting a terminal disclaimer. Ex. 1006, at 82-83.

47. The Examiner allowed the Application in the next Office Action, finding that the prior art of record failed to show (i) selecting a data source from a plurality of data sources based on the accuracy (relevancy) of search results from the selected data source, (2) monitoring the selected data source(s) for updates, and (3) establishing a communication channel by, e.g., sending an email to a second user alerting them of the update. Ex. 1006, at 23-24 (reasons for allowance). Notably, the reasons for allowance were not based on the requirement of "non-textual data" in the claims.

IX. CLAIM CONSTRUCTION

48. I understand that claim terms should normally be given their ordinary and accustomed meaning as would have been understood by a POSITA as of the priority date in view of the specification and prosecution history. In formulating my opinions, I have applied the ordinary and accustomed meaning as would have been understood by a POSITA as of the priority date in view of the specification and prosecution history to the claims of the '712 Patent.

X. OVERVIEW OF PRIOR ART

A. The “Collection Selection” Problem

49. By 1995, hundreds of different data sources (referred to as “collections”) had become available on the Internet, but it was impractical to search

all of them for a given topic. The problem presented back in 1995 -- named the "collection selection" problem -- was how to narrow a search to a few useful collections. As one author explained:

> Users of commercial retrospective information retrieval systems have always faced the *collection selection* problem. The user must either search all collections or *choose the subset to be searched*.

J. Callan et al., *Searching Distributed Collections With Inference Networks*, Presented at the 18th Annual International Conference on Research and Development in Information Retrieval, 1995 (the "Callan Paper") (emphasis added) (Ex. 1007, at 1).

50. The Callan Paper proposed the following four-step solution to the "collection selection" problem:

> 1. Use the query to retrieve a ranked list of collections,
> 2. Select the top group of collections,
> 3. Search the top group of selections, in parallel or sequentially, and
> 4. Merge the results from the various collections into a single ranking.

Id., at 2.

51. In 1996, the Callan Paper was cited (as reference 1) in Lu et al., *Measures in Collection Ranking Evaluation*, Technical Report TR-96-39, Computer Science Department, University of Massachusetts, 1996 (the "Lu Paper")(Ex. 1008), which also described the "collection selection" problem:

> As hundreds or even thousands of collections become available on the Internet, it is practically impossible to query all of them when

> searching for information on a given topic. ... Consequently *the need to narrow the search to a few useful collections raises the issue of how to locate the most useful collections*. Several approaches have been proposed to solve this problem. ... In this paper, we survey the measures for collection ranking used in the current research ... Ex. 1008, at 1 (emphasis added).

52. The Lu Paper described several methods for ranking collections, including methods that scored collections based on relevancy. *Id*., at 2-8.

53. In 1999, the Lu Paper was cited (as reference 17) in French et al., *Comparing the Performance of Database Selection Algorithms*, Proceedings of the 22nd Annual International ACM SIGIR Conference On Research And Development In Information Retrieval (the "French Paper") (Ex. 1009). There, the authors again addressed the "collection selection" problem by comparing two solutions, i.e., *gGLOSS* and *CORI*. Ex. 1009, at 1-2. The *gGLOSS* approach ranked databases based on a "goodness" score that was a function of the similarities between the query and documents in a database. *Id*., at 2. The *CORI* approach represented each database based on the frequency of terms appearing in the documents in the database, and ranked the databases by analyzing the query against this frequency information. *Id*., at 2.

54. In 2004, the French Paper was cited[3] in D'Souza et al., *Collection Selection for Managed Distributed Document Databases*, Information Processing &

[3] *See*, list of References in D'Souza et al., Ex. 1010, at 19.

Management, Volume 40, Issue 3 (May, 2004) (Ex. 1010). There, the authors again addressed the "collection selection" problem, describing still further solutions that had emerged since the French Paper. In this regard, the authors stated:

> Broadcasting queries to all collections could well be costly or impractical, thus it is attractive to first rank the collections in the database ... The user can then query the top-ranked collections directly ... *Ranking the collections then provides a form of collection selection*.
>
> * * *
>
> Several techniques have been proposed for collection selection. ... There are two general kinds of techniques based on indexing. One is to gather the set of distinct terms, or lexicon, from each collection, together with some inter- and intra-collection statistics such as the number of documents containing each term; these statistics may then be used to estimate the appropriateness of each collection to a query. ... The other kind of selection technique is to index documents in the database of collections, then select at query time the collection with the most promising documents. Ex. 1010, at 2 (emphasis added).

55. Among other points relevant to the claims of the '712 Patent, D'Souza et al. observed that long queries "generated by *query expansion*" are "*better at collection selection*." *Id*., at 4 (emphasis added).

56. Thus, by 2004, the "collection selection" problem had been written about extensively, and numerous solutions had been developed.[4]

[4] The Callan Paper (Ex. 1007), Lu Paper (Ex. 1008), French Paper (Ex. 1009), and D'Souza (Ex. 1010) all relate to the field of database searching and the collection selection problem. As a person of ordinary skill in the art in the relevant field, I would have had an interest in the subject matter of these publications at the time of the invention of the '712 Patent. In addition, a person of ordinary skill in the art would have been able to find these publications with reasonable diligence at around the time of the invention of the '712 Patent.

B. U.S. Patent No. 6,018,733 - Kirsch

57. In 1997, Infoseek Corporation -- a search engine founded by Steve Kirsch -- filed an application that matured into U.S. Patent No. 6,018,733 entitled "Methods for Iteratively Performing Collection Selection In Full Text Searches." (Kirsch) (Ex. 1011). In the Background of his patent, Kirsch explained that "[w]hile access to much of the information available through the Internet is free for public use, numerous proprietary or fee-based access document collections exist." *Id*., 1:32-35. Kirsch noted that the "collection selection" problem "is complicated further when the opportunity and desire exists to search any combination of public and private document collections." *Id*., at 2:21-23.

58. In Kirsch, a user query "formed using any combination of a free text" (*id*., 6:40-43) is "expanded as appropriate into a list of specifically identified search terms." *Id*., 7:6-7. Kirsch disclosed a plurality of different "document collections," each of which is a "data source."[5] Kirsch taught initially directing the query against multiple document collections. *Id*., at 14:48-50 ("the user query 84 is preconditioned upon issuance from the user workstation 82 to *specify the particular set of collections 88, 90, 92 that the query is to be resolved against.*") Kirsch taught that

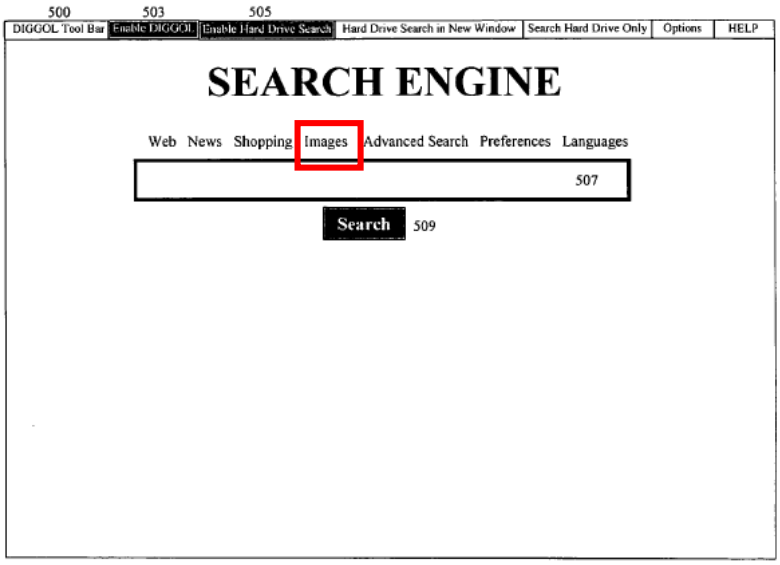
[5] Kirsch equates "document collections" with information (or data) sources. *See*, Ex. 1011, at 1:33-43 ("Private *document collections* are likely to continue to exist as significant *sources* of unique information.")(emphasis added).

the user query was then "used to *select* a *qualified set* of most likely relevant *document collections*." *Id*., 6:29-30 and 10:19-21 ("the present invention supports automatic ... collection selection.")

59. In Kirsch, the selection of a document collection from the set of collections was based on the document collection exceeding a relevance score and a rank. In order for a document collection to be selected, the collection needed to meet a "*minimum level of relevance*." *Id*., 12:11-12; 9:67-10:4 (discussing a "minimum relevance ranking threshold.") A further "post-search constraint" required a selected collection to *rank high enough to fall within a "maximum number of collections returned.*" *Id*., 12:11-12.

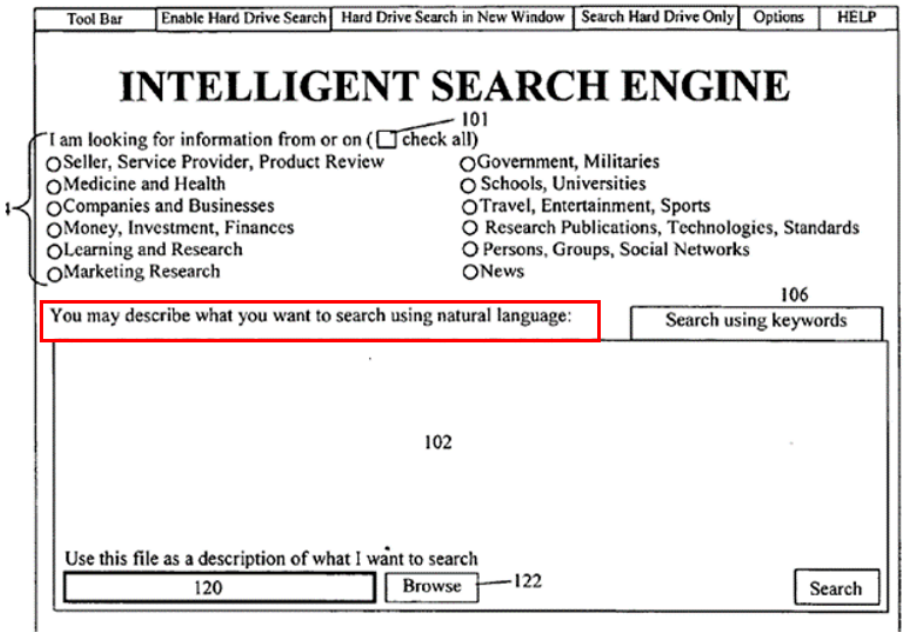
C. U.S. Patent Publication No. 2006/0106793 - Liang

60. U.S. Patent Publ. No. 2006/0106793 ("Liang") (Ex. 1012) entitled "Internet and Computer Information Retrieval And Mining With Intelligent Conceptual Filtering, Visualization And Automation," disclosed an "Intelligent Search Engine," which could be used for, *inter alia*, querying non-textual data such as images:



Ex. 1012, Figure 5 (red box added).

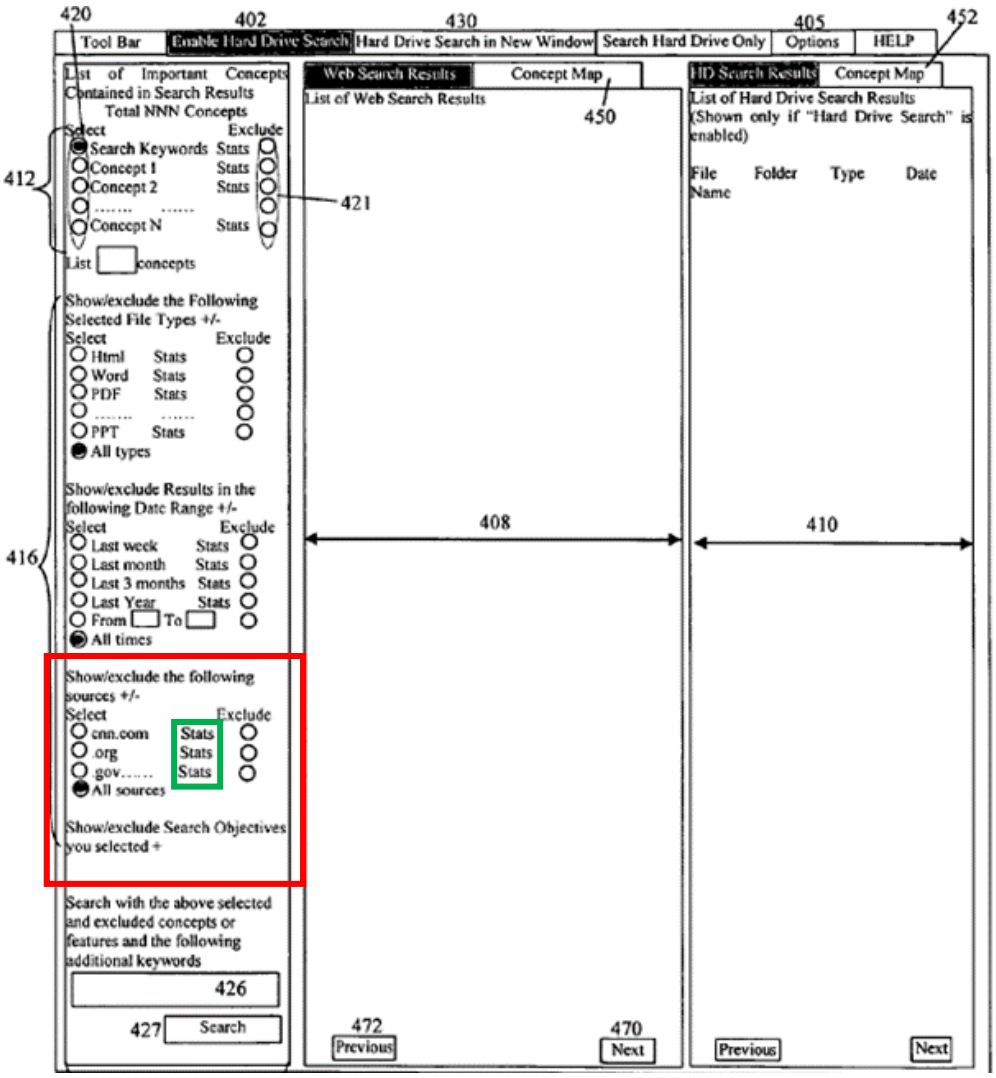
61. Figure 1 of Liang (shown below with red box added) illustrates an embodiment of Liang's user interface:



As shown above, Liang included a user interface through which a user could enter a natural language search query (red box). Liang also taught *expanding search*

keywords using *synonyms*.[6] Ex. 1012, [0013].

62. Liang taught execution of the query on multiple sources (see red box added below), as depicted in Figure 4 of the reference:



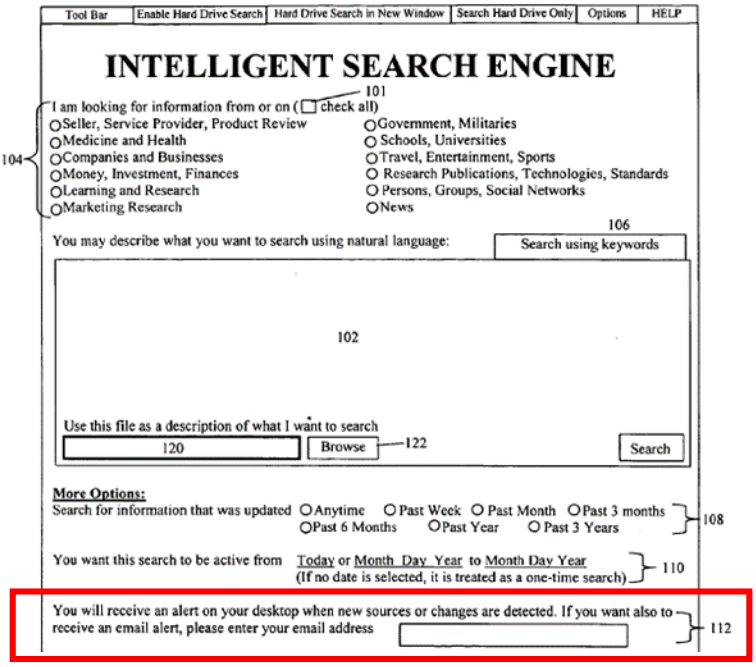
63. A POSITA would have understood Liang as teaching that a query is initially executed on a plurality of data sources. In describing Figure 4, Liang states

[6] Liang taught expanding search terms with "synsets." Ex. 1012, [0013]. A synset is a set of synonyms.

that "sections 416 allow a user to ***filter*** the search results by other filtering features such as file types, dates of modification, ***sources***, among other things." *Id*., [0178] (emphasis added). In order for a user to be able to filter based on source**s** (plural) in Liang, the query would first need to have been executed on a plurality of sources. The fact that the "All Source**s**" box is checked in Figure 4 further suggests that the initial query is applied to multiple data sources in Liang.

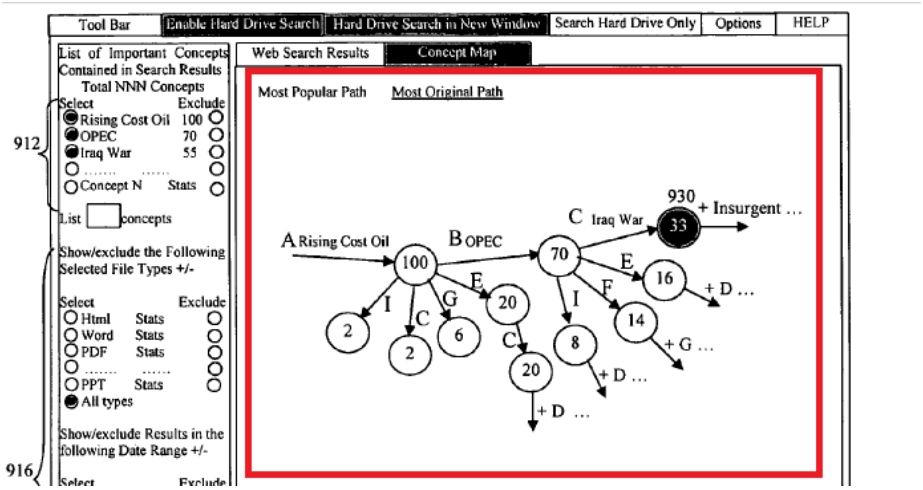
64. Liang also provides "Stats" (see green box in Figure 4 above) for each of the data sources that were searched. Liang explains that these "Stats" correspond to statistics, such as "the total number of appearances of an important concept in the search results." *Id*., [0184].

65. Liang provides users functionality to monitor data sources for updates using the query, and to receive an alert (e.g., by email) when changes are detected. This aspect is disclosed in Figure 1 of Liang (reproduced below with red box added):



66. With respect to Figure 1, Liang stated that "it offers a user ... the option to keep his search active for a period of time to *monitor for* new sources and *changes to existing sources* by specifying a date range in 110, and when a change is detected, the option to alert the user on his local PC or send an email to an email account that the user provides in 112." *Id*., [0052] (emphasis added). Liang also taught that "[b]oth the time period and the *monitoring frequency* may be chosen by a pull-down menu, or text box and check boxes." *Id*., [0347] (emphasis added).

67. Liang further taught techniques for visualization of search results, as shown for example in Figure 9 of the reference (shown in part below - red box added by Petitioner):



D. U.S. Patent No. 5,721,902 (Schultz)

68. U.S. Patent No. 5,721,902 to Schultz (Schultz) (Ex. 1013) taught a method for querying a database that included both textual data and non-textual multimedia data such as images. *Id*., 12:49-53 ("query server 116 scans its document index database 117 based on the user's search query in order to identify document identification numbers representing textual documents and multi-media files relevant to the search query"); *see also*, 12:66-13:4 ("[t]he term "document file" is used in the present application to mean a file in database 118 which stores either a textual document or, alternatively, a file which stores a multi-media record such as, for example, *a still image*, ...") (emphasis added).

69. As explained in connection with Figure 4A, in response to the query Schultz returned a results list to the user including "the relevance scores corresponding to each document identification number on the search results list" (see

item 342 in red boxes below) and "an indication whether the document file corresponds to a textual document or a multi-media file such as, for example, a file representing a still *image*" (see item 344 in green boxes below). *Id*., 14:29-37 (emphasis added).

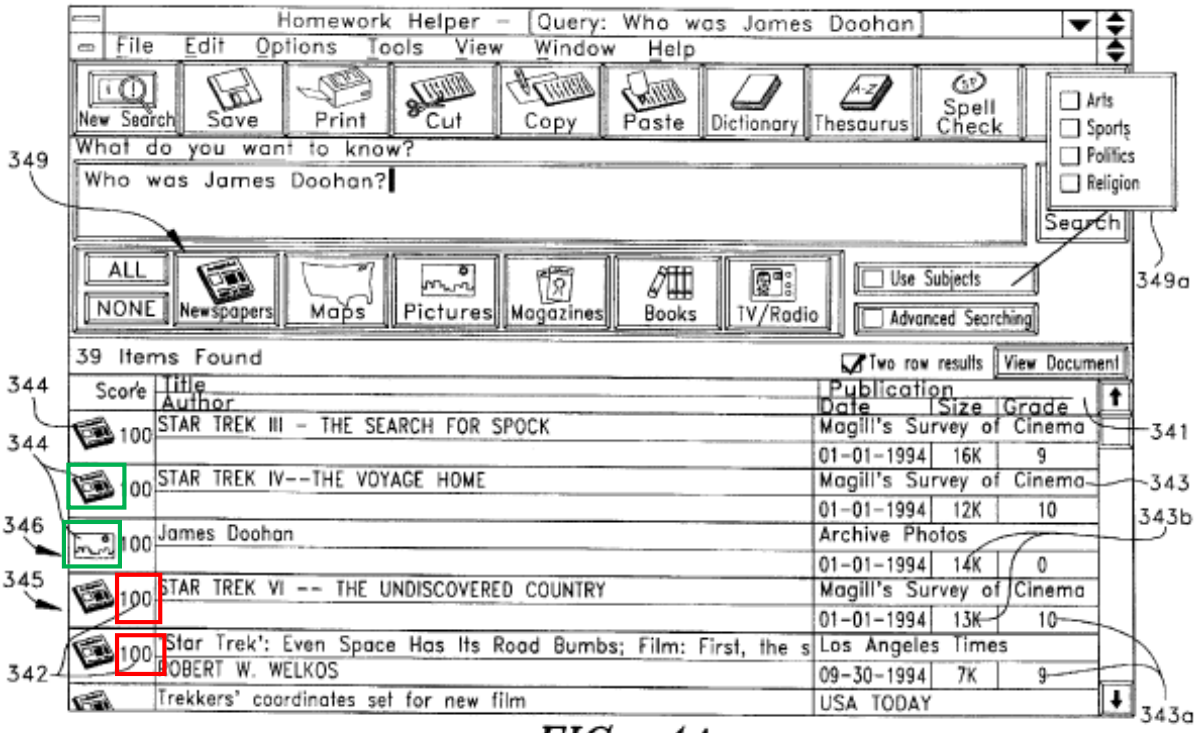
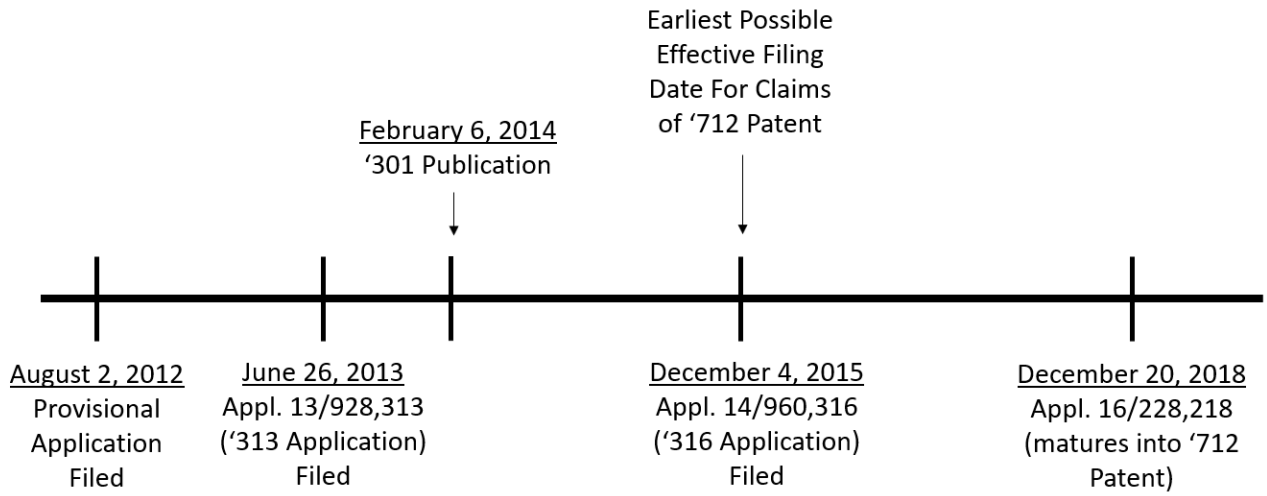


FIG. 4A

70. Schultz also taught that the results list can be sent from a server to a user over a communication channel, and the user can then provide feedback by clicking on a particular search result causing the server to return the corresponding document to the user over the *same* communication channel that the server used to send the search results list to the user. *Id*., 1017, 14:13-15:16.

E. U.S. Patent Publ. No. 2014/0040301 (the '301 Publication)

71. The '313 Application (discussed in Section VII(C), *supra*) was published on *February 6, 2014* as U.S. Patent Publ. No. 2014/0040301 ("the '301 Publication")(Ex. 1014). The timeline below depicts the sequence of applications to which the '712 Patent claims priority, along with the publication date of the '301 Publication.



72. As shown above, the publication date of the '301 Publication is more than one (1) year before the earliest possible effective filing date of the '712 Patent. I understand that the '301 Publication (which matured into the '977 Patent) is prior art to the '712 Patent.

73. As explained in Ground 13 below, the '301 Publication discloses *every* limitation in the Claims of the '712 Patent, other than the requirement that the querying be "for non-textual data."

F. U.S. Patent No. 7,765,228 (Eichstaedt)

74. U.S. Patent No. 7,765,228 to Eichstaedt et al. taught servers that "compare old query results" to "current content item(s)" after a "predefined period" and deliver an email alert to the user with the new content item. Ex. 1015, 8:19-20, 8:37-54 and Figures 5-6.

G. U.S. Patent No. 7,844,603 (Lucovsky)

75. U.S. Patent 7,844,603 to Lucovsky (Lucovsky) (Ex. 1016) is entitled "Sharing User Distributed Search Results." Lucovsky discloses mechanisms by which a first user could share search results with a second user via an email, a message board, an instant message and a blog. *Id*., at Figures 4-9.

H. WO Publication No. 01/65336 (Mellin)

76. WO Publication No. 01/65336 (Mellin) (Ex. 1017), taught the use of an email address directed to a group in order to facilitate dissemination of information to all members of the group. *See*, e.g., *Id.*, at 3:24-26 ("A group is a collection of users, known as group members. Each group has an email address with the system.")

I. U.S. Patent No. No. 7,043,482 (Vinsonneau)

77. U.S. Patent No. 7,043,482 (Vinsonneau) (Ex. 1018) taught a method for querying data banks where the number of bots used to process search queries depends on the volume or quantity of data being searched. *Id*., 4:8-9.

J. U.S. Patent No. 6,243,713 (Nelson)

77A. U.S. Patent No. 6,243,713 entitled "Multimedia Document Retrieval

By Application Of Multimedia Queries To A Unified Index Of Multimedia Data For A Plurality Of Multimedia Data Types" (Nelson) disclosed use of *a query that included an image* ("query image"). Ex. 1019, at 3:6-17. The query can include "query feature operators 154 such as "/shape' and "/color" [that] enable the user to selectively define which aspects of the query images are of interest. The "/shape' operator delimits that portion of the query to images having similar shapes as the sailboat image ... whereas the "/color" operator delimits that portion of the query to images having colors similar to the sunset image." *Id*., at 6:35-56. The query *can be expanded* to include "other images (or image attributes) of similar shape or color, texture ... ." *Id*., at 7:14-21.

XI. SUMMARY OF GROUNDS FOR UNPATENTABILITY

Ground	Claim(s)	References - §103
1	1, 2, 4-8, 10, 11, 13-17 and 19-20	Kirsch, Schultz, Eichstaedt and Lucovsky
2	1, 4-8, 10, 13-17 and 19	Kirsch, Schultz, Eichstaedt and Mellin
3	2, 11 and 20	Kirsch, Schultz, Eichstaedt and Mellin
4	3 and 12	Kirsch, Schultz, Eichstaedt and Lucovsky, in view of Liang
5	9 and 18	Kirsch, Schultz, Eichstaedt and Lucovsky, in view of Vinsonneau
6	9 and 18	Kirsch, Schultz, Eichstaedt and Mellin, in view of Vinsonneau
7	1-4, 7-8, 10-13, 16-17 and 19-20	Liang, Nelson and Lucovsky
8	1-8, 10-17 and 19-20	Liang, Nelson, Kirsch, Schultz and Lucovsky
9	1, 4, 7, 8, 10, 13, 16, 17 and 19	Liang, Nelson and Mellin

Ground	Claim(s)	References - §103
10	2, 3, 5, 6, 11, 12, 14, 15 and 20	Liang, Nelson, Mellin, Kirsch And Schultz
11	9 and 18	Liang, Nelson, Lucovsky and Vinsonneau
12	9 and 18	Liang, Nelson, Kirsch, Schultz and Lucovsky, in view of Vinsonneau
13	1-20	'301 Publication and Schultz

XII. ALL OF THE CHALLENGED CLAIMS ARE OBVIOUS

A. Ground 1 - Claims 1, 2, 4-8, 10, 11, 13-17 And 19-20 Are Obvious Over Kirsch, Schultz, Eichstaedt And Lucovsky

1. Claim 1

78. As set forth below, all of the elements of Claim 1 are met by Kirsch, Schultz, Eichstaedt and Lucovsky.

a. Element 1.1 - "A method of analyzing data"

79. To the extent the preamble is limiting, Kirsch teaches a method for searching databases, which maps to "[a] method of analyzing data." Ex. 1011, 7:47-49 (search terms are provided "for *analysis* against a collection meta-index 16.")(emphasis added). When Kirsch selects the "most likely relevant documents from the identified collection set" (*id*., 6:34-37), it is also "analyzing data."

b. Element 1.2 - "receiving, from a first user, a query for non-textual data"

80. Kirsch taught a user query for querying documents. Ex. 1011, 6:40-43. Schultz taught a user query for querying non-textual data such as images. Ex. 1013, 12:49-53 and 12:66-13:4. Kirsch and Schultz both relate to search technology, and it would have been obvious to incorporate Schultz's techniques for querying non-textual image data into Kirsch. A POSITA would have been motivated to combine the references in order to apply Kirsch's collection selection techniques to collections that included image data. A POSITA would have had a reasonable expectation of

success in applying Kirsch's collection selection techniques to collections that included image data as taught in Schultz, as the combination involves nothing more than the routine combination of existing software functionality.

c. Element 1.3 - "expanding search terms of the query"

81. Kirsch teaches that "search terms are thus ... expanded as appropriate into a list of specifically identified search terms." Ex. 1011, 7:6-7. Schultz teaches determining "expansion words" for a word in the query. Ex. 1013, 6:30-32. Thus, Kirsch and Schultz both teach this claim element.

d. Element 1.4 - "executing the query on a plurality of data sources"

82. Kirsch teaches searching a plurality of different "document collections," each of which is a "data source." Kirsch equates each document collection with a "source of unique information" which corresponds to a "data source" (see note 1, *supra*):

> While access to much of the information available through the Internet is free for public use, numerous proprietary or fee-based access *document collections exist*. Although such private document collections may all be accessible through the Internet, ... the collections are generally forced to be accessible as discrete entities in order to maintain fee-based access control. ... . Private document collections are likely to continue to exist as significant *sources of* unique *information*. Ex. 1011, 1:33-43 (emphasis added).

83. Kirsch similarly equates each document collection with a "database" which also corresponds to a "data source" (see note 1, *supra*):

> A known problem in providing access to multiple *databases* is the relative difficulty or inefficiency in identifying an optimal database or set of databases that should be searched to obtain the best search report for a some particular unstructured, or ad hoc, database query. ... The user must manually determine and *select a particular set of databases* that must, by definition, contain the desired results to a query. ... " *Id*., 1:64 -2:20 (emphasis added).

84. Kirsch teaches that the query is initially directed against multiple document collections (which map to "a plurality of data sources" in the claim.) *Id.*, at 14:48-50 ("the user query 84 is preconditioned upon issuance from the user workstation 82 to *specify the particular set of collections 88, 90, 92 that the query is to be resolved against*").

85. Schultz taught applying a user query for non-textual data to a collection that included text and images. Ex. 1013, 12:49-53, 12:66-13:4 and 13:20-24 ("the scan of document index database 117 by query server 116 in step 335 simultaneously identifies document identification numbers representing both textual documents *and multi-media* records relevant to the search query.")

86. It would have been obvious to incorporate Schultz's techniques for querying a collection having image data into Kirsch's system for querying a plurality of collections. A POSITA would have been motivated to combine the references in order to apply Kirsch's collection selection techniques to collections that included image data. A POSITA would have had a reasonable expectation of success in incorporating Schultz's techniques for querying a collection having image data into

Kirsch's system for querying a plurality of collections, as the combination involves nothing more than the routine combination of existing software functionality.

e. Element 1.5 - "selecting at least one data source from the plurality of data sources, the at least one data source being selected when results of the query are greater than or equal to an accuracy thresholds"

87. Kirsch teaches that "a user query 12 is used to *select* a *qualified set* of most likely relevant *document collections*." Ex. 1011, 6:29-30 and 10:19-21 ("the present invention supports automatic ... collection selection.") The selection of "a qualified set" of document collections in Kirsch maps to "selecting at least one data source from the plurality of data sources" in the claims. Kirsch teaches that in order for a document collection to be selected, the document collection must meet a "*minimum level of relevance*." *Id*., 2:12; 9:67-10:4 (discussing "*minimum relevance ranking threshold*.") A further "post-search constraint" required a selected collection to *rank high enough to fall within a "maximum number of collections returned.*" *Id*., 12:11-12.

88. Schultz taught relevance scoring and relevance ranking for image objects within a collection. Ex. 1013, 14:29-41, Figure 4A ("[t]his information is displayed ... in relevance ranked order.")

89. It would have been obvious to apply the relevance scoring and relevance ranking techniques for images taught in Schultz to Kirsch's system in order to select collection(s) having at least one image with a relevance score over a

minimum level of relevance and/or a relevance rank over a minimum ranking threshold. In this context, Schultz's relevance scoring and relevance ranking techniques are analogous to Kirsch's “minimum level of relevance” and “minimum relevancy ranking threshold,” and it would have been a matter of simple substitution to use Schultz's metrics for relevance scoring and/or relevance ranking of images for collection selection within Kirsch. A POSITA would have had a reasonable expectation of success in applying the relevance scoring and relevance ranking techniques for images taught in Schultz to Kirsch's system in order to select collection(s) having at least one image with a relevance score over a minimum level of relevance and/or a relevance rank over a minimum ranking threshold, as the combination involves nothing more than the routine combination of existing software functionality.

90. As explained in Section VI(B), *supra*, based on the specification and Patent Owner's infringement contentions, it is clear that in the context of the '712 Patent, a minimum relevancy threshold and/or a minimum relevancy ranking is an example of “an accuracy thresholds.” Thus, the combination of Kirsch and Schultz teach "selecting at least one data source from the plurality of data sources ... when results of the query are greater than or equal to an accuracy thresholds."

f. Element 1.6 - "monitoring, based on a set schedule, the at least one data source to extract non-textual data from the at least one data source when at least an update to stored data that matches the query, newly added non-textual data that matches the query, or a combination thereof"

91. Eichsteadt teaches providing an email notification (or alert) to a user "when new information becomes available regarding ... search results." Ex. 1015, 2:5-15. Eichsteadt teaches that "queries can be applied to new content information ... as the new content information becomes available." *Id*., 2:43-45. In order to prepare alerts, a poller server "wakes up after a predefined period" (e.g., after each "15 minute" slot elapses) and compares "old query results to the current content item(s). For example, if a current content item is different from an old query result, the poller server can replace the old query result with the current content item. ... The poller server also sends a request to the delivery server" which delivers "an alert with the content item." *Id*., 7:67, 8: 14-21, 8:28-54, 12:38-41 and Figures 5-6.

92. Eichsteadt's disclosures of (i) applying an old query to new content information after a predefined period (e.g., after expiration of 15 minute slots) and (ii) delivering an alert to a user with the new content item, map to "monitoring, based on a set schedule, the at least one data source to extract data when at least an update to stored data that matches the query," in the claim.

93. Kirsch, Schultz and Eichsteadt are all directed to search query technology. It would have been obvious to a POSITA to incorporate Eichsteadt's

monitoring functionality into the combination of Kirsch and Schultz described above, so that a user could receive periodic updates of new images from the "selected document collections." A POSITA would have had a reasonable expectation of success in incorporating Eichsteadt's monitoring functionality into the combination of Kirsch and Schultz described above, so that a user could receive periodic updates of new images from the "selected document collections," as the combination involves nothing more than the routine combination of existing software functionality.

g. Element 1.7 - "establishing a communication channel with a second user based at least in part on the data extracted from the at least one monitored data source"

94. As explained in Section VI, *supra*, the '712 Patent provides examples of "establishing a communication channel" that include sending a message via email to user(s). Ex. 1001, 14:56-57. Thus, sending an email alert with updated query results is an example of "establishing a communication channel ... based at least in part on the data extracted from the at least one monitored data source." In addition, a POSITA would have understood the claim language as being broad enough to include establishing a communication with a second user *via the first user*.[7]

[7] During prosecution of the '977 Patent, the Applicant made clear that the claimed establishing of the communication channel with the second user could be performed "via the first user." *See*, Ex. 1003, at 138, 142-143. While the Applicant later broadened the claim by deleting the *requirement* that the communication channel be

95. Lucovsky (Ex. 1016) discloses a first user sharing search results with a second user via an email (Figures 4-6), message board (Figure 7), instant messaging (Figure 8) and a blog (Figure 9). Kirsch, Schultz, Eichsteadt and Lucovsky are all directed to search query technology. It would have been obvious to incorporate Lucovsky's search results sharing techniques with Kirsch, Schultz and Eichsteadt, so that a first user could share updates with a second user. A POSITA would have had a reasonable expectation of success in incorporating Lucovsky's search results sharing techniques with Kirsch, Schultz and Eichsteadt, so that a first user could share updates with a second user, as the combination involves nothing more than the routine combination of existing software functionality. Moreover, to the extent element this requires that the communication channel be established with a different user without traversing the user submitting the query, the limitation is taught by Eichsteadt. Eichsteadt discloses that when new content matching a query is identified, a communication that includes the update is sent not only to the user that generated the query, but also to *other users* "who desire the same, or very similiar, query results." Ex. 1015, 4:13-51. Sending an update to any one of the "other users" maps to "establishing a communication channel with a second user."

96. Accordingly, Claim 1 is obvious over Kirsch, Eichsteadt and Lucovsky.

---

established with the second user "via the first user" before allowance, the prosecution history makes clear that establishing the communication channel with the second user *via the first user* is *within the scope of the claim*.

2. Claims 10 and 19

97. Claims 10 and 19 recite all of the limitations of Claim 1 discussed above, but in the format of an apparatus or computer program product rather than a method. Claim 10 further recites "a memory unit" and "at least one processor ... configured" to perform the recited functions, and Claim 19 further recites "a non-transitory computer-readable medium having program code recorded thereon" to perform the recites functions. These aspects of Claims 10 and 19 are met by Kirsch. Ex. 1011, 6:24-29 ("The methodology 10 is preferably implemented as an application program ... executing on one or more conventional networked computer systems.") A POSITA would have understood a conventional network computer system to include memory coupled to a processor, as claimed.

98. Accordingly, Claims 10 and 19 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

3. Claims 2, 11 and 20

99. Claims 2, 11 and 20 are cast in different formats but otherwise contain substantively identical limitations. The analysis below of Claim 2 applies identically to Claims 11 and 20.

a. Element 2.1 - “analyzing the extracted non-textual data”

100. Kirsch teaches that the document list returned from the selected collection(s) may be refined “based on preferences established from the user

workstation 82. These preferences may define refinement criteria such as minimal acceptable relevancy score, total number of documents identified, maximum number of documents from a particular collection, and the like. The *refined document list* is then provided by the search interface 122 through the Internet 86 as a report 121 to the user workstation 82. Thus, the end user can *analyze* the report 121 and selectively retrieve identified documents from the various collections 88, 90, 92 as desired." Ex. 1011, 15:50-59 (emphasis added). When Kirsch's system refines the document list based on user preferences, it is "analyzing" data. Similarly, Kirsch states that the user analyzes the document list, which also maps to "analyzing" data.

101. As noted above, Eichsteadt teaches extracting data as part of the process of monitoring for updates. It would have been obvious to apply Kirsch's teachings for refining and/or analyzing a document list (as set forth in the paragraph above) to Eichsteadt so that user preferences could be used to refine (or narrow) the subject matter of alerts. A POSITA would have had a reasonable expectation of success in applying Kirsch's teachings for refining and/or analyzing a document list to Eichsteadt so that user preferences could be used to refine (or narrow) the subject matter of alerts, as the combination involves nothing more than the routine combination of existing software functionality.

b. Element 2.2 - "receiving feedback from the user based at least in part on the established communication channel"

102. As mentioned above in Section XII(A)(1)(f), Lucovsky discloses a first user sharing search results with a second user via an email. An example of such an email is depicted in Figure 5 of the reference (shown below):

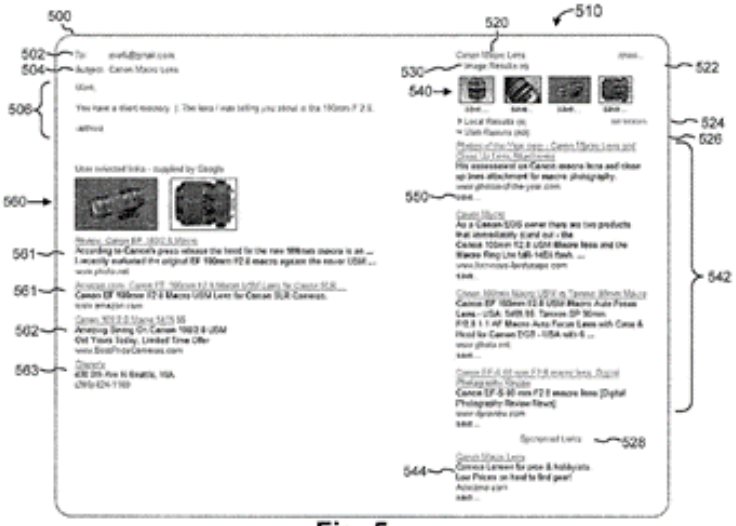
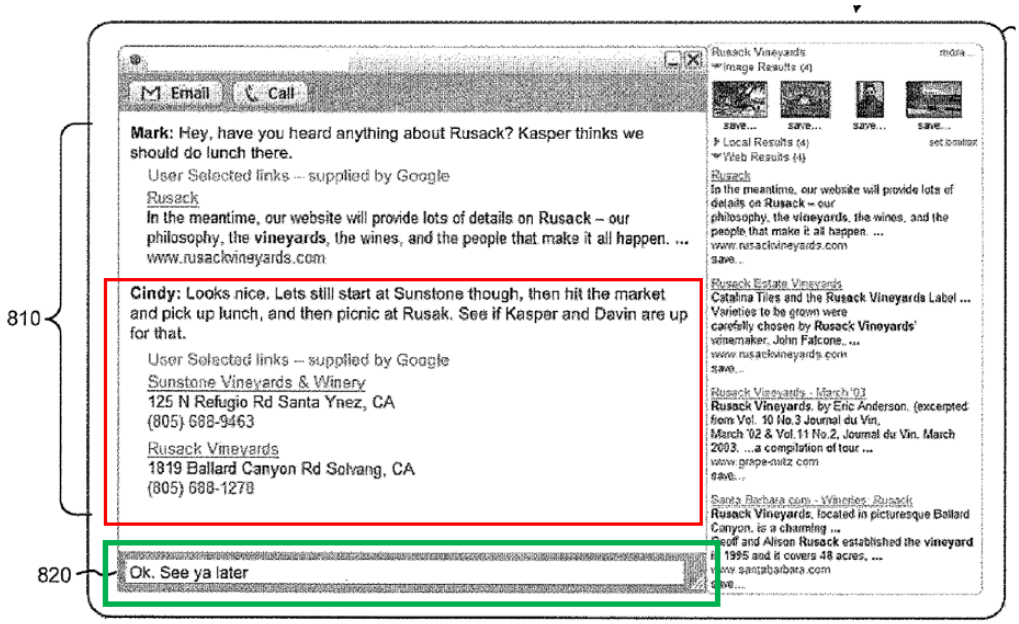


Fig. 5

103. As shown in Figure 8 of the reference (shown below with annotations added), Lucovsky discloses a user interface through which one user (Cindy) sends an email with search results (red box below) to a second user (Mark). Lucovsky discloses that the second user (Mark) *responds* to the first user's email *with the search results* with the message "OK, See ya later" (green box below).



The responsive message (green box above) maps to "receiving feedback from the user based at least in part on the established communication channel" in the claim.

104. In view of the above, Claims 2, 11 and 20 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

4. Claims 4 and 13

105. Claims 4 and 13 are cast in different formats but otherwise contain substantively identical limitations. The analysis below of Claim 4 applies identically to Claim 13.

a. Element 4.1 - "expanding the query comprises expanding the search terms to include at least misspellings, synonyms, sub-topics, antonyms, or a combination thereof"

106. Kirsch teaches that "[c]ommon variably represented terms, such as online, are also equivalently recognized; online being made equivalent to "on line" and "on-line" for purposes of term equivalence mapping. ... The various search terms are thus reduced and expanded as appropriate into a list of specifically identified search terms." Ex. 1011, 6:64-7:7. Expanding the query to include "equivalently recognized" search terms in Kirsch maps to "expanding the search terms to include ... synonyms."

107. Schultz teaches expansion of a query term using a semantic network. The term "is applied to a semantic network in step 142d to determine *expansion words* corresponding to the query term. For purposes of the present invention, the term "semantic network" means *a database in which each term carries with it semantic information such as, for example*, part of speech and *synonym information*." Ex. 1013, 23:42-48. Thus, Schultz also taught expanding a search terms to include synonyms.

108. Thus, Claims 4 and 13 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

5. Claims 5 and 14

109. Claims 5 and 14 are cast in different formats but otherwise contain substantively identical limitations. The analysis below of Claim 5 applies identically to Claim 14.

a. Element 5.1 - "refining the search terms based at least in part on the executed query, and in which the monitoring is performed to monitor for matches to a refined query "

110. Lucovsky teaches that the "user may continue to edit the content, enter or *refine search queries*." Ex. 1016, 7:49-50. (emphasis added). Schultz also taught functionality for performing a recursive query, in which search terms are refined based on query results. Ex. 1013, 22:41-54 and Figure 6A. It would have been obvious to apply a refined (narrowed) query as taught in Lucovsky and Schultz to the monitoring taught in Eichsteadt in order to narrow the updates to those reflected by a refined query. A POSITA would have had a reasonable expectation of success in applying a refined (narrowed) query as taught in Lucovsky and Schultz to the monitoring taught in Eichsteadt in order to narrow the updates to those reflected by a refined query, as the combination involves nothing more than the routine combination of existing software functionality.

111. Thus, Claims 5 and 14 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

6. Claims 6 and 15

112. Claims 6 and 15 are cast in different formats but otherwise contain substantively identical limitations. The analysis below of Claim 6 applies identically to Claim 15.

a. Element 6.1 - " the plurality of data sources include open-source data sources that are publicly available and closed-source data sources that at not publicly accessible"

113. Kirsch applies its teachings to document collections that include both (i) "information available through the Internet [that] is free for public use" (which maps to "open-source data sources that are publicly available" in the claim), and (ii) "proprietary or fee-based access document collections" (which map to "closed-source data sources that are not publicly accessible" in the claim.) Ex. 1011, 1:33-42.

114. Thus, Claims 6 and 15 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

7. Claims 7 and 16

115. Claims 7 and 16 are cast in different formats but otherwise contain identical limitations. The analysis below of Claim 7 applies identically to Claim 16.

a. Element 7.1 - "the query comprises at least a natural language query, a non-textual data query, or a combination thereof"

116. Schultz teaches that the "search query is preferably in a natural language format." Ex. 1013, 4:8-9. Kirsch, Schultz, Eichsteadt and Lucovsky are all directed to search query technology. It would have been obvious to use Schultz's natural language search queries in Kirsch in order to make Kirsch more user-friendly for users who prefer not to formulate a Boolean search. A POSITA would have had

a reasonable expectation of success in incorporating Schultz's natural language search queries in Kirsch, as the combination involves nothing more than the routine combination of existing software functionality.

117. Thus, Claims 7 and 16 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

8. Claims 8 and 17

118. Claims 8 and 17 are cast in different formats but otherwise contain substantively identical limitations. The analysis below of Claim 8 applies identically to Claim 17.

a. Element 8.1 - "in which the non-textual data comprises at least one of an image, audio, a video, or a combination thereof"

119. In Schultz, the non-textual data include images. Ex. 1013, 12:49-53 and 12:66-13:4.

120. Thus, Claims 8 and 17 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

B. Ground 2 - Claims 1, 4-8, 10, 13-17 And 19 Are Obvious Over Kirsch, Schultz, Eichsteadt And Mellin

1. Claim 1

121. As explained in Sections XII(A)(1)(a)-(f), *supra*, Elements 1.1 - 1.6 of Claim 1 are met by Kirsch and Eichsteadt. Instead of relying on Lucovsky for Element 1.7 as set forth in Ground 1 above, Ground 2 maps Element 1.7

("establishing a communication channel with a second user based at least in part on the data extracted from the at least one monitored data source") to Mellin which taught the use of a group email address to facilitate collaboration among members of a group. *See*, Mellin (Ex. 1017), at 3:24-26. In a case where the author of the query (mapped to "a first user" in the claim) was part of a group collaboration, it would have been obvious for the author to designate a group email address to receive search result updates so that the updated search results were distributed to all collaborators. Any collaborator on the group email address other the author of the search query maps to "a second user" in the claim.

122. Accordingly, Claim 1 is obvious over Kirsch, Schultz, Eichsteadt and Mellin.

2. Claims 4-8, 10, 13-17 And 19

123. As set forth in Sections XII(A)(2), the additional limitations of Claims 10 and 19 are met by Kirsch.

124. As set forth in Section XII(A)(4), the additional limitations of Claims 4 and 13 are met by Kirsch.

125. As set forth in Section XII(A)(5), the additional limitations of Claims 5 and 14 are met by Schultz.

126. As set forth in Section XII(A)(6), the additional limitations of Claims 6 and 15 are met by Kirsch.

127. As set forth in Section XII(A)(7), the additional limitations of Claims 7 and 16 are met by Schultz.

128. As set forth in Section XII(A)(8), the additional limitations of Claims 8 and 17 are met by Kirsch.

129. Accordingly, Claims 4-8, 10, 13-17 and 19 are also obvious over Kirsch, Schultz, Eichsteadt and Mellin.

C. Ground 3 - Claims 2, 11 and 20 Are Obvious Over Kirsch, Schultz, Eichsteadt And Mellin

130. As noted above, Claims 2, 11 and 20 contain substantively identical limitations. The analysis below of Claim 2 applies identically to Claims 11 and 20.

1. Element 2.1 - "analyzing the extracted data"

131. This feature is taught by, or obvious in view of, Kirsch and Eichsteadt as set forth in Section XII(A)(3)(a) above.

2. Element 2.2 - "receiving feedback from the user based at least in part on the established communication channel"

132. As explained above in Section XII(B), it would have been obvious in view of Mellin (Ex. 1017) for an author of a query to designate a group email address to receive search result updates so that the updated search results were distributed to all collaborators. Any collaborator on the group email address other the author of the search query maps to "a second user" in the claim.

133. Schultz (Ex. 1013) teaches that a results list can be sent from a server to a user computer. The channel used to send the results to the user maps to "the

established communication channel" in the claim. Schultz teaches that, using the search results displayed on the user's computer, the user can then click on a particular search result and the server will return the corresponding document to the user over the *same communication channel*. In this scenario, when the user (e.g., a collaborator in Mellin's group email address mapped to the "second user") clicks on a particular result, the server receives "feedback" (in the form of the selected search result) from the user. This "feedback" is sent back to the server using the *same* communication channel that the server used to send the search results list to the user. By sending the feedback on the same communication channel as was used to send the search results list ("the established communication channel"), the feedback is received based in part on the established communication channel. Ex. 1013, 14:13-15:16.

134. It would have been obvious to use Schultz's techniques for communicating with the server when the server sends a group email (as taught in Mellin) to a group of collaborators, so as to allow a given collaborator to retrieve a particular result on a search results list. A POSITA would have been motivated to combine these techniques in order to simplify retrieval of search results for users. A POSITA would have had a reasonable expectation of success in combining these techniques, as the combination involves nothing more than the routine combination of existing software functionality.

135. Thus, Claims 2, 11 and 20 are also obvious over Kirsch, Schultz, Eichsteadt and Mellin.

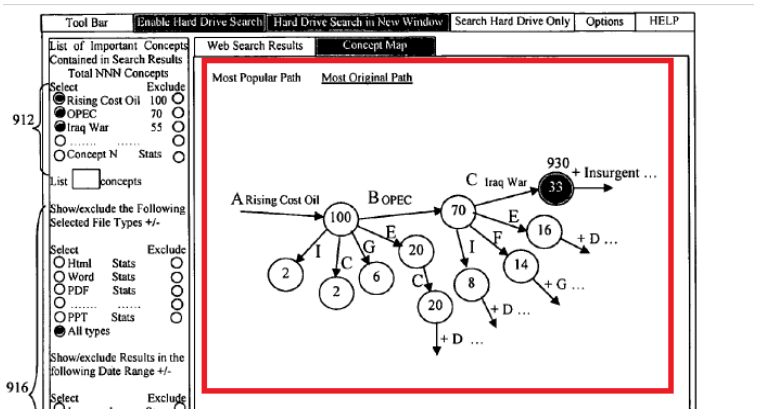
D. Ground 4 - Claims 3 And 12 Are Obvious Over Kirsch, Schultz, Eichsteadt And Lucovsky, In View Of Liang

136. Claims 3 and 12 contain substantively identical limitations. The analysis below of Claim 3 applies identically to Claim 12.

137. Claim 3 depends from Claim 2. As set forth in Section XII(A)(3), *supra*, Claim 2 is obvious over Kirsch, Schultz, Eichsteadt And Lucovsky.

1. Element 3.1 - "visualizing at least the extracted data, the analyzed data, or a combination thereof"

138. Liang (Ex. 1012) teaches visualization of search results using graphical diagrams (see Figure 9 of Liang below - red box added).



139. Kirsch, Schultz, Eichsteadt, Lucovsky and Liang are all directed to search query technology. It would have been obvious to incorporate Liang's visualization techniques with the other references so that a user could visualize search results. A POSITA would have had a reasonable expectation of success in

incorporating Liang's visualization techniques with the other references so that a user could visualize search results, as the combination involves nothing more than the routine combination of existing software functionality.

140. Thus, Claims 3 and 12 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky, in view of Liang.

E. Ground 5 - Claims 9 and 18 Obvious Over Kirsch, Schultz, Eichsteadt And Lucovsky, And Further In View Of Vinsonneau

141. Claims 9 and 18 are contain substantively identical limitations. The analysis below of Claim 9 applies identically to Claim 18.

142. Claim 9 depends from Claim 1. As set forth in Section XII(A)(1), *supra*, Claim 1 is obvious over Kirsch, Schultz, Eichsteadt and Lucovsky.

1. Element 9.1 - "the monitoring comprises deploying extractors to each of specific data source, a number of extractors being dynamic based at least in part on a traffic volume of each specific data source"

143. Vinsonneau (Ex. 1018) taught a method for querying data banks where the number of bots used to process search queries depends on the volume or quantity of data being searched. *Id*., 4:8-9. Vinsonneau and Eichsteadt are both directed to search technology. In view of Vinsonneau, it would have been obvious to dynamically scale the number of extractors (or "poller servers") in Eichsteadt for a particular data source based on traffic volume in order to improve the performance of Eichsteadt as traffic volume varied. A POSITA would have had a reasonable

expectation of success in dynamically scaling the number of extractors (or "poller servers") in Eichsteadt for a particular data source based on traffic volume as taught in Vinsonneau, as the combination involves nothing more than the routine combination of existing software functionality.

144. Thus, Claims 9 and 18 are obvious over Kirsch, Schultz, Eichsteadt and Lucovsky, in view of Vinsonneau.

F. Ground 6 - Claims 9 and 18 Obvious Over Kirsch, Schultz, Eichsteadt And Mellin, And Further In View Of Vinsonneau

145. Claims 9 and 18 are contain substantively identical limitations. The analysis below of Claim 9 applies identically to Claim 18.

146. Claim 9 depends from Claim 1. As set forth in Section XII(B), *supra*, Claim 1 is obvious over Kirsch, Schultz, Eichsteadt and Mellin. With respect to the additional limitation of Claim 9, it would have been obvious in view of Vinsonneau to dynamically scale the number of extractors (or "poller servers") in Eichsteadt for a particular data source based on traffic volume, in order to improve the performance of Eichsteadt as traffic volume varied. A POSITA would have had a reasonable expectation of success in dynamically scaling the number of extractors (or "poller servers") in Eichsteadt for a particular data source based on traffic volume as taught in Vinsonneau, as the combination involves nothing more than the routine combination of existing software functionality.

147. Thus, Claims 9 and 18 are also obvious over Kirsch, Schultz, Eichsteadt and Mellin, in view of Vinsonneau.

G. Ground 7 - Claims 1-4, 7-8, 10-13, 16-17 And 19-20 Are Obvious Over Liang, Nelson And Lucovsky

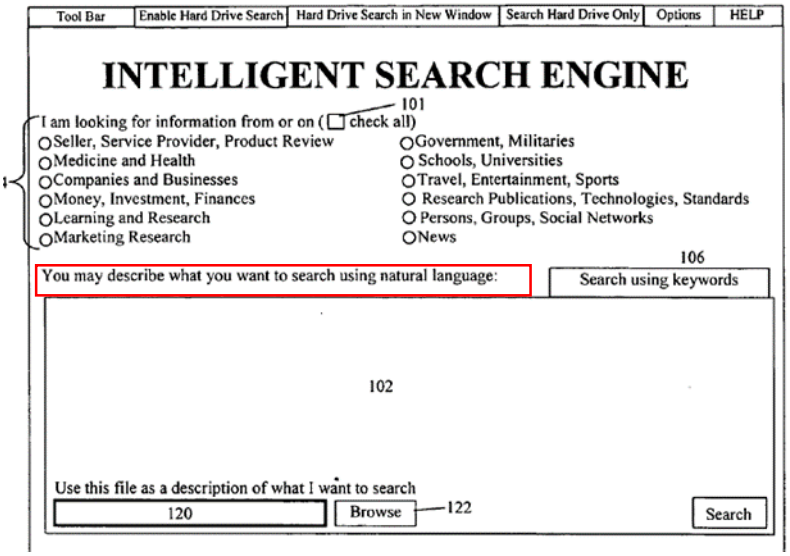
1. Claim 1

a. Element 1.1 - "A method of analyzing data"

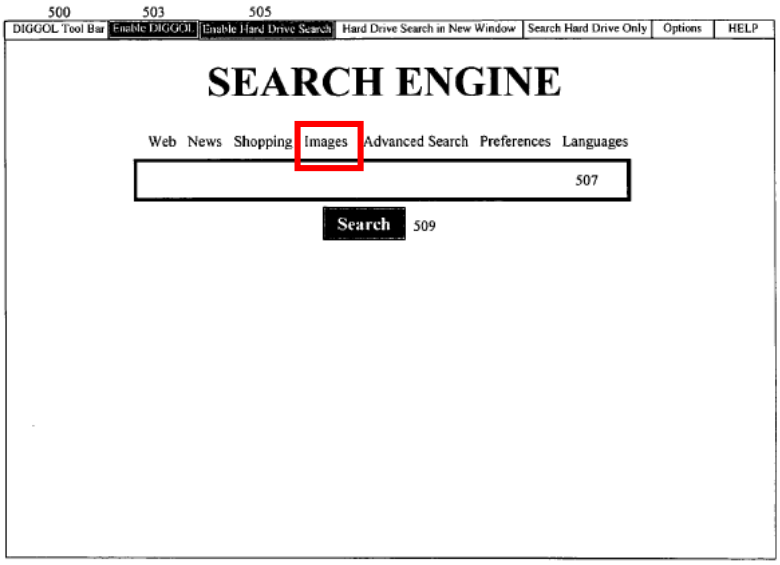
148. To the extent the preamble is limiting, Liang teaches a method of analyzing data. *See*, Ex. 1012, at [0167] ("this invention uses the natural language understanding module 202, the search keyword string extraction module 204 and the search keyword string generation module 206 to *analyze* search results ...")

b. Element 1.2 - "receiving, from a first user, a query for non-textual data"

149. Figure 1 of Liang teaches an "Intelligent Search Engine" that a user can search using a natural language query:



Id., Fig. 1 (red box added). Liang also taught that its search engine which could be used for, *inter alia*, querying non-textual data such as images:



Ex. 1012, Figure 5 (red box added). Thus, Liang teaches "receiving, from a first user, a query for non-textual data."

150. To the extent this claim element requires that the query itself include non-textual data, such a query is taught by Nelson. Ex. 1019, at 3:6-17 (teaching that the system retrieves "an image with image characteristics (e.g. color, texture, luminance) which match the input query image") and 6:35-56. Nelson further taught "a compound query" that would require that the word 'beach' appear within 10 words of an image of a sunset." *Id*., at 3:11-12. Nelson and Liang are both related to search technology. It would have been obvious to incorporate Nelson's techniques for using query images and compound queries into Liang's retrieval system. A POSITA would have been motivated to combine the references in order to expand

the capabilities of Liang such that a user can include both text and a query image in a query. A POSITA would have had a reasonable expectation of success with respect to this combination, as the combination involves nothing more than the routine combination of existing software functionality.

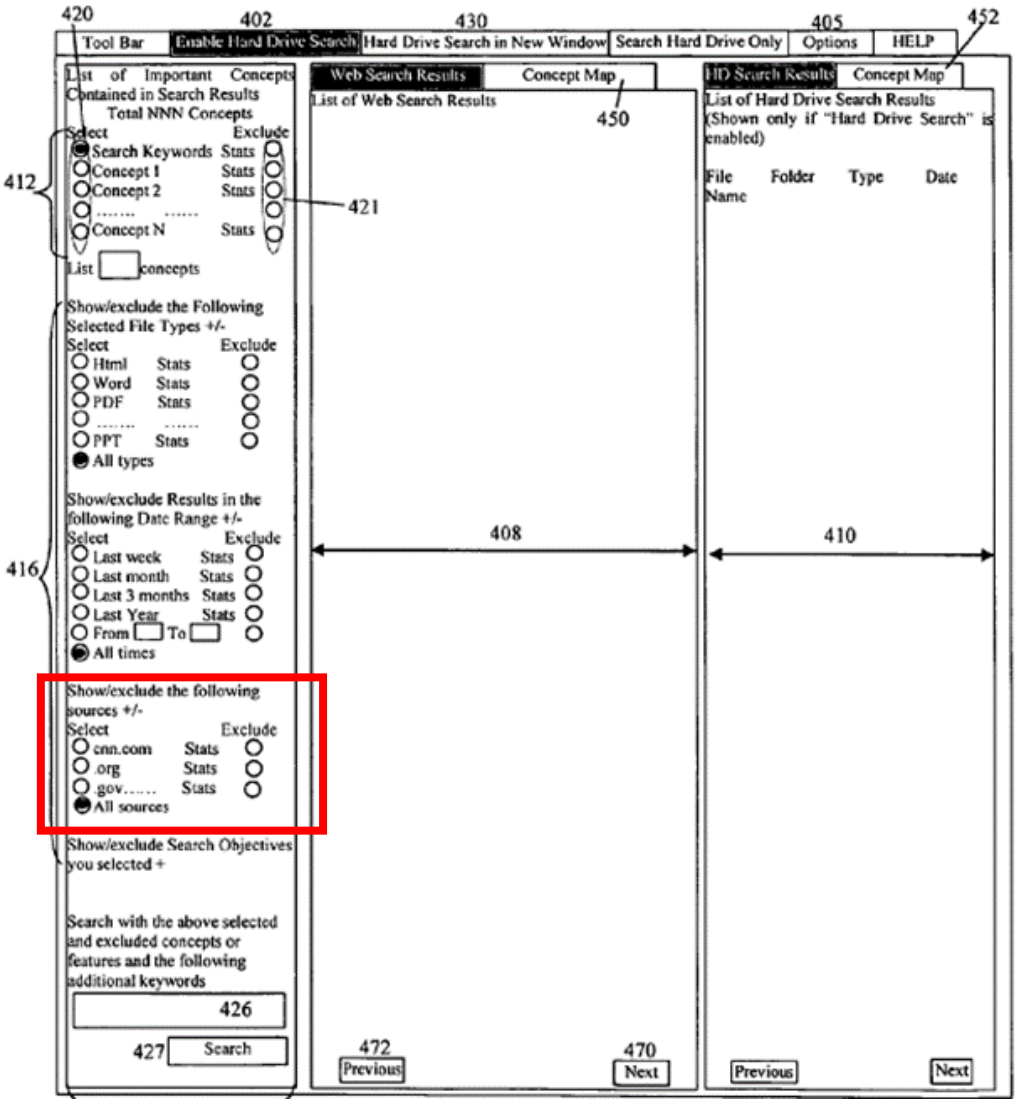
c. Element 1.3 - "expanding search terms of the query"

151. Liang "broadens a search by *expanding search keywords* into concepts comprising of the synsets, hypemym, and/or hyponym/troponym of a keyword ...," which maps to "expanding search terms of the query" in the claim. Ex. 1012, [0013].

152. To the extent the claim requires that the query itself include non-textual data, such a query is taught by Nelson, which would have been obvious to combine with Liang, as explained above. *See*, Section (X)(G)(1)(b), *supra*. Nelson further taught that its query can be expanded to include "other images (or image attributes) of similar shape or color, texture, and so forth." *Id*., at 7:14-21. It would have been obvious to further incorporate Nelson's techniques for expanding the query image into Liang's retrieval system. A POSITA would have been motivated to combine Liang's query expansion techniques with those in Nelson in order to increase the robustness of search results from a "compound query." *See*, Section (X)(G)(1)(b), *supra*. A POSITA would have had a reasonable expectation of success with respect to this combination, as the combination involves nothing more than the routine combination of existing software functionality.

d. Element 1.4 - "executing the query on a plurality of data sources"

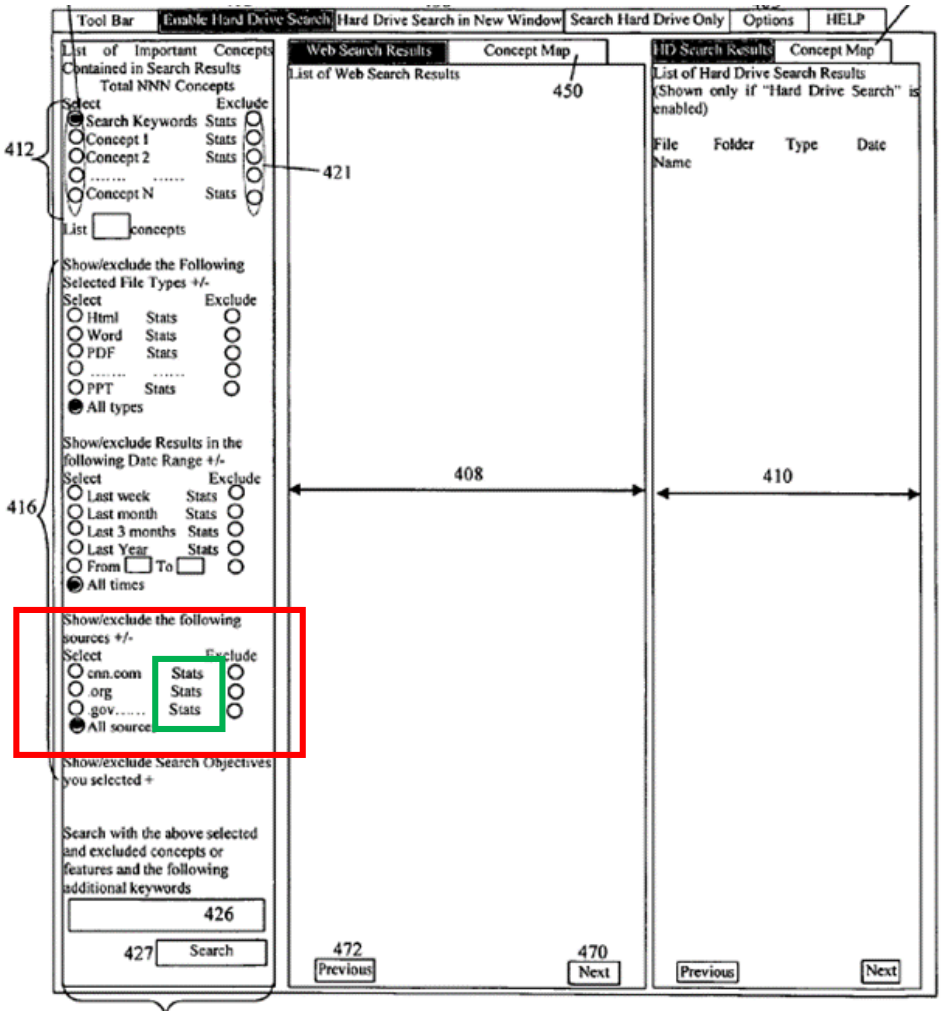
153. Liang teaches execution of a query on multiple sources (see red box below). Liang suggests initially executing the query on a plurality of data sources, after which the user is provided with the option to narrow the number of data sources used for the query.



Ex. 1012, Figure 4 (annotations added). In describing Figure 4, Liang states that the "sections 416 allow a user to *filter* the search results by other filtering features such as file types, dates of modification, *sources*, among other things." Ex. 1012, [0178] (emphasis added). In order for a user to be able to filter based on sources (plural) in Liang, the query would first need to have been executed on a plurality of sources. The fact that the "All Sources" box is checked in Figure 4 further suggests that the initial query is applied to multiple data sources in Liang. In any event, it would have been obvious for a user to select multiple sources for an initial query in Liang, so as to cover a range of sources with potentially relevant data.

e. Element 1.5 - "selecting at least one data source from the plurality of data sources, the at least one data source being selected when results of the query are greater than or equal to an accuracy thresholds"

154. Liang provides "Stats" (green box) for each of the data sources that were searched:

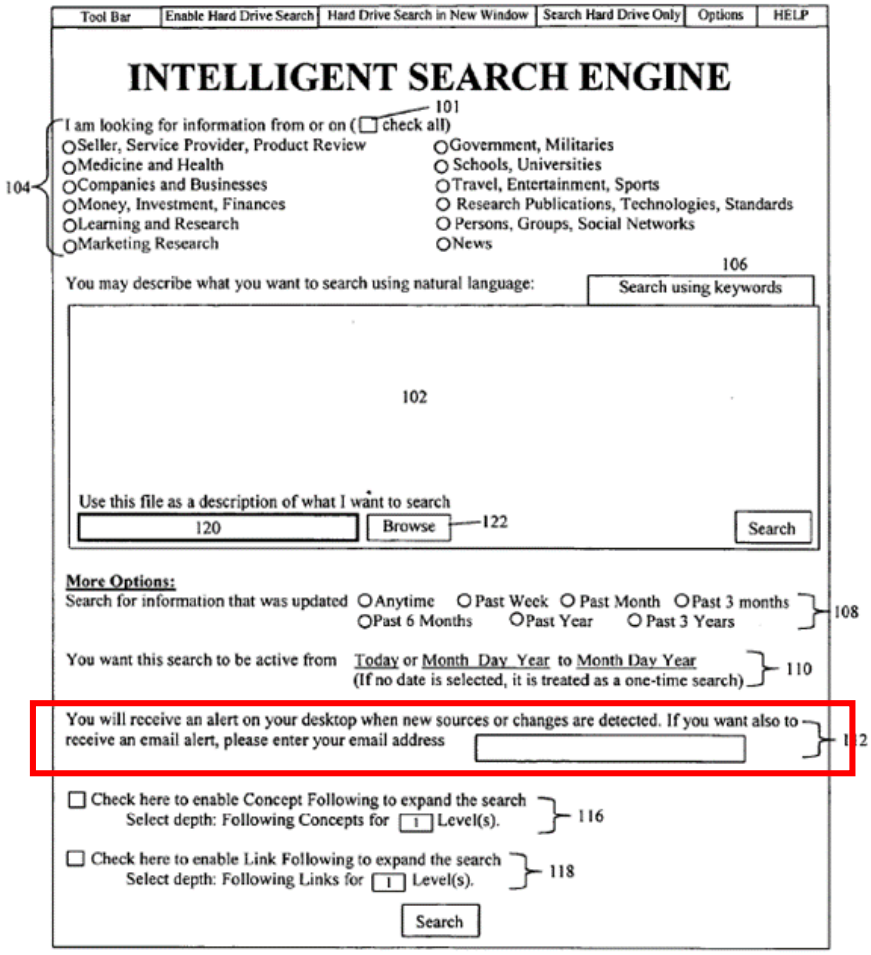


Ex. 1012, Figure 4 (annotations added). Liang explains that these "Stats" correspond to statistics, such as "the total number of appearances of an important concept in the search results" or "the number of web pages or files in the search results ... that match the filtering feature." Ex. 1012, [0184]. To the extent the claim requires that the query itself include non-textual data (as taught in Nelson, *see*, Sections (X)(G)(1)(b) and (c)), Nelson disclosed the use of filtering features "such as "/shape' and "/color" [that] enable the user to selectively define which aspects of the query images are of interest." *Id*., at 6:35-56. The number of appearances of a concept for a particular

source or the number of files/webpages matching Nelson's filtering feature for images, each correspond to an “accuracy” parameter for each source. Liang's system automatically "selects the first S, where S is a positive integer and can be set by default or by user" sources "from which the most number of web pages and files are retrieved in the search results." Ex. 1012, [0229]. These selected sources in Liang correspond to sources with query results (e.g., number of retrieved web pages and files) greater than or equal to an accuracy threshold (the top "S" sources).

f. Element 1.6 - “monitoring, based on a set schedule, the at least one data source to extract non-textual data from the at least one data source when at least an update to stored data that matches the query, newly added non-textual data that matches the query, or a combination thereof”

155. Figure 1 of Liang teaches alerting a user by email when “changes are detected.”



Ex. 1012, Figure 1 (annotations added). With respect to Figure 1, Liang states that "it offers a user ... the option to keep his search active for a period of time to *monitor for* new sources and *changes to existing sources* by specifying a date range in 110, and when a change is detected, the option to alert the user on his local PC or send an email to an email account that the user provides in 112." *Id.*, [0052] (emphasis added). Liang also teaches that "[b]oth the time period and the *monitoring frequency* may be chosen by a pull-down menu, or text box and check boxes." *Id.*, [0347]

(emphasis added). The "monitoring frequency" in Liang maps to "monitoring, based on a set schedule" in the claim.

156. Both Liang and Nelson teach returning non-textual data matching a query. *See*, Section XII(G)(1)(b). To the extent Liang does not teach returning images in response to a query, it would have been obvious to incorporate Nelson's techniques for returning images that match characteristics of a query image into Liang's retrieval system. Both Liang and Nelson are directed to search technology. A POSITA would have been motivated to combine the references in order to expand the capabilities of Liang such that a user can include both text and a query image in a query, and receive query results that include both text and images. A POSITA would have had a reasonable expectation of success with respect to this combination, as the combination involves nothing more than the routine combination of existing software functionality.

g. Element 1.7 - "establishing a communication channel with a second user based at least in part on the data extracted from the at least one monitored data source"

157. As explained in Sections VI(A) and XII(A)(1)(g), *supra*, this limitation is met by sending an email alert with updated query results to the second user "via the first user."

158. Lucovsky (Ex. 1016) discloses a first user sharing search results with a second user via an email (Figures 4-6), message board (Figure 7), instant messaging

(Figure 8) and a blog (Figure 9). Liang, Nelson and Lucovsky are all directed to search query technology. It would have been obvious to incorporate Lucovsky's search results sharing techniques with Liang and Nelson, so that a first user could share updates with a second user. A POSITA would have had a reasonable expectation of success in incorporating Lucovsky's search results sharing techniques with Liang and Nelson, as the combination involves nothing more than the routine combination of existing software functionality.

159. Accordingly, Claim 1 is obvious over Liang, Nelson and Lucovsky.

2. Claims 10 and 19

160. As noted in Section XII(A)(2), *supra*, Claim 10 further recites "a memory unit" and "at least one processor ... configured" to perform the recited functions, and Claim 19 further recites "a non-transitory computer-readable medium having program code recorded thereon" to perform the recites functions. These aspects of Claims 10 and 19 are met by Liang. Ex. 1012, [0039] ("the corresponding apparatus element may be configured in hardware, software, firmware or combinations thereof.") A POSITA would have understood "hardware, software" or "combinations thereof" to include memory that stored software code executed by a processor, as claimed.

161. Accordingly, Claims 10 and 19 are also obvious over Liang, Nelson and Lucovsky.

3. Claims 2, 11 and 20

162. Claims 2, 11 and 20 are contain substantively identical limitations. The analysis below of method Claim 2 applies identically to Claims 11 and 20.

a. Element 2.1 - "analyzing the extracted data"

163. Liang taught analyzing search results which maps to "analyzing the extracted data" in the claim. *See*, Ex. 1012, at [0167]. Alternatively, it would have been obvious to apply Liang's techniques for analyzing search results to "the extracted data" resulting from Liang's monitoring techniques. A POSITA would have had a reasonable expectation of success in applying Liang's techniques for analyzing search results to "the extracted data" resulting from Liang's monitoring techniques, as the combination involves nothing more than the routine combination of existing software functionality.

b. Element 2.2 - "receiving feedback from the user based at least in part on the established communication channel"

164. As mentioned above in Section XII(A)(3)(b), Lucovsky discloses a first user sharing search results with a second user via an email, and the second user providing a responsive message thereto. Such a responsive message maps to "receiving feedback from the user based at least in part on the established communication channel" in the claim.

165. Thus, Claims 2, 11 and 20 are obvious over Liang, Nelson and Lucovsky.

4. Claims 3 and 12

166. As explained above, the additional limitations of Claims 3 and 12 recite visualization of search results. As explained in Section XII(D), *supra*, Liang teaches visualization of search results using graphical diagrams. Thus, Claims 3 and 12 are also obvious over Liang, Nelson and Lucovsky.

5. Claims 4 and 13

167. The additional limitations of Claims 4 and 13 recite "expanding the search terms to include ... synonyms." Liang states that it "broadens a search by *expanding search keywords* into concepts comprising of the *synsets*, hypemym, and/or hyponym/troponym of a keyword ..." Ex. 1012, [0013]. A "synset" is a synonym set. Thus, Claims 4 and 13 are also obvious over Liang, Nelson and Lucovsky.

6. Claims 7 and 16

168. Claims 7 and 16 further recite that the query is a natural language query, a non-textual data query, or a combination thereof. As explained in Section XII(G)(1)(b), *supra*, Figure 1 of Liang teaches a natural language query. Thus, Claims 7 and 16 are also obvious over Liang and Lucovsky. Further, as explained in Section XII(G)(1)(b), *supra*, Nelson taught a compound query that included a

combination of a query image (image of a sunset) and text (the word 'beach'). Ex. 1019, at 3:11-12. Thus, Claims 7 and 16 are obvious over Liang, Nelson and Lucovsky.

7. Claims 8 and 17

169. Claims 8 and 17 further recite that the non-textual data comprises "an image, audio, a video, or a combination thereof." As explained in Section XII(G)(1)(b), *supra*, Liang taught that its search engine which could be used for, *inter alia*, querying non-textual data such as images. Further, as explained in Section XII(G)(1)(b), Nelson taught using an image as the query itself. Ex. 1019, at 3:6-17. Thus, Claims 8 and 17 are also obvious over Liang, Nelson and Lucovsky.

H. Ground 8 - Claims 1-8, 10-17 and 19-20 Are Obvious Over Liang, Nelson, Kirsch, Schultz and Lucovsky

1. Claims 1, 10 and 19

170. As set forth in Sections XII(G)(1) and (2), all of the limitations of Claims 1, 10 and 19 are obvious in view of Liang, Nelson and Lucovsky.

a. Element 1.4

171. With respect to Element 1.4 (*see* Section XII(G)(1)(d)), *supra*), to the extent Liang does not expressly teach executing the query on multiple data sources, it would have been obvious in view of Kirsch. As described in Section XII(A)(1)(d), *supra*, Kirsch teaches that a query is initially directed against multiple document collections. Ex. 1011, 6:29-30.

172. Liang, Nelson and Kirsch are all directed to search technology. It would have been obvious to incorporate techniques for querying a collection having image data as taught in Liang/Nelson into Kirsch's system for querying a plurality of collections. A POSITA would have been motivated to combine the references in order to apply the collection selection techniques of Kirsch to collections that included image data (in order to identify (from a larger set of document collections) a subset of document collections most relevant to the query). A POSITA would have had a reasonable expectation of success in applying the collection selection techniques of Kirsch to collections that included image data (in order to identify (from a larger set of document collections) a subset of document collections most relevant to the query), as the combination involves nothing more than the routine combination of existing software functionality.

b. Element 1.5

173. With respect to Element 1.5 (*see* Section XII(H)(1)(e)), *supra*), to the extent this feature is not expressly taught in Liang, it would have been obvious in view of Kirsch, Nelson and Schultz. As noted in Section XII(A)(5), *supra*, (i) Kirsch taught automatic selection of data sources based on a “minimum level of relevance,” “minimum relevancy ranking threshold” and/or requirement that a selected collection rank high enough to fall within a "maximum number of collections returned.” Nelson taught a query that included a query image. *See*, Section

XII(G)(1)(b), *supra*. Schultz taught relevance scoring and relevance ranking for image objects within a collection. Ex. 1013, 14:29-41, Figure 4A. It would have been obvious to use the collection selection techniques in Kirsch together with Schultz's scoring and ranking metrics for non-textual data in the context of Liang/Nelson to narrow the sources searched in Liang/Nelson to ones selected using Kirsch's techniques in order to filter out less relevant sources. A POSITA would have had a reasonable expectation of success in using the collection selection techniques in Kirsch together with Schultz's scoring and ranking metrics for non-textual data in the context of Liang/Nelson to narrow the sources searched in Liang/Nelson to ones selected using Kirsch's techniques in order to filter out less relevant sources, as the combination involves nothing more than the routine combination of existing software functionality.

c. Element 1.6

174. It would have been obvious to apply Liang's monitoring to the "qualified set" of data sources selected using Kirsch's collection selection techniques and Schultz's metrics. A person would have been motivated to limit the number of data sources used for monitoring in order avoid identification of less relevant results during the monitoring. A POSITA would have had a reasonable expectation of success in applying Liang's monitoring to the "qualified set" of data sources selected using Kirsch's collection selection techniques and Schultz's metrics, as the

combination involves nothing more than the routine combination of existing software functionality.

175. Accordingly, Claims 1, 10 and 19 are also obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

2. Claims 2, 11 and 20

176. The additional limitations of dependent claims 2, 11 and 20 are taught by Liang and Lucovsky. See, Section XII(G)(3), *supra*. Claims 2, 11 and 20 are therefore obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

3. Claims 3 and 12

177. As explained above, the additional limitations of Claims 3 and 12 (visualization of search results) are met by Liang. *See*, Section XII(G)(3), *supra*. Thus, Claims 3 and 12 are also obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

4. Claims 4 and 13

178. As explained in Section XII(G)(4), the additional limitations of Claims 4 and 13 are taught by Liang. Alternatively, as explained in Section XII(A)(3), *supra*, Kirsch teaches expanding the search terms to include synonyms."

179. Thus, Claims 4 and 13 are also obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

5. Claims 5, 14 and 20

180. As explained in Section XII(A)(5), *supra*, the additional limitations of Claims 5, 14 and 21 are met by Lucovsky. It would have been obvious to incorporate Lucovsky's search refinement techniques into the other references, so that a user could refine search terms based on the query and monitor for matches to the refined query. A POSITA would have had a reasonable expectation of success in incorporating Lucovsky's search refinement techniques into the other references, so that a user could refine search terms based on the query and monitor for matches to the refined query, as the combination involves nothing more than the routine combination of existing software functionality.

181. Thus, Claims 5, 14 and 21 are obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

6. Claims 6 and 15

182. As explained in Section XII(A)(5), *supra*, the additional limitations of Claims 6 and 15 are taught by Kirsch. It would have been obvious to apply the teachings of Liang, Nelson and Schultz to document collections that include both public and proprietary data, as taught in Kirsch. Ex. 1011, 1:33-42. A POSITA would have been motivated to implement these aspects of Kirsch in combination with the other references in order to allow for searching of both public and private collections. A POSITA would have had a reasonable expectation of success in

applying the teachings of Liang and Schultz to document collections that include both public and proprietary data, as taught in Kirsch, as the combination involves nothing more than the routine combination of existing software functionality.

183. Thus, Claims 6 and 15 are obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

7. Claims 7 and 16

184. As explained in Section XII(G)(6), the additional limitations of Claims 7 and 16 are taught by Liang (natural language search query) amd Nelson (compund query). Thus, Claims 7 and 16 are also obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

8. Claims 8 and 17

185. As explained in Section XII(G)(7), *supra*, the additional limitations of Claims 8 and 17 are met by Liang and Nelson. Thus, Claims 8 and 17 are obvious over Liang, Nelson and Lucovsky in combination with Kirsch and Schultz.

I. Ground 9 - Claims 1, 4, 7, 8, 10, 13, 16, 17 and 19 Are Obvious Over Liang, Nelson and Mellin

1. Claim 1

186. As set forth in Section XII(G)(1)(a)-(f), the Elements 1.1-1.6 of Claim 1 are taught by Liang and Nelson. Instead of relying on Lucovsky for Element 1.7 as set forth in Ground 7 above, Ground 9 maps Element 1.7 to Mellin which taught the use of a group email address to facilitate collaboration among members of a

group. *See*, Mellin (Ex. 1017), at 3:24-26. In a case where the author of the query (mapped to "a first user" in the claim) was part of a group collaboration, it would have been obvious for the author to designate a group email address to receive search result updates so that the updated search results were distributed to all collaborators. Any collaborator on the group email address other the author of the search query maps to "a second user" in the claim. Accordingly, Claim 1 is obvious over Liang, Nelson and Mellin.

2. Claims 4, 7, 8, 10, 13, 16, 17 and 19

187. As set forth in Section XII(G)(2), the additional limitations of Claims 10 and 19 are met by Liang.

188. As set forth in Section XII(G)(5), the additional limitations of Claims 4 and 13 are met by Liang.

189. As set forth in Section XII(G)(6), the additional limitations of Claims 7 and 16 are met by Liang and Nelson.

190. As set forth in Section XII(G)(7), *supra*, the additional limitations of Claims 8 and 17 are met by Liang and Nelson.

191. Accordingly, Claims 4, 7, 8, 10, 13, 16, 17 and 19 are obvious over Liang, Nelson and Mellin.

J. Ground 10 - Claims 2, 3, 5, 6, 11, 12, 14, 15 And 20 Are Obvious Over Liang, Nelson and Mellin In Combination With Kirsch And Schultz

1. Claims 2, 11 and 20

192. As set forth in Section XII(I), Claims 1, 10 and 19 are obvious over Liang, Nelson and Mellin. As set forth in Sections XII(A)(3)(a) and XII(C)(2), *supra*, the additional limitations of Claims 2, 11 and 20 are taught by Kirsch and Schultz. For the reasons set forth in Section XII(H), supra, it would have been obvious to combine Liang and Nelson with Kirsch and Schultz. Accordingly, Claims 2, 11 and 20 are obvious over Liang, Nelson and Mellin, in view of Kirsch and Schultz.

2. Claims 3, 5, 6, 12, 14 and 15.

193. As set forth in Section XII(D), *supra*, the additional limitations of Claims 3 and 12 are taught in Liang.

194. As set forth in Section XII(A)(5), *supra*, the additional limitations of Claims 5 and 14 are taught by Schultz.

195. As set forth in Section X(A)(6), *supra*, the additional limitations of Claims 6 and 15 are taught by Schultz.

196. Accordingly, Claims 3, 5, 6, 12, 14 and 15 are obvious over Liang, Nelson and Mellin, in view of Kirsch and Schultz.

K. Ground 11 - Claims 9 and 18 Are Obvious Over Liang, Nelson, Lucovsky And Vinsonneau

197. As set forth in Section XII(G)(1), *supra*, the limitations of Claim 1 are taught by Liang, Nelson and Lucovsky, and as set forth in Section XII(E), *supra*, the limitations of Claims 9 and 18 are taught in Vinsonneau. It would have been obvious to combine Vinsonneau's scaling with Liang's monitoring techniques in order to improve the performance of Liang as traffic volume varied. A POSITA would have had a reasonable expectation of success in combining Vinsonneau's scaling techniques with Liang's monitoring techniques in order to improve the performance of Liang as traffic volume varied, as the combination involves nothing more than the routine combination of existing software functionality. Thus, Claims 9 and 18 are also obvious over Liang, Nelson, Lucovsky and Vinsonneau.

L. Ground 12 - Claims 9 and 18 Are Obvious Over Liang, Nelson, Lucovsky, Kirsch And Schultz, In View Of Vinsonneau

198. As set forth in Section XII(H)(1), *supra*, the limitations of Claim 1 are taught by Liang, Nelson, Lucovsky, Kirsch and Schultz, and as set forth in Section XII(E), *supra*, the limitations of Claims 9 and 18 are taught in Vinsonneau. It would have been obvious to combine Vinsonneau's scaling with Liang's monitoring in order to improve the performance of Liang as traffic volume varied. A POSITA would have had a reasonable expectation of success in combining Vinsonneau's scaling techniques with Liang's monitoring techniques in order to improve the performance

of Liang as traffic volume varied, as the combination involves nothing more than the routine combination of existing software functionality. Thus, Claims 9 and 18 are also obvious over Liang, Nelson, Lucovsky, Kirsch and Schultz, in view of Vinsonneau.

M. Ground 13 - Claims 1-20 Are Obvious Over The '301 Publication And Schultz

1. Claim 1

199. As set forth below, all of the elements of Claim 1 are met by the '301 Publication and Schultz.

a. Element 1.1 - "A method of analyzing data"

200. The '301 Publication describes "a method of analyzing data." Ex. 1014, [0009].

b. Element 1.2 - "receiving, from a first user, a query for non-textual data"

201. The '301 Publication discloses that a query is generated by the user. *Id*., [0009]; [0026]. Schultz taught a user query for querying non-textual data such as images. Ex. 1013, 12:49-53 and 12:66-13:4. The '301 Publication and Schultz both relate to search technology, and it would have been obvious to incorporate Schultz's techniques for querying non-textual image data into the '301 Publication. A POSITA would have been motivated to combine the references in order to apply Schultz's image querying techniques into the '301 Publication. A POSITA would have had a

reasonable expectation of success in incorporating Schultz's techniques for querying non-textual image data into the '301 Publication, as the combination involves nothing more than the routine combination of existing software functionality.

c. Element 1.3 - "expanding search terms of the query"

202. The '301 Publication teaches "expanding search terms of the query." Ex. 1014, [0009].

d. Element 1.4 - "executing the query on a plurality of data sources"

203. The '301 Publication teaches "executing the query on one or more data sources." *Id*., [0009].

e. Element 1.5 - "selecting at least one data source from the plurality of data sources, the at least one data source being selected when results of the query are greater than or equal to an accuracy thresholds"

204. The '301 Publication teaches that "one or more of the data sources are selected based on the accuracy of the search results." *Id*., [0026]. Schultz taught relevance scoring and relevance ranking for image objects within a collection. Ex. 1013, 14:29-41, Figure 4A.

205. It would have been obvious to apply the relevance scoring and relevance ranking techniques for images taught in Schultz to the system in the '301 Publication in order to select collection(s) having at least one image with a relevance score over a minimum level of relevance and/or a relevance rank over a minimum ranking threshold. It would have been a matter of simple substitution to use Schultz's

metrics for relevance scoring and/or relevance ranking of images for collection selection within the '301 Publication. A POSITA would have had a reasonable expectation of success in applying the relevance scoring and relevance ranking techniques for images taught in Schultz to the system in the '301 Publication in order to select collection(s) having at least one image with a relevance score over a minimum level of relevance and/or a relevance rank over a minimum ranking threshold, as the combination involves nothing more than the routine combination of existing software functionality.

f. Element 1.6 - "monitoring, based on a set schedule, the at least one data source to extract non-textual data from the at least one data source when at least an update to stored data that matches the query, newly added non-textual data that matches the query, or a combination thereof"

206. The '301 Publication teaches "monitoring a specific data source selected from the one or more data sources. The monitoring is performed to monitor for matches to the query." Ex. 1014, [0009]. The '301 Publication also teaches "extracting data from the at least one monitored specific data source ..." *Id*., Claim 2. Finally, the '301 Publication teaches that the "user may also set a schedule for the extractor to determine if there is new content ... The schedule may be once a week, hour, month, minute, or other time." *Id*., [0048].

207. Schultz taught a user query for querying non-textual data such as images. Ex. 1013, 12:49-53 and 12:66-13:4. The '301 Publication and Schultz both

relate to search technology, and it would have been obvious to incorporate Schultz's techniques for querying non-textual image data into the '301 Publication's monitoring step so as to permit monitoring of non-textual data. A POSITA would have had a reasonable expectation of success in incorporating Schultz's techniques for querying non-textual image data into the '301 Publication's monitoring step so as to permit monitoring of non-textual data, as the combination involves nothing more than the routine combination of existing software functionality.

g. Element 1.7 - "establishing a communication channel with a second user based at least in part on the data extracted from the at least one monitored data source"

208. The '301 Publication teaches that "a communication channel may be established with specific individuals and/or *other users* based on the analyzed data." Ex. 1014, [0032]. The "other users" in [0032] of the '301 Publication map to a "second user," as claimed. The '301 Publication further teaches "extracting data from the at least one monitored specific data source; analyzing the extracted data; establishing a communication channel with a user based at least in part on the analyzed data." *Id*., Claim 2.

209. Accordingly, Claim 1 was obvious in view of the '301 Publication and Schultz.

2. Claims 10 and 19

210. The additional limitations of Claim 10 (memory and a processor) and Claim 19 (program code executed by a processor) are taught in the '301 Publication. Ex. 1014, Claims 10 and 19.

3. Claims 2, 11 and 20

211. The additional limitations of these claims are met by the '301 Publication, which teaches (i) "[t]he method of claim 1, further comprising ... analyzing the extracted data" (*id*., Claim 2), and (ii) that "customers and/or users may provide feedback. ... [T]he feedback may also be generated based on the communication channel established at block 116." *Id*., [0032].

4. Claims 3 and 12

212. The additional limitations of these claims are met by the '301 Publication, which teaches "[t]he method of claim 2, further comprising visualizing at least the extracted data, the analyzed data, or a combination thereof." *Id*., Claim 3.

5. Claims 4 and 13

213. The additional limitations of these claims are met by the '301 Publication, which teaches "[t]he method of claim 1, in which expanding the query comprises expanding the search terms to include at least misspellings, synonyms, sub-topics, antonyms, or a combination thereof." *Id*., Claim 4.

6. Claims 5 and 14

214. The additional limitations of these claims are met by the '301 Publication, which teaches "[t]he method of claim 1, further comprising: refining the search terms based at least in part on the executed query, and in which the monitoring is performed to monitor for matches to a refined query." *Id*., Claim 5.

7. Claims 6 and 15

215. The additional limitations of these claims are met by the '301 Publication, which teaches "[t]he method of claim 1, in which the plurality of data sources include open-source data sources that are publicly available and closed-source data sources that at not publicly accessible." *Id*., Claim 6.

8. Claims 7 and 16

216. The additional limitations of these claims are met by the '301 Publication, which teaches "[t]he method of claim 1, in which the query is a natural language query..." *Id*., Claim 7.

9. Claims 8 and 17

217. The additional limitations of these claims are met by Schultz, which teaches non-textual data in the form of an image. Ex. 1013, 12:49-53 and 12:66-13:4.

10. Claims 9 and 18

218. The additional limitations of these claims are met by the '301 Publication, which teaches "[t]he method of claim 1, in which the monitoring comprises deploying extractors to each of specific data source, a number of extractors being dynamic based at least in part on a traffic volume of each specific data source." Ex. 1014, Claim 9.

XIII. DECLARATION

219. I declare that all statements made herein of my own knowledge are true and that all statements made on information and belief are believed to be true, and that these statements were 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 tile 18 of the United States Code.

Dated: August 22, 2024

Dr. Jaideep Srivastava, Ph.D.