

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

BRIGHT DATA LTD.,

Petitioner

v.

OXYLABS, UAB,

Patent Owner

Case IPR2023-01441

U.S. Patent No. 7,010,526

PETITION FOR *INTER PARTES* REVIEW

Mail Stop PATENT BOARD
Patent Trial and Appeal Board
United States Patent and Trademark Office
P.O. Box 1450
Alexandria, VA 22313-1450

<u>PETITIONER’S LIST OF EXHIBITS</u>	
EX.1001	U.S. Patent No. 7,010,526
EX.1002	Prosecution History of U.S. Patent No. 7,010,526
EX.1003	U.S. Pub. No. 2003/0163454 to Jacobsen et al. (“Jacobsen”)
EX.1004	U.S. Patent No. 7,082,427 to Seibel et al. (“Seibel”)
EX.1005	U.S. Pub. No. 2002/0049773 to Jaramillo (“Jaramillo”)
EX.1006	U.S. Patent No. 6,606,659 to Hegli et al. (“Hegli”)
EX.1007	“Creating” Definition and Meaning - Merriam-Webster; accessed at https://www.merriam-webster.com/dictionary/creating on September 26, 2023
EX.1008	Declaration of Dr. Tim A. Williams
EX.1009	Appendix of Challenged Claims
EX.1010	U.S. Pub No. 2002/0091923 to Chipman et al. (“Chipman”)
EX.1011	U.S. Patent No. 6,691,105 to Viridy (“Viridy”)
EX.1012	U.S. Pub. No. 2003/0112270 to Newell et al. (“Newell”)
EX.1013	Docket Control Order (Dkt. 28) in the case of <i>Bright Data Ltd. v. Oxylabs, UAB</i> , No. 2:23-cv-171 (E.D. Tex.)(June 21, 2023)
EX.1014	U.S. District Courts–Combined Civil and Criminal Federal Court Management Statistics (June 30, 2023), accessed at https://www.uscourts.gov/statistics/table/na/federal-court-management-statistics/2023/06/30-1
EX.1015	Order (Dkt. 129) in the case of <i>Oxylabs, UAB v. Bright Data Ltd.</i> , No. 2:22-cv-11 (E.D. Tex.)(April 14, 2023)

EX.1016	Oxylabs' Answer to Bright Data's Counterclaim and Oxylabs' Counterclaims (Dkt. 15) in the case of <i>Bright Data Ltd. v. Oxylabs, UAB</i> , No. 2:23-cv-171 (E.D. Tex.)(April 28, 2023)
---------	---

TABLE OF AUTHORITIES

Cases

<i>Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH</i> , IPR2019-01469, Paper 6 (PTAB Feb. 13, 2020)	73
<i>Apple Inc. v. Fintiv, Inc.</i> , IPR2020-00019, Paper 11 (PTAB Mar. 20, 2020) . <i>passim</i>	
<i>Becton, Dickinson & Co. v. B. Braun Melsungen AG</i> , IPR2017-01586, Paper 8 (PTAB Dec. 15, 2017)	76
<i>Thorne Res., Inc. v. Trs. of Dartmouth Coll.</i> , IPR2021-00491, Paper 18 (PTAB Aug. 12, 2021)	73

Statutes

35 U.S.C. § 314	1, 68
35 U.S.C. § 325	68

Other Authorities

Director Katherine K. Vidal, ‘Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation’ (June 21, 2022)	68, 69, 70, 72
--	----------------

Rules

37 C.F.R. § 42.103	3
--------------------------	---

37 C.F.R. § 42.104	3
37 C.F.R. § 42.108(c).....	1
37 C.F.R. § 42.15(a).....	3
37 C.F.R. § 42.8	1

TABLE OF CONTENTS

I.	INTRODUCTION	1
II.	STATUTORY PREDICATES	1
	A. Mandatory Notices (37 C.F.R. §42.8)	1
	1. Real Parties-in-Interest	1
	2. Related Matters.....	1
	3. Lead and Back-Up Counsel.....	2
	4. Service Information	2
	B. Payment of Fees (37 C.F.R. §42.103)	3
	C. Certification of Standing (37 C.F.R. §42.104(a))	3
	D. Identification of Challenges (37 C.F.R. §42.104(b)(1)-(2)).....	3
III.	OVERVIEW OF THE '526 PATENT.....	4
	A. Summary of the '526 Patent.....	5
	B. Claims of the '526 Patent	7
	C. Prosecution History of the '526 Patent.....	8
	D. Level of Skill in the Art.....	13
	E. Claim Construction.....	13
IV.	OVERVIEW OF PRIOR ART	16
	A. Applicant Admitted Prior Art.....	16
	B. Jacobsen (EX.1003).....	17
	C. Seibel (EX.1004)	19
	D. Jaramillo (EX.1005)	20
	E. Hegli (EX.1006)	21
V.	ANALYSIS OF THE GROUNDS.....	22
	A. GROUND 1: ANTICIPATION OF CLAIMS 1, 3, 19, 20 34, 36, 52, AND 53 BY JACOBSEN (EX.1003).....	22
	1. Claim 1	22
	2. Claim 3	27

3. Claim 19	28
4. Claim 20	29
5. Claim 34	29
6. Claim 36	31
7. Claim 52	31
8. Claim 53	32
B. GROUND 2: OBVIOUSNESS OF CLAIMS 2 AND 35 BASED ON JACOBSEN (EX.1003)	
1. Claim 2	32
2. Claim 35	34
C. GROUND 3: OBVIOUSNESS OF CLAIMS 2 AND 35 BASED ON JACOBSEN (EX.1003) AND JARAMILLO (EX.1005).....	
1. Motivations to Combine Regarding Jacobsen and Jaramillo.....	34
2. Claim 2	36
3. Claim 35	38
D. GROUND 4: OBVIOUSNESS OF CLAIMS 6-9, 11, 16-17, 39-42, 44, and 49-50 BASED ON JACOBSEN (EX.1003) AND HEGLI (EX.1006)	
1. Motivation to Combine Jacobsen and Hegli	38
2. Claim 6	41
3. Claim 7	44
4. Claim 8	45
5. Claim 9	45
6. Claim 11	46
7. Claim 16	47
8. Claim 17	48
9. Claim 39	48
10. Claim 40	48
11. Claim 41	49
12. Claim 42	49
13. Claim 44	49
14. Claim 49	49
15. Claim 50	50
E. GROUND 5: ANTICIPATION OF CLAIMS 1, 6, 9, 11, 13, 19, 20, 34, 39, 42, 44, 46, 52, AND 53 BY SEIBEL (EX.1004)	
1. Claim 1	50

2. Claim 6	60
3. Claim 9	60
4. Claim 11	61
5. Claim 13	61
6. Claim 19	61
7. Claim 20	62
8. Claim 34	63
9. Claim 39	64
10. Claim 42	65
11. Claim 44	65
12. Claim 46	65
13. Claim 52	65
14. Claim 53	66
F. GROUND 6: OBVIOUSNESS OF CLAIMS 2 AND 35 BASED ON SEIBEL (EX.1004)	66
1. Claim 2	66
2. Claim 35	68
VI. DISCRETIONARY DENIAL WOULD NOT BE APPROPRIATE.....	68
A. The factors under §314 weigh against discretionary denial	68
B. The factors under §325 weigh against discretionary denial	73
VII. CONCLUSION	76

I. INTRODUCTION

Petitioner Bright Data Ltd. respectfully requests *inter partes* review and cancellation of claims 1-3, 6-9, 11, 13, 16-17, 19-20, 34-36, 39-42, 44, 46, 49-50, and 52-53 (“Challenged Claims”) of U.S. Patent No. 7,010,526 to Burckart, et al. (EX.1001, “the ’526 Patent”). As outlined in this petition, there is a reasonable likelihood that at least one of the Challenged Claims is unpatentable. 35 U.S.C. §314(a); 37 C.F.R. §42.108(c). This petition is supported by the expert declaration of Dr. Tim A. Williams (EX.1008).

II. STATUTORY PREDICATES

A. Mandatory Notices (37 C.F.R. §42.8)

Pursuant to 37 C.F.R. §42.8, Petitioner Bright Data Ltd. hereby provides the following mandatory notices in connection with this *inter partes* review of Patent No. 7,010,526 titled “Knowledge-based data mining system.”

1. Real Parties-in-Interest

Petitioner certifies that Bright Data Ltd. is the real party-in-interest.

2. Related Matters

There is a related district court case involving the ‘526 Patent, styled as *Bright Data Ltd. v. Oxylabs, UAB*, Case No. 2:23-cv-171 (E.D. Tex.). That case also

involves U.S. Patent No. 7,020,667 assigned to Patent Owner. That case also involves U.S. Patent No. 11,206,317 assigned to Petitioner.

There are no related administrative matters involving the '526 Patent.

3. Lead and Back-Up Counsel

Lead Counsel	Back-Up Counsel
Thomas M. Dunham Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036 Tel. (202) 838-1567 Fax (202) 478-1715 tomd@cherianllp.com (for correspondence from Patent Owner) tom@dunham.cc (for correspondence from P-TACTS) USPTO Customer No. 144371 USPTO Reg. No. 39,965	Robert W. Molitors Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036 Tel. (202) 838-1563 Fax (202) 478-1715 rmolitors@cherianllp.com USPTO Customer No. 144371 USPTO Reg. No. 66,726 Elizabeth A. O'Brien Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036 Tel. (202) 873-1726 Fax (202) 478-1715 elizabetho@cherianllp.com USPTO Customer No. 144371 USPTO Reg. No. 76,105

4. Service Information

Electronic Mail	1. tomd@cherianllp.com (for correspondence from Patent Owner) or tom@dunham.cc (for correspondence from P-TACTS) 2. rmolitors@cherianllp.com
-----------------	---

	3. elizabetho@cherianllp.com
Postal (and hand delivery) mailing addresses:	Thomas M. Dunham Robert W. Molitors Elizabeth A. O'Brien Cherian LLP 1901 L Street NW, Suite 700 Washington, D.C. 20036
Telephone	Tel. (202) 838-1567
Facsimile:	Fax (202) 478-1715

Additionally, Petitioner consents to electronic service via email at the email addresses noted above.

B. Payment of Fees (37 C.F.R. §42.103)

The Office is authorized to charge Deposit Account No. 603803 (Customer No. 144371) the fees set forth in 37 C.F.R. §42.15(a), or any other applicable fees, for this Petition.

C. Certification of Standing (37 C.F.R. §42.104(a))

Petitioner certifies that the '526 Patent is available for *inter partes* review and that Petitioner is not barred or estopped from requesting *inter partes* review of the Challenged Claims on the grounds alleged herein.

D. Identification of Challenges (37 C.F.R. §42.104(b)(1)-(2))

Petitioner requests cancellation of the Challenged Claims on the following grounds:

Ground 1	Claims 1, 3, 19-20, 34, 36, and 52-53	§102 Anticipation by Jacobsen
Ground 2	Claims 2 and 35	§103 Obviousness based on Jacobsen + General knowledge of a person of ordinary skill in the art
Ground 3	Claims 2 and 35	§103 Obviousness based on Jacobsen + Jaramillo
Ground 4	Claims 6-9, 11, 16-17, 39-42, 44, and 49-50	§103 Obviousness based on Jacobsen + Hegli
Ground 5	Claims 1, 6, 9, 11, 13, 19-20, 34, 39, 42, 44, 46, and 52-53	§102 Anticipation by Seibel
Ground 6	Claims 2 and 35	§103 Obviousness based on Seibel + General knowledge of a person of ordinary skill in the art

III. OVERVIEW OF THE '526 PATENT

The '526 Patent issued from U.S. Application No. 10/141,686 filed on 5/8/2002. EX.1001, codes (21) and (22).

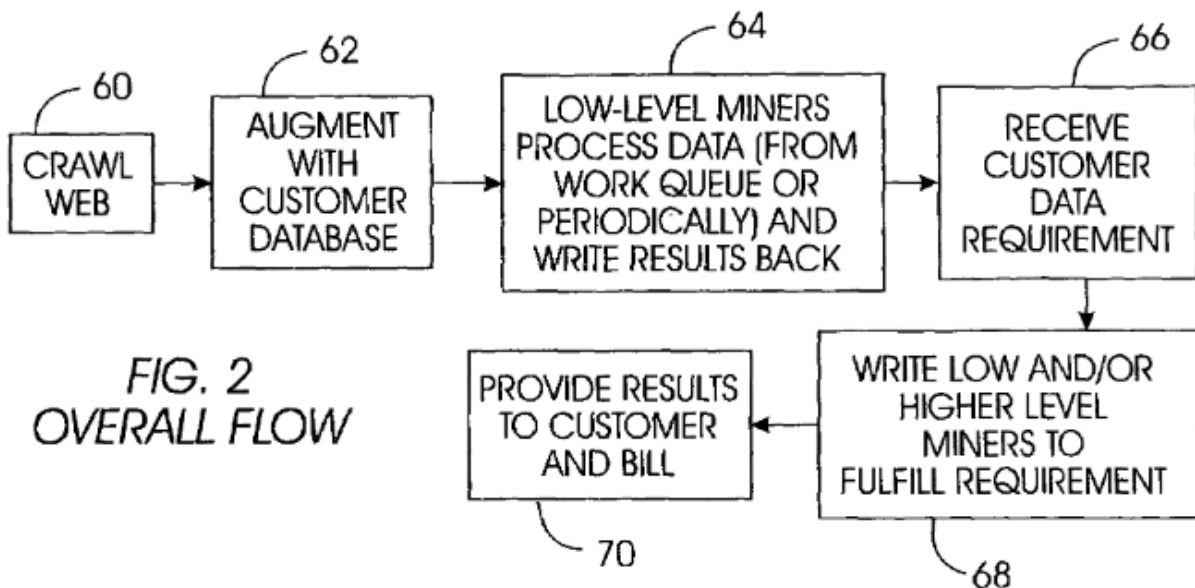
A. Summary of the '526 Patent

The '526 Patent titled “Knowledge-Based Data Mining System” relates to data mining systems where data is gathered into a data store using a web crawler where the data may then be classified into entities. EX.1001 at Abstract; EX.1008 at ¶32. Data miners may use rules to process the entities and append respective keys to the entities that represent characteristics of the entities. *Id.*

In the background section, the '526 Patent describes “data mining” as “the process of extracting user-desired information from a corpus of information” *Id.* at 1:11-12; EX.1008 at ¶33. The '526 Patent describes that data mining can be labor intensive because it requires a great deal of mundane sifting through query results by a person. *Id.* at 1:32-35. The '526 Patent notes that expertise in a particular area is often required “to separate the wheat from the chaff.” *Id.* at 1:35-36. As recognized by the specification: “one expert is required to process documents using his or her expert criteria to winnow out a subset of the documents, and a second expert must then use his or her expert criteria to locate the required information in the subset from the first expert.” *Id.* at 1:37-45.

The '526 Patent discloses a system that “can be used for enterprise data analysis, competitive intelligence, trending, discovery, web portal services, clustering and taxonomy.” *Id.* at 3:13-16; EX.1008 at ¶34. FIG. 2, reproduced below, shows the “overall logic of system 10 operation”. EX.1001 at 8:13-14. At block 60,

the crawler 12 crawls the Web 14 to add data to the data store 18. *Id.* at 8:14-15. At block 64, low level miners access and process the data and write the results back to the data store 18. *Id.* at 8:18-31. For example, the low level miners may write back respective keys to the corresponding entities. *Id.* At block 66, customer requests for information can be received. *Id.* at 8:32-33. At block 68, additional low level miners or high level miners, if required and not yet written, can be created. *Id.* at 8:33-39. For example, high level miners may specify dependencies on the outputs of other miners, e.g., results according to a geospatial key appended to the results of the low level miners. *Id.* at 8:40-49. At block 70, results are provided to the customer and the customer is billed. *Id.* at 8:54-56.



B. Claims of the '526 Patent

The '526 Patent includes 54 claims, with claims 1, 21, and 34 being independent.

Exemplary independent claim 1¹ recites a method as described below:

1. A method for data mining, comprising:

crawling the Internet to return Web pages to a data store;

accessing the data store with at least a first data miner to process

data in the store and output at least one result:

receiving a customer request for information;

creating at least a second data miner in response to the customer

request for information, the second data miner accessing at

least the result of the first data miner to generate an output;

and

¹ Exemplary independent claim 1 recites an age-old method that is regularly performed by humans sifting through search results “to separate the wheat from the chaff” and merely using a computer as a tool. EX.1001 at 1:32-36; EX.1008 at ¶30, note 1. Exemplary claim 1 merely mimics what humans do to search for information. EX.1001 at 1:38-59.

at least partially based on the output of the second data miner,
responding to the customer request for information.

This Petition challenges claims 1, 2-3, 6-9, 11, 13, 16-17, 19-20, 34, 35-36, 39-42, 44, 46, 49-50, and 52-53 (independent claims underlined). An Appendix of Challenged Claims is provided in EX.1009.

C. Prosecution History of the '526 Patent

The underlying application that issued as the '526 Patent was filed on May 8, 2002 with 54 claims and titled as “Knowledge-Based Data Mining System.” EX.1002 at 119-156.²

On March 24, 2005, the examiner issued a Non-final Rejection objecting to the title of the specification and rejecting claims 1, 3-5, 19-20, 32-34, 36-38, and 52-53 as being anticipated by U.S. Pub. No. 2002/0091923 to Chipman et al. (EX.1010, “Chipman”). EX.1002 at 89. The examiner also rejected claims 2 and 35 as being obvious over Chipman and U.S. Pub. No. 2002/0049773 to Jaramillo (EX.1005, “Jaramillo”); claims 6-11, 16-17, 21-27, 29-30, 39-44, and 49-50 as being obvious over Chipman and U.S. Patent No. 6,606,659 to Hegli (EX.1006, “Hegli”); claims

² The numbering of the claims did not change during prosecution. EX.1002 at 15.

E.g., original claim 1 corresponds to issued claim 1.

12-15, 18, 28, 31, 45-48, and 51 as being obvious over Chipman and Hegli and U.S. Patent No. 6,691,105 to Virdy (EX.1011, “Virdy”); and claim 54 as being obvious over Chipman and U.S. Pub. No. 2003/0112270 to Newell et al. (EX.1012, “Newell”). EX.1002 at 94, 95, 100, 101.

On March 29, 2005, Applicant filed Remarks and amended claims 1 and 34 to recite “creating” a second miner rather than “establishing” it. EX.1002 at 73-85.³ Similar amendments were made in dependent claims 3 and 36 for consistency. *Id.* With respect to claims 1 and 34, Applicant argued that these amendments overcome Chipman because in the relied-upon section of Chipman alleged to teach establishing a miner, no new minors are created; rather, the request is passed to pre-existing search engines. EX.1002 at 83. Applicant argued that picking a pre-existing search engine is an overly broad interpretation of claims 1 and 34. *Id.* With respect to claim 21, Applicant argued that in Chipman, no representation of the Web pages is ever “tagged” with anything, since “tag” clearly connotes “append.” EX.1002 at 84. Applicant also argued that in Hegli, no score is appended to representations of the

³ As support for the amendments, Applicant relied on, *e.g.*, the portion of the specification describing Fig. 2 which states: “At block 68, additional low level miners can be written in response, or high level miners, if required and not yet written, can be created.” EX.1002 at 83; *see* EX.1001 at 8:33-35.

Web pages from which the scores are derived. *Id.* Applicant also argued that there is no prior art motivation for the combination of Chipman and Hegli because Chipman is an information retrieval system that does not appear to be concerned at all with the spam/pornography access control of Hegli. *Id.*

On June 16, 2005, the examiner issued a Final Rejection maintaining its earlier objection to the title and the anticipation and obviousness rejections. EX.1002 at 47-66. With respect to claims 1 and 34, the examiner stated that Chipman's disclosure of a request being parsed and stored would be equivalent to the claim limitation "creating a miner in response to a user request information." EX.1002 at 63. With respect to claim 21, the examiner stated that Chipman discloses a Web page including a "class" identifier, where the "'class' is stored in the catalog that allow[s] a user to search for and identify the accessed item". EX.1002 at 63. The examiner also stated that Hegli teaches "tagging Web pages with a key representing a characteristic of the Web page" because tagging Web pages with a key or tagging a category with a key are interchangeable since each category represents a particular type of Web page. EX.1002 at 64. The examiner also stated that there is a motivation to combine Chipman and Hegli because MLS is limited to real estate agents only in order to prevent users without legitimate reasons from accessing the desired data. EX.1002 at 64-65.

On July 5, 2005, Applicant filed a Notice of Appeal and an Appeal Brief requesting review of the Examiner's objection/rejections. EX.1002 at 28-44. With respect to claims 1 and 34, Applicant challenged how parsing and storing a search request could be equated with creating a data miner to explore that search request. EX.1002 at 34. With respect to claim 21, Applicant challenged how Chipman's disclosure of storing strings in the RIDB, which is separate from the Web pages found by the crawler, could be equated with tagging a Web page. EX.1002 at 34-35. Applicant also challenged whether assigning a category to a page as taught by Hegli meets the language of claim 21 requiring tagging/append. *Id.* Applicant also challenged the motivation to combine Chipman and Hegli and whether tagging could be used as a way to exclude certain users (buyers and sellers) from accessing MLS for real-estate agents only, when this is usually done using access control principles such as passwords. EX.1002 at 36. Applicant also challenged the motivation to

combine Chipman and Jaramillo⁴, Chipman and Hegli and Viridy⁵, and Chipman and Newell⁶.

In response to the Appeal Brief dated July 5, 2005, the Examiner issued a Notice of Allowance on October 4, 2005. EX.1002 at 9-11. The reasons for allowance for claims 1-20 and 34-54 states that “the prior art of record fails to teach or fairly suggest creating at least a second data miner in response to the customer request for information.” *Id.* at 10. The reasons for allowance for claims 21-33 states that “the prior art of record fails to teach or fairly suggest at least one lower level data mining element receiving information from the data store and based thereon tagging at least one representation of a Web page with at least one key representing

⁴ Applicant argued that Jaramillo is focused on limiting access, not searching per se. EX.1002 at 36. Applicant also stated that Jaramillo bills only by subscription, not query, and so even if Chipman and Jaramillo were combined, they would not teach the limitations of claim 2. *Id.*

⁵ Applicant argued that Viridy is directed to organizing businesses and does not relate to data mining of the present invention or to access control of Hegli. EX.1002 at 36.

⁶ Applicant argued that Newell is directed to a litigation management system and completely unrelated to the technology of Chipman. EX.1002 at 37.

at least one characteristic of the Web page and at least one higher level data mining element using the key to output a result useful for responding to a request for information.” *Id.*

D. Level of Skill in the Art

A person of ordinary skill in the art (“POSA”) would have had, as of May 8, 2002, either (1) a Bachelor’s Degree or higher in the field of electrical engineering, computer engineering, or computer science and at least one year of related practical experience; or (2) at least four years of practical experience in data management. EX.1008 at ¶¶30-31.

E. Claim Construction

The claim terms at issue in the Challenged Claims have not been construed by any court.

Independent claims 1 and 34 both recite “...creating at least a second data miner in response to the customer request for information...”. During prosecution, in the Remarks dated March 29, 2005 and in the Appeal Brief dated July 5, 2005, Applicant expressly distinguished the prior art by arguing that “picking a pre-existing search engine” is an overly broad interpretation of the claim language. EX.1002 at 83 and 34. Applicant also expressly distinguished the prior art by arguing that “no new miners are created” in Chipman. *Id.* Applicant also argued that passing

the user search request to a pre-existing search engine is not the same as creating a data miner to explore that search request. *Id.* Applicant also stated that parsing and storing a search request is not the same as creating a data miner to explore that search request. *Id.* Based on the prosecution history, a POSA would understand the limitations of claims 1 and 34 to require creating a **new** data miner to search the results of a first data miner to provide customized results to the user based on the customized parameters of the user search request.⁷ EX.1008 at ¶49. This understanding is consistent with the Merriam-Webster dictionary definition of “creating” as “to make or bring into existence something new.” EX.1007 at 2.

Independent claim 21 recites “...tagging at least one representation of a Web page with at least one key...”. During prosecution, in the Remarks dated March 29, 2005 and in the Appeal Brief dated July 5, 2005, Applicant expressly stated that “tag” clearly connotes “append” and distinguished the prior art by arguing there was no representation of the Web page that was “tagged” in Chipman or Hegli. EX.1002 at 84 and 34-35. Applicant argued that assigning a category to a page is not the same

⁷ To the extent that Patent Owner considers the use of a pre-existing template to search the results of a first data miner to satisfy creating a second data miner, both Jacobsen and Seibel still anticipate this limitation as discussed below regarding independent claims 1 and 34. EX.1008 at ¶49, note 7.

as “tagging” a representation of a Web page as required by the claim. *Id.* Based on the prosecution history, a POSA would understand the limitation of claim 21 to require appending at least one representation of a Web page with at least one key. EX.1008 at ¶50.

Dependent claims 2 and 35 both recite “...billing the customer for responding to the customer request for information”. During prosecution, in the Appeal Brief dated July 5, 2005, Applicant expressly distinguished the prior art by stating that the Jaramillo reference “bills only by subscription, not by query” and that “[a]ccordingly, even if the references [Chipman and Jaramillo] were combined as proposed Claim 2 would not result.” EX.1002 at 36.⁸ This statement amounts to clear disclaimer of billing the customer on a subscription basis.⁹ EX.1008 at ¶51. Based

⁸ This is a factually inaccurate characterization of the Jaramillo reference as discussed below regarding dependent claims 2 and 35 in Ground 3. EX.1008 at ¶51, note 8.

⁹ Notably, with respect to Fig. 2, the specification states that “[t]he results are provided to the customer at block 70 and the customer is billed, on a per request basis or on a subscription basis.” EX.1001 at 8:54-56; EX.1008 at ¶51, note 9.

on the prosecution history, a POSA would understand the limitations of claims 2 and 35 to require billing the customer on a per request basis.¹⁰ *Id.*

IV. OVERVIEW OF PRIOR ART

A. Applicant Admitted Prior Art

The “Background” section of the ’526 Patent describes data mining by crawling the web. For example, “they access Web pages and pages to which the accessed pages hyperlink, generating an inverted index of words that occur on the Web pages.” EX.1001 at 1:20-23. The ’526 Patent further describes that “[q]ueries are responded to by accessing the index using the requested key words as entering arguments, and then returning from the index the URLs that satisfy the queries.” *Id.* at 1:25-28.

The ’526 Patent describes issues with crawling for information there “[d]espite the relevancy ranking used by most commercial search engines, finding particular types of information typically entails a great deal of mundane sifting through query results by a person.” *Id.* at 1:32-35. The ’526 Patent describes the

¹⁰ The limitations of claims 2 and 35 are rendered obvious regardless of whether Patent Owner agrees these prosecution history statements qualify as disclaimer. EX.1008 at ¶51, note 10. The grounds discussed below address billing per request or on a subscription basis.

needs for “cascading expert rules to sift through a large body of information.” *Id.* at 2:4-6. The cascading search may consist of “[e]liminating the useless pages [that] would require an expert to look at each page and determine whether it was the type of page that might contain information” and “[a]nother expert might then be required to examine the pages passed on from the first expert to determine if, in fact, the pages contained the answer to the specific question that was posed.” *Id.* at 1:64-2:4.

B. Jacobsen (EX.1003)

U.S. Pub. No. 2003/0163454 to Jacobsen et al. (“Jacobsen”) was filed on February 26, 2002 and published on August 28, 2003. EX.1003. Thus, Jacobsen qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

Jacobsen is analogous art to the ’526 Patent at least because Jacobsen and the ’526 Patent both relate to the searching and crawling the Internet. In addition, Jacobsen and the ’526 Patent were both classified in Intl. Cl. “G06F” relating to electrical data processing and to the U.S. Cl. “707” relating to data processing: database, data mining, and file management or data structures. EX.1001 at codes (51) and (52); EX.1003 at codes (51) and (52).

Jacobsen titled “Subject Specific Search Engine” is an invention described as “[a] subject-specific search engine [that] utilizes a smart web crawler and includes a capability of filtering out sites not relevant to the particular subject.”

EX.1003 at Abstract. Jacobsen is “directed to the general field of Systems, methods, and computer program products for performing Internet-based Searching.” EX.1003 at ¶[0002]. Jacobsen discloses problems including that (a) “[t]he only problem is to find the truly relevant resources. Search engines are what make the Internet useful, because without these tools the chances of finding relevant resources would be significantly diminished.” (*Id.* at ¶[0004]); (b) “current search engine technology often leaves one dissatisfied and frustrated, particularly where one would like to find resources on a given Subject in a Specific context” (*Id.* at ¶[0005]); and (c) “[a] general purpose Search engine (GPSE) will typically return numerous irrelevant links” (*id.*). Jacobsen addresses these problems by disclosing “a subject specific search engine (SSSE), where the strategy adopted is to collect and index only the pages deemed relevant for a specific subject, e.g., law or medicine.” *Id.* at ¶[0010].

Jacobsen describes the concept of separating the “straw” from the “chaff” like in the ’526 Patent, where the smart crawler is designed so as to filter each site it traverses and “[t]he web page corresponding to each site that is passed through the filter and deemed preliminarily relevant may *then be filtered one or more additional times.*” *Id.* at ¶[0025], emphasis added, and ¶[0075].

C. Seibel (EX.1004)

U.S. Patent No. 7,082,427 to Seibel et al. (“Seibel”) was filed on May 24, 2001, issued on July 25, 2006, and claims priority to U.S. Provisional application No. 60//206,772, filed on May 24, 2000. EX.1004. Thus, Seibel qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

Seibel is analogous art to the ’526 Patent at least because Seibel and the ’526 Patent both relate to the analysis and filtering of content on the Internet. In addition, Seibel and the ’526 Patent were both classified in Intl. Cl. “G06F” relating to electrical data processing and to the U.S. Cl. “707” relating to data processing: database, data mining, and file management or data structures. EX.1001 at codes (51) and (52); EX.1004 at codes (51) and (52).

Seibel titled “Text Indexing System to Index, Query the Archive Database Document by Keyword Data Representing the Content of the Documents and by Contact Data Associated with the Participant Who Generated the Document” is an invention described as “[a] text indexing system for collecting business intelligence ... [where] [t]he text indexing system has various components, including a data acquisition process that extracts textual data from various Internet sources, a database for storing the extracted data, a text indexing server that executes keyword searches of the database, and an output repository.” EX.1004 at Abstract. Seibel notes that “information is found in disparate locations and

formats” (id. at 1:53-54) and thus is directed to an “indexing system has various components, including a data acquisition process that extracts textual data from various Internet sources.” *Id.* at 2:12-14.

D. Jaramillo (EX.1005)

U.S. Publication No. 2002/0049773 to Jaramillo (“Jaramillo”) was filed on May 21, 2001 and published on April 25, 2002. EX.1005. Thus, Jaramillo qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

Jaramillo is analogous art to the ’526 Patent at least because Jaramillo and the ’526 Patent both relate to the searching and filtering of content on the Internet. In addition, Jaramillo and the ’526 Patent were both classified in Intl. Cl. “G06F” relating to electrical data processing and to the U.S. Cl. “707” relating to data processing: database, data mining, and file management or data structures. EX.1001 at codes (51) and (52); EX.1005 at codes (51) and (52).

Jaramillo titled “Method of Providing a Reference Standard Within a Web Site on an Internet” is describes a method that “includes the steps of providing a web site having a reference standard database on a computer network.... The method further includes the steps of searching the reference standard database for the reference standard using the search criteria and providing the user with the reference standard identified by the search of the reference standard database.” *Id.* at ¶[0012]. Jaramillo describes the advantages of the present invention as including

“an interactive search process is used to locate standards within the reference standard database using different search criteria, including name of company, product, service, keyword or language” (*id.* at ¶[0013]).

Jaramillo further describes that such services may be offered for a fee on a pay per use basis or on subscription basis. For example, “[o]ne example of a user 12 is an individual who accesses the web site 40 and pays a fee for the service provided by the web site 40 on a per-use basis” and “[a]nother example is for a user 12 that subscribes to the web site 40 for predetermined access to the reference standard services by paying a predetermined subscription fee.” *Id.* at ¶[0026].

E. Hegli (EX.1006)

U.S. Patent No. 6,606,659 to Hegli et al. (“Hegli”) was filed on January 28, 2000, and issued on August 12, 2003. EX.1006. Thus, Hegli qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

Hegli is analogous art to the ’526 Patent at least because Hegli and the ’526 Patent both relate to the searching and filtering of content on the Internet. In addition, Hegli and the ’526 Patent were both classified in Intl. Cl. “G06F” relating to electrical data processing. EX.1001 at codes (51); EX.1006 at codes (51).

Hegli titled “System and Method for Controlling Access to Internet Sites” is directed to a system that “includes a database of Internet sites that have been categorized so that the system determines the category of information that a user is

accessing.” EX.1006 at Abstract. Hegli discloses “analysis of lexical elements appearing on Internet sites that are strongly associated with a particular category” and “[i]n this context, a lexical element is a word or plurality of words that appear on the site under analysis.” *Id.* at 4:20-25.

V. ANALYSIS OF THE GROUNDS

A. GROUND 1: ANTICIPATION OF CLAIMS 1, 3, 19, 20 34, 36, 52, AND 53 BY JACOBSEN (EX.1003)

Jacobsen anticipates claims 1, 3, 19, 20, 34, 36, 52, and 53 of the ’526 patent. EX.1008 at ¶69.

1. Claim 1

a. “A method for data mining, comprising: ...”

To the extent the preamble is limiting, Jacobsen discloses this claim limitation. EX.1008 at ¶70. As stated in the ’526 patent, “[i]n a data mining system data is gathered into a data store using, e.g., a Web crawler.” EX.1001 at Abstract. Jacobsen discloses systems and methods where a “subject-specific search engine utilizes a smart web crawler and includes a capability of filtering out sites not relevant to the particular subject.” EX.1003 at Abstract; EX.1008 at ¶70.

b. “... *crawling the Internet to return Web pages to a data store; ...*”

Jacobsen discloses this claim limitation. EX.1008 at ¶71. For example, Jacobsen discloses smart crawler 16 operates in basically the same way as prior-art crawlers, *e.g.*, by visiting sites and transmitting information back to host computer 17, except that it does not operate under the “everything is relevant” principle. *Id.* at ¶[0045]. “[S]mart crawler 16 traverses the Internet 1 and then performs a screening operation, denoted by site filter 11. Site filter 11 determines, based on terminology of the profession to which the SSSE is directed (*e.g.*, law), whether or not each site is considered to be relevant to the profession.” *Id.* As one non-limiting example, Jacobsen describes the invention as being able to provide a particular profession (*e.g.*, law) with only relevant content, similar to a specialized law library as opposed to a regular library. *Id.* at ¶¶[0018]-[0019]; *see also, e.g.*, claim 1.

FIG. 6A, reproduced below, shows filtering operations according to the present invention. EX.1003 at ¶[0041]; EX.1008 at ¶72. Step S11 is an optional step of removing extraneous material, *e.g.*, advertisements. EX.1003 at ¶[0061]. At Step S12, the main part of the filter takes the page and compares it with a lexicon of terms (*e.g.*, legal terms) whose presence will indicate that the page may be

relevant. *Id.* If comparison is favorable in Step S13, then the page is saved in Step S15. *Id.*

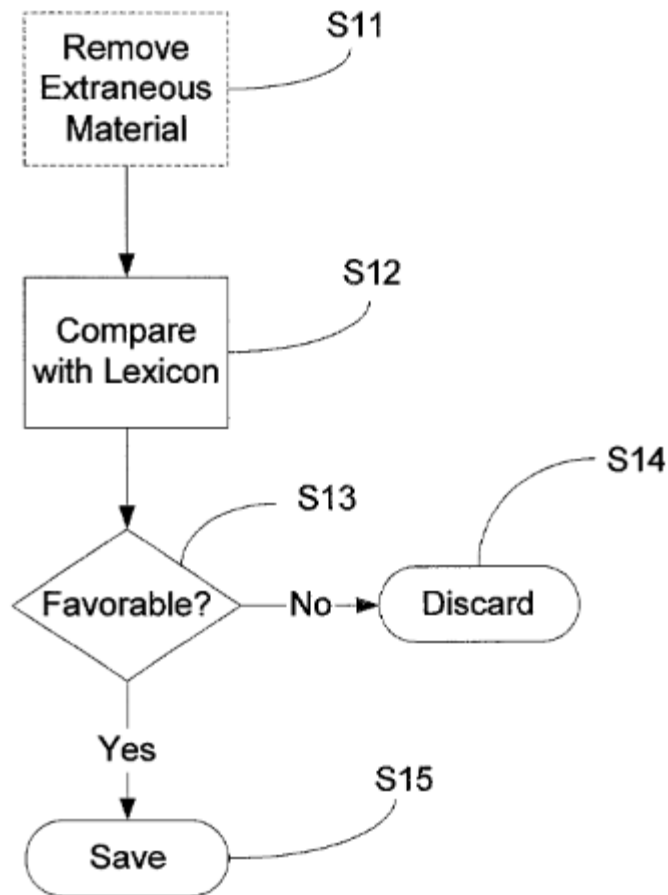


Figure 6A

As disclosed in Jacobsen, using a lexicon of legal terms would save relevant pages for a legal data store, similar to a specialized law library. EX.1008 at ¶73.

Jacobsen also discloses that the components of the filter itself may be the same for any context/profession, where only the lexicon may be changed. *Id.* For example, it is only necessary to change the lexicon to include medical terms, as opposed to

legal terms, to create a data store for a different profession, *e.g.*, medicine.

EX.1003 at ¶[0066]. This may be done for example by referencing a different file in a memory for each lexicon for each profession. *Id.*

c. “... *accessing the data store with at least a first data miner to process data in the store and output at least one result; ...*”

Jacobsen discloses this claim limitation. EX.1008 at ¶74. Jacobsen discloses a first data miner, *e.g.*, word scoring feature. EX.1008 at ¶75. This word scoring feature involves accessing the saved (relevant) web pages for processing and outputting at least one result, *e.g.*, word scores. *Id.*

As disclosed in Jacobsen, “[t]he basic idea of this technique is to assign numerical scores to words and to sum the scores of the words on a page to determine a score for the page.” EX.1003 at ¶[0068]. A word score is saved for each word on each page. EX.1003 at ¶[0069].

d. “... *receiving a customer request for information; ...*”

Jacobsen discloses this claim limitation. EX.1008 at ¶77. Jacobsen discloses receiving a customer request for information, *e.g.*, a user query. *Id.*

Jacobsen explains the process of crawlers where “[a]ny front-end query looks up the search terms in the information stored in the host’s database 9.” *Id.* at ¶[0007]. Jacobsen also discloses a user query, where its method includes

“permitting a user to enter a query; and searching the database for information according to the query.” *Id.* at claim 28.

e. “... creating at least a second data miner in response to the customer request for information, the second data miner accessing at least the result of the first data miner to generate an output; and ...”

Jacobsen discloses this claim limitation. Jacobsen discloses creating a second data miner, *e.g.*, site ranking feature, in response to the customer request for information. EX.1008 at ¶79. The second data miner, *e.g.*, site ranking feature, accesses the word scores, which correspond to the result of the first data miner, in order to generate an output. *Id.*

Jacobsen discloses that “[w]hen a user enters a query, the inventive SSSE determines a set of (relevant) pages that contain the query terms.” EX.1003 at ¶[0069]. “For each page, the word scores are summed for the words of the query to compute a site ranking for that page.” *Id.* “The site rankings for the pages are then used in determining how to display the search results to the user.” *Id.*

The site ranking feature is a customized feature based on the words in the user query. EX.1008 at ¶81. The site ranking feature provides users with ease in retrieving the most relevant information. *Id.* at ¶[0026]; EX.1008 at ¶81. The site ranking feature is used to search the results of the first data miner to provide

customized results to the user based on the user search request. EX.1008 at ¶81.

The site ranking feature is described as “dynamic” based on the word scores and in response to user queries. EX.1003 at ¶[0069]; EX.1008 at ¶81.

f. “... at least partially based on the output of the second data miner, responding to the customer request for information.”

Jacobsen discloses this claim limitation. EX.1008 at ¶82. Jacobsen discloses that the site ranking feature is used to determine an order in which to display the results. EX.1003 at ¶[0070]. “That is, the most relevant sites, as determined by their rankings, would be displayed earlier in the display and/or more prominently than less high ranking sites.” *Id.* This hierarchical display is meant to provide a user-friendly display of results. *See, e.g.,* EX.1003, FIGS. 8A and 8B. With respect to, for example, FIG. 8A, Jacobsen discloses that the user can click on the document 21 to access the information or site. EX.1003 at ¶[0071].

Thus, Jacobsen discloses all the limitations recited in claim 1. EX.1008 at ¶83.

2. Claim 3

Incorporating the above discussion of claim 1 in Ground 1, this additional claim limitation is disclosed in Jacobsen. EX.1008 at ¶84. As discussed above, the methods of Jacobsen may be provided to the user as a service and the acts of crawling, accessing, receiving, creating, and responding are undertaken at a search

facility. *See, e.g.*, EX.1003 at ¶[0008] and ¶[0044]. Jacobsen repeatedly analogizes the claimed invention to a specialized law library. *E.g.*, ¶¶[0018]-[0019]. Thus, Jacobsen discloses all the limitations of claim 3. EX.1008 at ¶85.

3. Claim 19

The above discussion of claim 1 in Ground 1 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having antecedent basis in claim 1, then this additional claim limitation is disclosed in Jacobsen. EX.1008 at ¶86.

As discussed above, Jacobsen discloses that the first data miner, *e.g.*, word scoring feature, uses a “word ranking scheme” to assign numerical scores to words. EX.1003 at ¶[0068]. “The technique works by examining each word (non-trivial word, i.e., not “stop words,” like “if,” “it,” “and,” and the like) on a given page and increasing its score if it appeared on a relevant page (i.e., a page that passed through filtering) containing a link to the given page.” *Id.* Additional examples include, increasing the word score according to how many relevant pages that linked to the given page contain the word and/or increasing the word score according to where the word appeared in a link leading to the page being examiner and/or increasing the word score if the word appears closer in proximity to the link to the page being examined. *Id.* Accordingly, Jacobsen discloses rules for

generating the result of the first data miner, *e.g.*, word scores. EX.1008 at ¶86.

Thus, Jacobsen discloses all the limitations of claim 19. EX.1008 at ¶87.

4. Claim 20

The above discussion of claim 1 in Ground 1 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having antecedent basis in claim 1, then this additional claim limitation is disclosed in Jacobsen. EX.1008 at ¶88.

As discussed above, Jacobsen discloses that the result of the first data miner, *e.g.*, word scores, are saved and later used by the second data miner. *E.g.*, EX.1003 at ¶¶[0068]-[0069]. Accordingly, the result of the first data miner is useful for generating an index of the data store, *e.g.*, pages with saved word scores. EX.1008 at ¶88. Thus, Jacobsen discloses all the limitations of claim 20. EX.1008 at ¶89.

5. Claim 34

a. “An information service, comprising: ...”

To the extent the preamble is limiting, Jacobsen discloses this claim limitation. EX.1008 at ¶90. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.a herein.

- b. *“... means for accessing, with at least a first data miner, a data store containing information representing at least one Web page to process data in the store and output at least one result; ...”*

Jacobsen discloses this claim limitation. EX.1008 at ¶¶91-95. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.b and V.A.1.c herein.

- c. *“... means for receiving a customer request for information; ...”*

Jacobsen discloses this claim limitation. EX.1008 at ¶96. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.d herein.

- d. *“... means for creating at least a second data miner in response to the customer request for information, the second data miner accessing at least the result of the first data miner to generate an output; and ...”*

Jacobsen discloses this claim limitation. EX.1008 at ¶¶97-99. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.e herein.

e. “... means for responding to the customer request for information at least partially based on the output of the second data miner.”

Jacobsen discloses this claim limitation. EX.1008 at ¶¶100-101. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.f herein. Thus, Jacobsen discloses all the limitations recited in claim 34.

6. Claim 36

Incorporating the above discussion of claim 34 in Ground 1, and for at least the reasons discussed with respect to claim 3 in Ground 1, Jacobsen discloses the additional limitations of claim 36. EX.1008 at ¶102.

7. Claim 52

The above discussion of claim 34 in Ground 1 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having antecedent basis in claim 34, then this additional claim limitation is disclosed in Jacobsen. EX.1008 at ¶103. Petitioner additionally incorporates by reference the above discussion of claim 19 in Section V.A.3 herein. For at least the reasons discussed with respect to claim 19, Jacobsen discloses the additional limitations of claim 52. EX.1008 at ¶103.

8. Claim 53

The above discussion of claims 34 and 52 in Ground 1 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having antecedent basis in claim 34, then this additional claim limitation is disclosed in Jacobsen. EX.1008 at ¶104. Petitioner additionally incorporates by reference the above discussion of claim 20 in Section V.A.4 herein. For at least the reasons discussed with respect to claim 20, Jacobsen discloses the additional limitations of claim 53. EX.1008 at ¶104.

B. GROUND 2: OBVIOUSNESS OF CLAIMS 2 AND 35 BASED ON JACOBSEN (EX.1003)

With Ground 1 as a predicate, the combination of Jacobsen and the general knowledge of a POSA renders obvious claims 2 and 35 of the '526 patent. EX.1008 at ¶105.

1. Claim 2

This claim limitation would be obvious to a POSA based the teachings of Jacobsen combined with his or her general knowledge. EX.1008 at ¶107.

The '526 Patent discloses: “[t]he results are provided to the customer at block 70 and the customer is billed, on a per request basis or on a subscription

basis.” EX.1001 at 8:54-56. As discussed above, the applicant’s statements during prosecution amount to clear disclaimer of billing the customer on a subscription basis. EX.1008 at ¶108. Regardless, Jacobsen in combination with a POSA’s general knowledge teaches the limitations of claim 2. *Id.*

Notably, the ‘526 Patent does not disclose any specific means or methods for billing the customer, which supports that it would have been within the general knowledge of a POSA to utilize known means or methods for doing so. EX.1008 at ¶109.

It would have been obvious for a POSA to incorporate the additional step of billing the user for responding to the user request for information on a per request basis into the methods of Jacobsen. EX.1008 at ¶110. For example, a POSA would have been motivated to bill the user for responding to the user request for information and effectuate payment prior to allowing the user to access the desired documents (*see e.g.*, EX.1003 at ¶¶[0071]-[0072]) in order to ensure payment. EX.1008 at ¶110. It would also have been obvious to bill the user for responding to the user request for information on a subscription basis, for example, if the user frequently used the system of Jacobsen. *Id.*

Billing the user for responding to the user request for information would allow Jacobsen’s services to make a profit in order to maintain the system so that it continues to provide services and products to users. EX.1008 at ¶111. This would

have been merely combining known methods (subject-specific search methods and known methods of billing customers) to yield predictable results (profit to continue provision of services). *Id.*

For at least these reasons, Jacobsen combined with the general knowledge of a POSA teaches the additional limitations of claim 2. EX.1008 at ¶112.

2. Claim 35

For at least the reasons discussed with respect to claim 2 in Ground 2, Jacobsen combined with the general knowledge of a POSA teaches the additional limitations of claim 35. EX.1008 at ¶114.

C. GROUND 3: OBVIOUSNESS OF CLAIMS 2 AND 35 BASED ON JACOBSEN (EX.1003) AND JARAMILLO (EX.1005)

With Ground 1 as a predicate, the combination of Jacobsen and Jaramillo renders obvious claims 2 and 35 of the '526 patent. EX.1008 at ¶115.

1. Motivations to Combine Regarding Jacobsen and Jaramillo

Jacobsen titled “Subject Specific Search Engine” is an invention described as “[a] subject-specific search engine [that] utilizes a smart web crawler and includes a capability of filtering out sites not relevant to the particular subject.” EX.1003 at Abstract; *see also* EX.1008 at ¶¶54-57.

Jaramillo titled “Method of Providing a Reference Standard Within a Web Site on an Internet” is an invention described as methods to facilitate searching of reference standards. EX.1005 at ¶[0022].¹¹ The methods also include “the steps of identifying if a user has permission to access the reference standard database on the web site and selecting a search criteria for a reference standard by the user, if the user has permission to access the reference standard database.” EX.1005 at Abstract.; *see also* EX.1008 at ¶¶61-64.

Jaramillo discloses it was known for search engines to use automated software programs that traverse the Internet cataloging documents and storing the information in a computer database, where it is filtered, categorized and stored. *Id.* at ¶[0006]. Jaramillo discloses it was known to charge a user to access a particular web site and for a user to pay a fee for the service provided by the web site on a per-use basis. *Id.* at ¶[0026]. Jaramillo also discloses it was known to charge a user

¹¹ Contrary to the applicant’s arguments during prosecution, Jaramillo relates to searching and accessing relevant documents, e.g., reference standards.

for a subscription to a website for predetermined access to a reference standard for paying a predetermined subscription fee. *Id.*¹²

A POSA would be motivated to combine the teachings of Jacobsen and Jaramillo to achieve the benefits both provide to Internet searching. EX.1008 at ¶¶116-120. A POSA would be motivated to improve the methods of Jacobsen with the billing methods of Jaramillo in order to ensure a user of Jacobsen has permission to use/access the subject-specific search engine services. *Id.* Further, billing the user would allow Jacobsen's services to make a profit in order to maintain the system so that it continues to provide services and products to users. *Id.* This would have been merely combining known methods (subject-specific search methods and known methods of billing customers) to yield predictable results (profit to continue provision of services). *Id.*

2. Claim 2

Billing the user for responding to the user request for information would have been obvious to a POSA based on the teachings of Jacobsen and Jaramillo discussed above. EX.1008 at ¶122.

¹² Contrary to the applicant's arguments during prosecution, Jaramillo discloses billing the customer as claimed at least because Jaramillo expressly discloses billing on a per-use basis, as well as billing by subscription.

The ‘526 Patent discloses: “[t]he results are provided to the customer at block 70 and the customer is billed, on a per request basis or on a subscription basis.” EX.1001 at 8:54-56. As discussed above, the applicant’s statements during prosecution amount to clear disclaimer of billing the customer on a subscription basis. EX.1008 at ¶123. Regardless, Jacobsen in combination with Jaramillo teaches the limitations of claim 2. *Id.*

The ‘526 Patent does not disclose any specific means or methods for billing the customer, which supports that it would have been within the general knowledge of a POSA to utilize known means or methods for doing so. EX.1008 at ¶124.

Jaramillo teaches billing the user for responding to the user request for information. EX.1008 at ¶125. For example, “[o]ne example of a user 12 is an individual who accesses the web site 40 and pays a fee for the service provided by the web site 40 on a per-use basis” and “[a]nother example is for a user 12 that subscribes to the web site 40 for predetermined access to the reference standard services by paying a predetermined subscription fee.” EX.1005 at ¶[0026].

It would have been obvious to a POSA to combine the teachings of Jacobsen and Jaramillo to add the step of billing the user for responding to the user request for information, as disclosed by Jaramillo, into the methods of Jacobsen. EX.1008 at ¶126. For example, like in Jaramillo, a POSA would have been motivated to bill the user for responding to the user request for information and effectuate payment

prior to allowing the user to access the desired documents (*see e.g.*, EX.1003 at ¶¶[0071]-[0072]) in order to ensure payment. EX.1008 at ¶126. Like in Jaramillo, billing the use could be done on a per request/use basis or on a subscription basis. *Id.*

For at least these reasons, Jacobsen combined with Jaramillo teaches the additional limitations of claim 2. *Id.*

3. Claim 35

For at least the reasons discussed with respect to claim 2 in Ground 3, Jacobsen combined with Jaramillo teaches the additional limitations of claim 35. EX.1008 at ¶¶127-128.

D. GROUND 4: OBVIOUSNESS OF CLAIMS 6-9, 11, 16-17, 39-42, 44, and 49-50 BASED ON JACOBSEN (EX.1003) AND HEGLI (EX.1006)

With Ground 1 as a predicate, the combination of Jacobsen and Hegli renders obvious claims 6-9, 11, 16-17, 39-42, 44, and 49-50 of the '526 patent. EX.1008 at ¶129.

1. Motivation to Combine Jacobsen and Hegli

Jacobsen titled “Subject Specific Search Engine” is an invention described as “[a] subject-specific search engine [that] utilizes a smart web crawler and includes a capability of filtering out sites not relevant to the particular subject.”

EX.1003 at Abstract. Jacobsen addresses these problems by disclosing a specific search engine (SSSE), where the strategy adopted is to collect and index only the pages deemed relevant for a specific subject, e.g., law or medicine.” *Id.* at ¶[0010]. The analysis disclosed by Jacobsen includes use of a lexicon of terms (e.g., legal terms) whose presence will indicate that the page may be relevant. *Id.* at ¶[0061]. The analysis disclosed by Jacobsen further includes use of a word scoring feature and a site ranking feature. *Id.* at ¶¶[0068]-[0069]; *see also* EX.1008 at ¶¶54-57.

Hegli titled “System and Method for Controlling Access to Internet Sites” is an invention described as “[a] method and system for providing flexible access to Internet sites.” EX.1006 at Abstract. Hegli is directed to a system that “includes a database of Internet sites that have been categorized so that the system determines the category of information that a user is accessing.” *Id.* Hegli discloses where “embodiments of this system include a training database that is created by analysis of lexical elements appearing on Internet sites that are strongly associated with a particular category” and “[i]n this context, a lexical element is a word or plurality of words that appear on the site under analysis.” *Id.* at 4:20-25; *see also* EX.1008 at ¶¶65-67.¹³

¹³ Contrary to the applicant’s arguments during prosecution, Hegli is related to a categorized database of Internet sites, including relevancy scores at least partially

A POSA would be motivated to combine the teachings of Jacobsen and Hegli to achieve the benefits both provide to Internet searches. EX.1008 at ¶¶130-133. A POSA would be motivated to improve the subject specific search engine of Jacobsen with the controlled and categorized searching of Hegli in order to improve functionality related to the collection, storage, and retrieval of information from the Internet. *Id.* A POSA reading Jacobsen would have been motivated to use the analysis of lexical elements and determination of categories of information as taught by Hegli. *Id.* For example, Jacobsen's analogy of a specialized law library would be further improved with Hegli's categorization methods in order to create sections of the library for particular subjects, *e.g.*, patent law. *Id.* Further, given that Jacobsen discloses different lexicons for different professions, a POSA would have been motivated to modify Jacobsen to allow certain customers to access the legal search results and other customers to access the medical search results depending on, for example, a customer's profession. *Id.* In some cases, a lawyer may need to also access medical search results and a doctor may need to also

defined by lexical terms, as well as controlling user access to categories. *See, e.g.*, EX.1006, claim 1. Hegli is not limited to the context of pornography. EX.1008 at ¶132, note 13. The '526 Patent mentions pornography, yet is not limited to the context of pornography. *Id.*

access legal search results, which would be provided for in a modified system of Jacobsen, based on the teachings of Hegli. *Id.* The combination of Jacobsen and Hegli would have been merely combining known known methods (search engines and lexical analysis) to yield predictable results (categorized search results). *Id.*

2. Claim 6

The combination of Jacobsen and Hegli teaches this claim limitation.

EX.1008 at ¶134. As discussed above, Jacobsen discloses word scores corresponding to the result output by the first data miner. EX.1008 at ¶135. It would have been obvious to incorporate the teachings of Hegli into Jacobsen such that the result output by the first data miner additional includes at least one key representing at least one characteristic (*e.g.*, word ID, category, relevance score, pair, adjacent) of at least one entity in the data store. EX.1008 at ¶¶135-138.

For example, Hegli teaches the use of analysis of lexical elements “that are strongly associated with a particular category.” EX.1006 at 4:21-23. Hegli teaches “[e]xamples of lexical elements include individual words, word pairs, adjacent words, and triplets of words.” *Id.* at 4:25-27. As discussed in Hegli, “the words, word pairs and word adjacencies from new Internet sites are tested against the category prototypes to determine if the new page should be categorized within any particular category.” *Id.* at 5:23-25. For example, the words “football team” would

be given a strong relevance score to the Sports category and a low relevance score to the Commerce category. *Id.* at 4:35-39; *see also id.* at 7:51-8:14.

With respect to Figure 3, reproduced below, Hegli teaches the use of word identification table 200 that includes lists of words and a corresponding ID for each word. EX.1006 at 8:31-33. Hegli further teaches the use of a category ID table 205 including the alpha score and beta score for each category. EX.1006 at 8:49-67. Hegli further teaches the use of the use of a word relevance table 210 that provides the relevance scores of word pairs and word adjacencies with particular categories in the system. EX.1006 at 9:1-20.

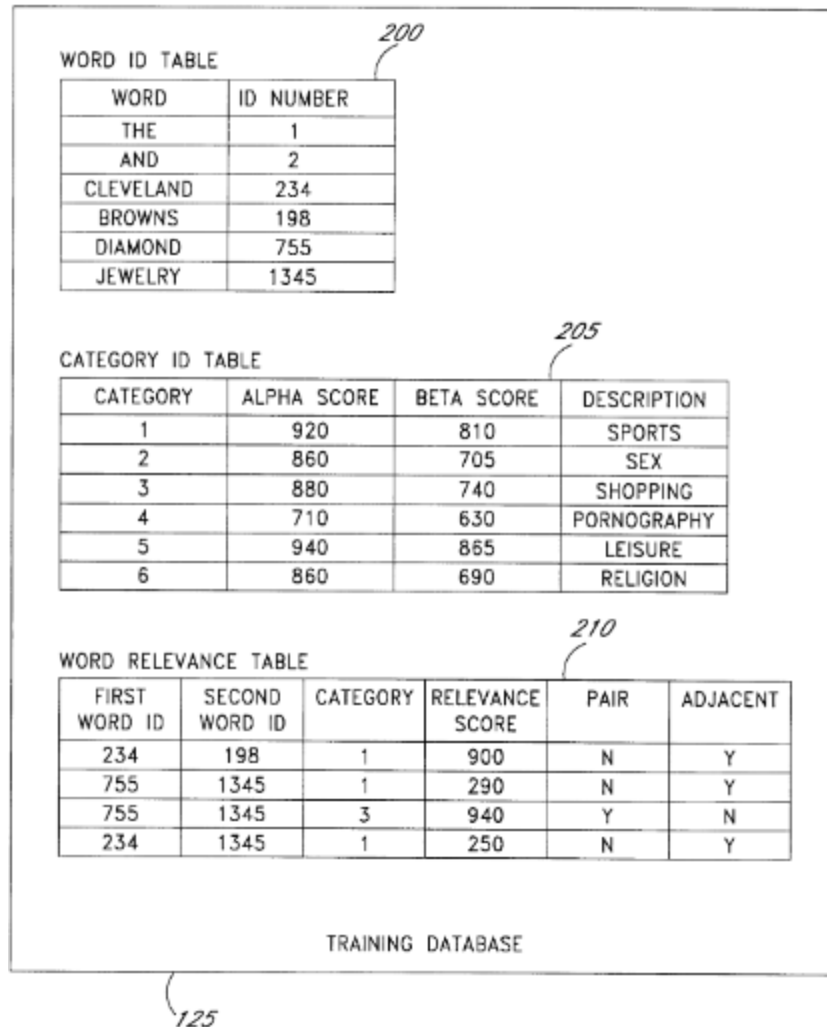


FIG. 3

Thus, the combination of Jacobsen and Hegli teaches the additional limitations of claim 6.¹⁴ EX.1008 at ¶¶135-138.

¹⁴ The examiner also relied on col. 4:21-29 and col. 9:13-20 of Hegli for teaching the limitations of claim 6, which was not disputed by the applicant during prosecution. The applicant only argued that a different primary prior art reference

3. Claim 7

Incorporating the above discussion of claim 6 in Ground 4, the combination of Jacobsen and Hegli teaches this claim limitation. EX.1008 at ¶¶139-141.

With respect to Fig. 3 of Hegli, shown above, Hegli teaches wherein the key is added to a table having therein plural entities, at least some entities having at least one attribute. EX.1008 at ¶140. For example, Hegli teaches the use of a category ID table 205. Hegli describes “category identification table 205 that provides a category ID number for each category within the system” where “table 205 also includes an alpha and beta score that provide the cut-off values for assigning a particular page to the selected category.” EX.1006. at 8:49-54.

Thus, the combination of Jacobsen and Hegli teaches the additional limitations of claim 7.¹⁵ EX.1008 at ¶141.

(Chipman) was not properly combinable with Hegli. The applicant did not argue that Hegli did not disclose the additional claim limitations.

¹⁵ The examiner also relied on Fig. 3 of Hegli for teaching the limitations of claim 7, which was not disputed by the applicant during prosecution. The applicant only argued that a different primary prior art reference (Chipman) was not properly combinable with Hegli. The applicant did not argue that Hegli did not disclose the additional claim limitations.

4. Claim 8

Incorporating the above discussion of claim 6 in Ground 4, the combination of Jacobsen and Hegli teaches this claim limitation. EX.1008 at ¶¶142-144.

With respect to Fig. 3 of Hegli, shown above, Hegli teaches wherein the value of the key is a binary value. EX.1008 at ¶143. For example, Hegli teaches the use of a word relevance table 210. The word relevance table 210 illustrates binary key values indicating whether words are adjacent, *i.e.*, “Y” or “N”, and whether the words are a pair, *i.e.*, “Y” or “N”.

Thus, the combination of Jacobsen and Hegli teaches the additional limitations of claim 8.¹⁶ EX.1008 at ¶¶142-144.

5. Claim 9

Incorporating the above discussion of claim 6 in Ground 4, the combination of Jacobsen and Hegli teaches this claim limitation. EX.1008 at ¶¶145-147.

¹⁶ The examiner also relied on Fig. 3 of Hegli for teaching the limitations of claim 8, which was not disputed by the applicant during prosecution. The applicant only argued that a different primary prior art reference (Chipman) was not properly combinable with Hegli. The applicant did not argue that Hegli did not disclose the additional claim limitations.

With respect to Fig. 3 of Hegli, shown above, Hegli teaches wherein the value of the key is a binary value. EX.1008 at ¶146. For example, Hegli teaches the use of a word relevance table 210. The word relevance table 210 illustrates non-binary key values indicating the relevance score.

Thus, the combination of Jacobsen and Hegli teaches the additional limitations of claim 9.¹⁷ EX.1008 at ¶¶145-147.

6. Claim 11

Incorporating the above discussion of claim 6 in Ground 4, the combination of Jacobsen and Hegli teaches this claim limitation. EX.1008 at ¶¶148-150.

For example, Hegli, in Fig. 6, illustrates “a process 328 of analyzing the word content of pages to determine their relevance to particular categories is illustrated.” *Id.* at 12:33-35. Hegli further discloses a quality rating of the relevance score where “the relevance score of a word pair within a particular category varies

¹⁷ The examiner also relied on Fig. 3 of Hegli for teaching the limitations of claim 9, which was not disputed by the applicant during prosecution. The applicant only argued that a different primary prior art reference (Chipman) was not properly combinable with Hegli. The applicant did not argue that Hegli did not disclose the additional claim limitations.

from 0 to 1,000, with 1,000 being a word pair that is perfectly associated with a category.” *Id.* at 12:52-54.

Thus, the combination of Jacobsen and Hegli teaches the additional limitations of claim 11.¹⁸ EX.1008 at ¶¶148-150.

7. Claim 16

Incorporating the above discussion of claim 6 in Ground 4, the combination of Jacobsen and Hegli teaches this claim limitation. EX.1008 at ¶¶151-153.

As discussed above, Jacobsen discloses the entity in the data store represents a Web page. *See, e.g.*, EX.1003 at ¶¶[0068]-[0069]. Jacobsen also discloses computing a word score associated with each word (except stop words) on each web page. *See, e.g.*, EX.1003 at claim 56. Also, for example, with respect to Fig. 8A, Jacobsen discloses the files 20, 22 represent the actual Web pages/sites. EX.1003 at ¶[0071]. As discussed above, Hegli additionally discusses the analysis of web pages in a data store. *See also, e.g.*, EX.1006 at 14:51-53 and 14:58-60.

¹⁸ The examiner also relied on Fig. 6 of Hegli for teaching the limitations of claim 11, which was not disputed by the applicant during prosecution. The applicant only argued that a different primary prior art reference (Chipman) was not properly combinable with Hegli. The applicant did not argue that Hegli did not disclose the additional claim limitations.

Thus, the combination of Jacobsen and Hegli teaches the additional limitations of claim 16. EX.1008 at ¶¶151-153.

8. Claim 17

Incorporating the above discussion of claim 6 in Ground 4, the combination of Jacobsen and Hegli teaches this claim limitation. EX.1008 at ¶¶154-156.

For example, with respect to Figs. 8A and 8B, Jacobsen discloses the “files or “site types” may consist of URLs of sites and contain the sites linked to from those URLs. EX.1003 at ¶[0073].

Thus, the combination of Jacobsen and Hegli teaches the additional limitations of claim 17. EX.1008 at ¶¶154-156.

9. Claim 39

For at least the reasons discussed with respect to claim 6 in Ground 4, Jacobsen combined with Hegli teaches the additional limitations of claim 39. EX.1008 at ¶157.

10. Claim 40

Incorporating the above discussion of claim 39 in Ground 4, and for at least the reasons discussed with respect to claim 7 in Ground 4, Jacobsen combined with Hegli teaches the additional limitations of claim 40. EX.1008 at ¶158.

11.Claim 41

Incorporating the above discussion of claim 39 in Ground 4, and for at least the reasons discussed with respect to claim 8 in Ground 4, Jacobsen combined with Hegli teaches the additional limitations of claim 41. EX.1008 at ¶159.

12.Claim 42

Incorporating the above discussion of claim 39 in Ground 4, and for at least the reasons discussed with respect to claim 9 in Ground 4, Jacobsen combined with Hegli teaches the additional limitations of claim 42. EX.1008 at ¶160.

13.Claim 44

Incorporating the above discussion of claim 39 in Ground 4, and for at least the reasons discussed with respect to claim 11 in Ground 4, Jacobsen combined with Hegli teaches the additional limitations of claim 44. EX.1008 at ¶161.

14.Claim 49

Incorporating the above discussion of claim 39 in Ground 4, and for at least the reasons discussed with respect to claim 16 in Ground 4, Jacobsen combined with Hegli teaches the additional limitations of claim 49. EX.1008 at ¶162.

15.Claim 50

Incorporating the above discussion of claim 39 in Ground 4, and for at least the reasons discussed with respect to claim 17 in Ground 4, Jacobsen combined with Hegli teaches the additional limitations of claim 50. EX.1008 at ¶163.

E. GROUND 5: ANTICIPATION OF CLAIMS 1, 6, 9, 11, 13, 19, 20, 34, 39, 42, 44, 46, 52, AND 53 BY SEIBEL (EX.1004)

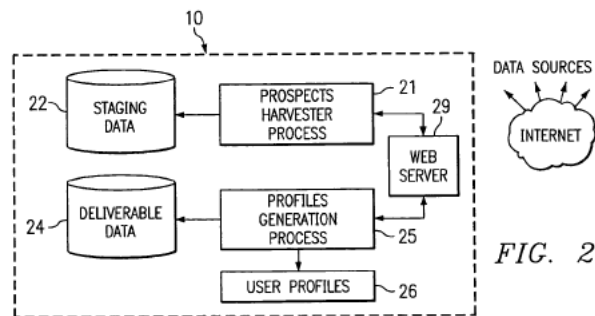
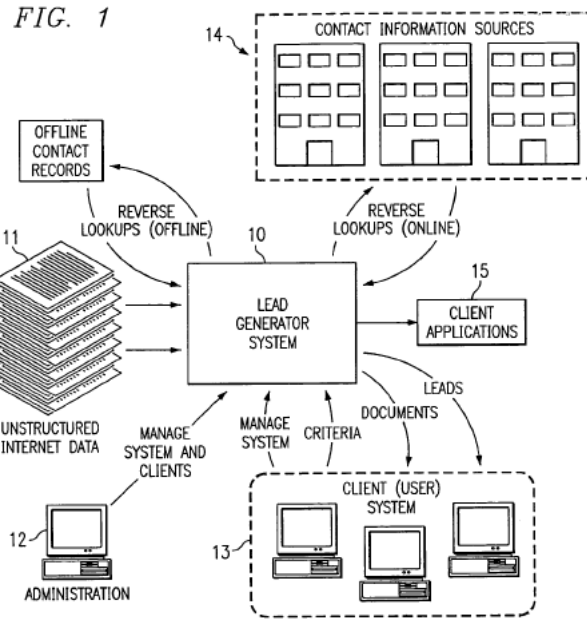
Seibel anticipates claims 1, 6, 9, 11, 13, 19, 20, 34, 39, 42, 44, 46, 52, and 53 of the '526 patent. EX.1008 at ¶164.

1. Claim 1

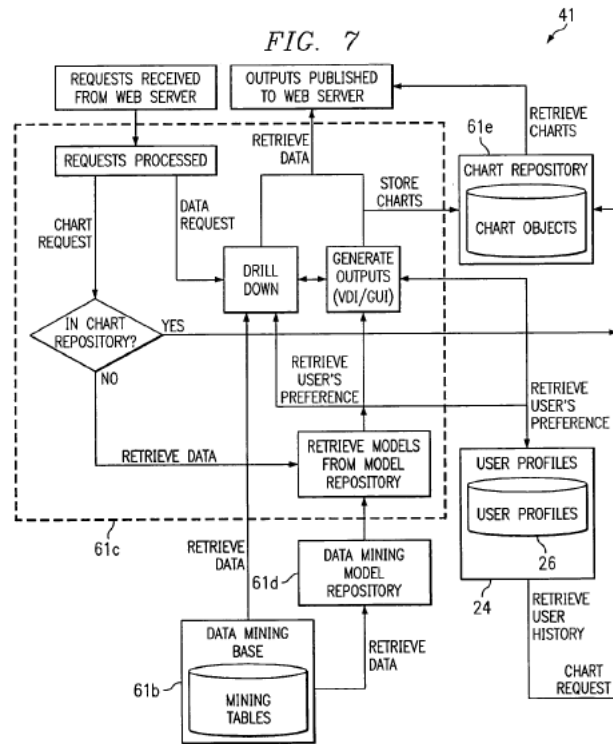
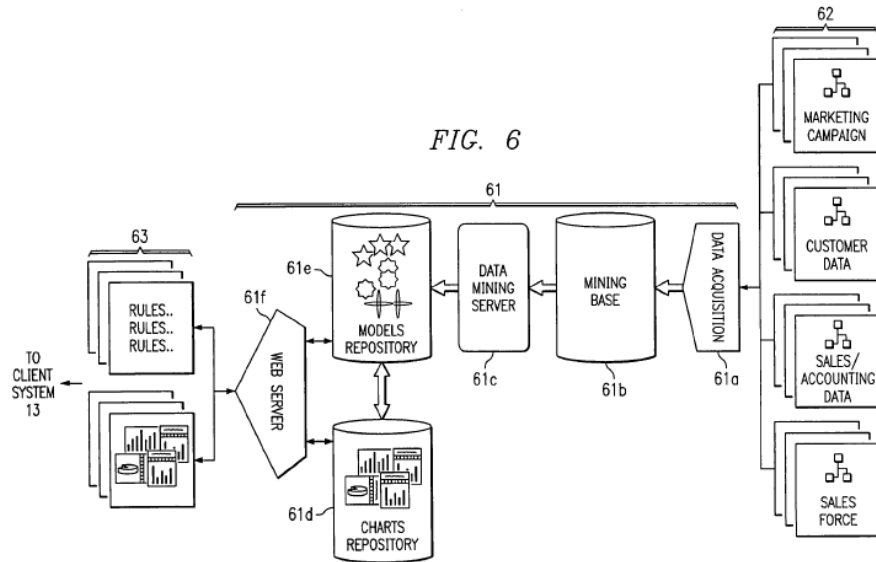
a. “A method for data mining, comprising: ...”

To the extent the preamble is limiting, Seibel discloses this claim limitation. EX.1008 at ¶165. As stated in the '526 patent, “[i]n a data mining system data is gathered into a data store using, *e.g.*, a Web crawler.” EX.1001 at Abstract.

Seibel describes a customer lead generation system 10 that uses the Internet and multiple data sources to mine text data to enhance businesses. EX.1004 at 2:47-57; EX.1008 at ¶¶166-168. Figs. 1 and 2 of Seibel are reproduced below:



Seibel discloses a method for data mining. EX.1008 at ¶169. For example, “FIGS. 6 and 7 illustrate a data mining system 61, which may be used within system 10 of FIGS. 1 and 2.” EX.1004 at 7:48-49. Figs. 6 and 7 of Seibel are reproduced below:



b. “... *crawling the Internet to return Web pages to a data store; ...*”

Seibel discloses this claim limitation. EX.1008 at ¶170. For example, Seibel discloses a lead generation system 10, shown in Figs. 1 and 2. The processes of the lead generation system 10 include crawling the Internet using a prospect harvester process 21 or crawler process 31. EX.1008 at ¶¶170-173.

Seibel describes “prospects harvester process 21 uses a combination of external data sources, client internal data sources and user parameter extraction interfaces, in conjunction with a search, recognition and retrieval system, to harvest contact information from the web and return it to a staging data base 22.” EX.1004 at 3:63-4:1.

Seibel further discloses that crawler process 31 is a text mining tool that provides access to data sources. *E.g., id.* at 4:63-65. Fig. 3, reproduced below, provides additional detail of the data sources of Figs. 1 and 2.

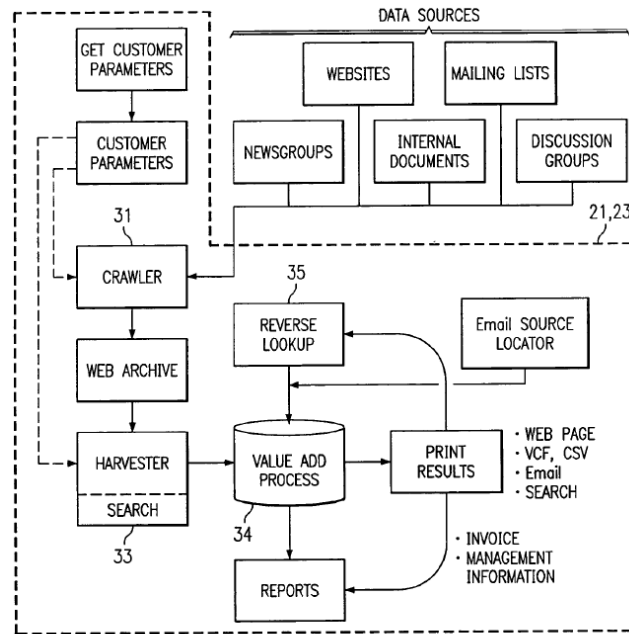
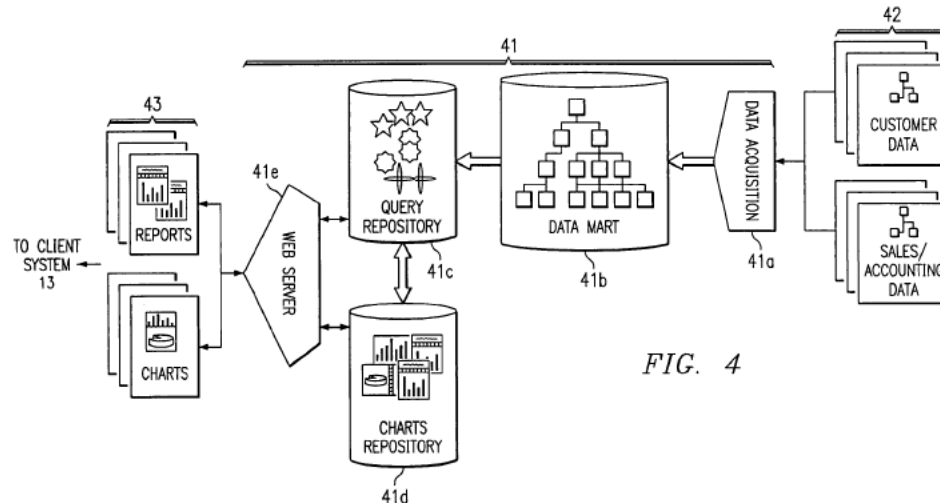


FIG. 3

Seibel further discloses that one such data source obtained from a crawler process are “web pages from Internet web sites” where “[t]his function of the crawler process is similar to the functionality associated with web-discussion-forums” and that “[s]earches are controlled by user-specified parameters such as web site URL, starting page, the number of traversal levels and crawling frequency.” *Id.* at 5:52-57.

Seibel further discloses a database server system 41, which may be used within lead generation system 10 of Figs. 1 and 2. EX.1008 at ¶174. Fig 4, reproduced below, shows database server system 41, including a data acquisition component 41a. Seibel discloses that “data acquisition component 41a collects and extracts data from different data sources, applying appropriate transformation,

aggregation and cleansing to the data collected. This component consists of predefined data conversions to accomplish most commonly used data transformations, for as many different types of data sources as possible.” EX.1004 at 6:12-18.



Thus, Seibel discloses crawling the Internet using processes 21, 31 to return Web pages to a data store, e.g., database server system 41. EX.1008 at ¶175.

c. “... *accessing the data store with at least a first data miner to process data in the store and output at least one result;*
...”

Seibel discloses this claim limitation. EX.1008 at ¶¶176-178. Seibel discloses a first data miner, e.g., profiles generation process 25, which processes customer data in the store to create a user (client) profiles database 26, which stores profiles of the client and its customers. *Id.*

For example, Seibel discloses that “system 41 could be used to implement the profiles generation process 25, which collects profile data about the [user/] client.” EX.1004 at 5:62-54. Seibel describes profiles generation process 25 which “analyzes the user’s in-house files and records related to the user's existing customers to identify and group those customers into profile categories based on the customer's buying patterns and purchasing volumes” and overlays the patterns and purchasing volumes of the existing customers on the salient contact information previously harvested to allow the aggregation of the revenue-based leads into prioritized demand generation sets. *Id.* at 4:23-39

Seibel further discloses that the “resulting profiles are then overlaid on the data obtained by [prospects harvester] process 21 to facilitate more accurate revenue projections and to enhance the sales and marketing process. The client may add certain value judgments (rankings) in a table that is linked to a unique lead id that can subsequently be analyzed by data mining or OLAP analytical tools.” *Id.* at 4:39-45.

d. “... receiving a customer request for information; ...”

Seibel discloses this claim limitation. EX.1008 at ¶¶179-181. Seibel discloses database server system 41, discussed above, which also includes a charts repository 41d. *Id.* at Fig. 4. Seibel discloses receiving a customer request for

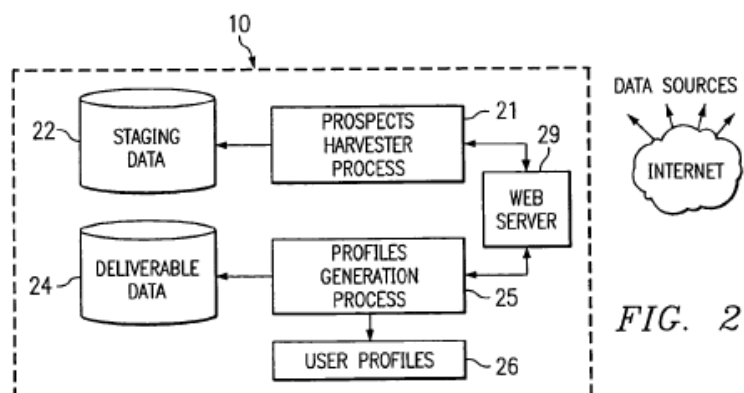
information which includes “chart requests” and “data requests”. EX.1004 at Fig. 5 and 6:58-7:2.

e. “... creating at least a second data miner in response to the customer request for information, the second data miner accessing at least the result of the first data miner to generate an output; ...”

Seibel discloses this claim limitation. EX.1008 at ¶¶182-186. Seibel discloses creating at least a second data miner in response to the customer request for information. For example, Seibel discloses a second data miner, *e.g.*, server 41e runs a query using different parameters (*e.g.*, a quarterly revenue query where the parameters passed are the specific region, the specific year, and the specific quarter) to extract data from data mart 41b to feed the data to the requested chart. *Id.*; EX.1004 at 6:44-7:2. Seibel further discloses that users may request charts by regions, by quarters, by products, by companies, and by chart style (*e.g.*, pie chart, bar chart, map view, or data view). *Id.*; EX.1004 at 7:29-39.

Server 41e extracts (*e.g.*, mines) data in response to a user (*e.g.*, customer) request. EX.1008 at ¶183. This is a customized extraction based on the parameters of the user request. *Id.* This provides customized results to the user based on the user request. *Id.*

Seibel discloses the second data miner accessing at least the result of the first data miner to generate an output. EX.1008 at ¶184. For example, profiles generation process 25 (*e.g.*, the first data miner) generates resulting deliverable data 24 and user profiles 26, as seen in Fig. 2:



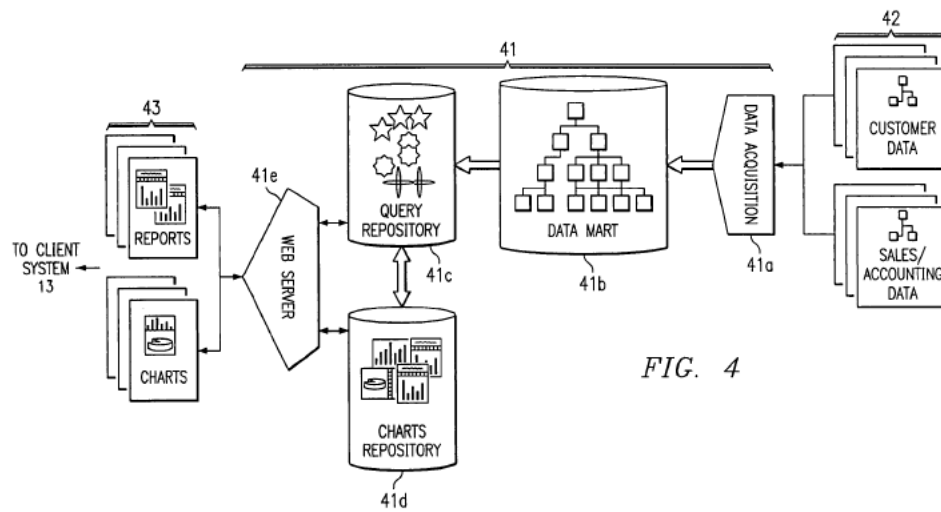
Seibel further discloses multiple other data miners, including server 41e, accessing those results to generate an output: “As explained above in connection with FIG. 2, the profiles generation process 25 may be used to generate a profiles database 26. This [profiles] database 26 stores information about the client and its customers that may be used to better identify prospective customers. ... as illustrated in FIG. 5, the database server 41e of database server system 41 may access database 24 [(which contains Profiles Database 26 - Fig. 5)] to determine user preferences in formulating queries and presenting outputs...” EX.1004 at 10:60-11:8; *see also id.* at Fig. 9; EX.1008 at ¶185.

Seibel further discloses that server 41e is distinct from profiles generation process 25 (*e.g.*, a first data miner). EX.1008 at ¶186. For example, Fig. 5 shows

that server 41e accesses user profiles and data 24 and 26, which, as discussed above, were already separately generated by profiles generation process 25.

f. “... at least partially based on the output of the second data miner, responding to the customer request for information.”

Seibel discloses this claim limitation. EX.1008 at ¶¶187-190. For example, Seibel discloses that server 41e generates output 43, which is sent to client system 13, as seen in Fig. 4, reproduced below:



Seibel further discloses a response to a customer request for information based at least partially on output 43. EX.1008 at ¶189. For example, Seibel discloses that output 43 can comprise business intelligence tailored to specific customer requests. EX.1004 at 7:29-39.

Thus, Seibel discloses all the limitations of claim 1.

2. Claim 6

Incorporating the above discussion of claim 1 in Ground 5, Seibel discloses this claim limitation. EX.1008 at ¶¶191-193. For example, Seibel discloses the profiles generation process 25 creates profiles of the client and its customers, e.g., the result output. *Id.* Seibel further discloses that “profiles include items such as purchasing criteria, buying cycles and trends, cross-selling and up-selling opportunities, and effort to expense/revenue correlations.” *Id.* at 4:37-39. Seibel further discloses profiles include data obtained by process 21, such as the prospect’s name, email address, Web site address, and other contact information. *Id.* at 4:9-22. Seibel further discloses that a client may add certain value judgments (rankings) in a table that is linked to a unique lead id for later analysis. *Id.* at 4:42-45. Thus, Seibel discloses the additional limitations of claim 6.

3. Claim 9

Incorporating the above discussion of claims 1 and 6 in Ground 5, Seibel discloses this claim limitation. EX.1008 at ¶¶194-196. As discussed above with respect to claim 6, as one example, the value of the key (*e.g.*, purchasing criteria) is a non-binary value. As another example, the value of the key (*e.g.*, prospect’s name) is a non-binary value. Thus, Seibel discloses the additional limitations of claim 9.

4. Claim 11

Incorporating the above discussion of claims 1 and 6 in Ground 5, Seibel discloses this claim limitation. EX.1008 at ¶¶197-199. As discussed above with respect to claim 6, as one example, Seibel discloses the client may add certain value judgments (rankings) to show that the profile is associated with a high value contact. *Id.* at 4: 42-44. Seibel also discloses that the contact data provided by process 21 may include a related document to verify the high level value of the contact. *Id.* at 4:18-22. Thus, Seibel discloses the additional limitations of claim 11.

5. Claim 13

Incorporating the above discussion of claims 1 and 6 in Ground 5, Seibel discloses this claim limitation. EX.1008 at ¶¶200-202. As discussed above with respect to claim 6, as one example, the value of the key (e.g., person's name) represents whether the associated entity represents a person. Thus, Seibel discloses the additional limitations of claim 13.

6. Claim 19

The above discussion of claim 1 in Ground 5 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having

antecedent basis in claim 1, then this additional claim limitation is disclosed in Seibel. EX.1008 at ¶¶203-205.

For example, Seibel discloses the profiles generation process 25 creates profiles of the client and its customers, e.g., the result output. EX.1008 at ¶204. Seibel further discloses that the process 25 may include rules to identify and group customers into profile categories based on the customer's buying patterns and purchasing volumes. *Id.* at 4:23-27. Process 25 may also include rules to aggregate leads into prioritized demand generation sets. *Id.* at 4:27-31. Process 25 may also be used to create profiles of the most profitable customers. *Id.* at 4:34-37. Thus, Seibel discloses the additional limitations of claim 19.

7. Claim 20

The above discussion of claim 1 in Ground 5 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having antecedent basis in claim 1, then this additional claim limitation is disclosed in Seibel. EX.1008 at ¶¶206-208. For example, Seibel discloses the profiles generation process 25 creates profiles of the client and its customers, e.g., the result of the first data miner. The resultant profiles are useful for generating at least one index (e.g., profiles database 26) which may be accessed during various data and

text mining processes to better identify prospective customers of the client. *Id.* at 4:47-52. Thus, Seibel discloses all the limitations of claim 20.

8. Claim 34

a. *“An information service, comprising: ...”*

To the extent the preamble is limiting, Seibel discloses this claim limitation. EX.1008 at ¶209. Petitioner incorporates by reference the above discussion of claim 1 in Section V.E.1.a herein.

b. *“... means for accessing, with at least a first data miner, a data store containing information representing at least one Web page to process data in the store and output at least one result; ...”*

Seibel discloses this claim limitation. EX.1008 at ¶¶210-215. Petitioner incorporates by reference the above discussion of claim 1 in Section V.E.1.b and V.E.1.c herein.

c. *“... means for receiving a customer request for information; ...”*

Seibel discloses this claim limitation. EX.1008 at ¶¶216-217. Petitioner incorporates by reference the above discussion of claim 1 in Section V.E.1.d herein.

- d. “... means for creating at least a second data miner in response to the customer request for information, the second data miner accessing at least the result of the first data miner to generate an output; ...”

Seibel discloses this claim limitation. EX.1008 at ¶¶218-223. Petitioner incorporates by reference the above discussion of claim 1 in Section V.E.1.e herein.

- e. “... means for responding to the customer request for information at least partially based on the output of the second data miner.”

Seibel discloses this claim limitation. EX.1008 at ¶¶224-228. Petitioner incorporates by reference the above discussion of claim 1 in Section V.E.1.f herein. Thus, Seibel discloses all the limitations of claim 34.

9. Claim 39

Incorporating the above discussion of claim 34 in Ground 5, and for at least the reasons discussed with respect to claim 6 in Ground 5, Seibel discloses the additional limitations of claim 39. EX.1008 at ¶229.

10.Claim 42

Incorporating the above discussion of claims 34 and 39 in Ground 5, and for at least the reasons discussed with respect to claim 9 in Ground 5, Seibel discloses the additional limitations of claim 42. EX.1008 at ¶230.

11.Claim 44

Incorporating the above discussion of claims 34 and 39 in Ground 5, and for at least the reasons discussed with respect to claim 11 in Ground 5, Seibel discloses the additional limitations of claim 44. EX.1008 at ¶231.

12.Claim 46

Incorporating the above discussion of claims 34 and 39 in Ground 5, and for at least the reasons discussed with respect to claim 13 in Ground 5, Seibel discloses the additional limitations of claim 46. EX.1008 at ¶232.

13.Claim 52

The above discussion of claim 34 in Ground 5 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having antecedent basis in claim 34, then this additional claim limitation is disclosed in Seibel for at least the reasons discussed with respect to claim 19 in Ground 5. EX.1008 at ¶233.

14.Claim 53

The above discussion of claims 34 and 42 in Ground 5 is incorporated herein. It is unclear which data miner this claim refers to, but if the PTAB determines this claim refers to the first data miner because this claim refers to “the result” having antecedent basis in claim 34, then this additional claim limitation is disclosed in Seibel for at least the reasons discussed above with respect to claim 20 in Ground 5. EX.1008 at ¶234.

F. GROUND 6: OBVIOUSNESS OF CLAIMS 2 AND 35 BASED ON SEIBEL (EX.1004)

With Ground 5 as a predicate, the combination of Seibel and the general knowledge of a POSA renders obvious claims 2 and 35 of the ‘526 Patent. EX.1008 at ¶235.

1. Claim 2

This claim limitation would be obvious to a POSA based the teachings of Seibel combined with his or her general knowledge. EX.1008 at ¶237.

The ‘526 Patent discloses: “[t]he results are provided to the customer at block 70 and the customer is billed, on a per request basis or on a subscription basis.” EX.1001 at 8:54-56. As discussed above, the applicant’s statements during prosecution amount to clear disclaimer of billing the customer on a subscription

basis. EX.1008 at ¶238. Regardless, Seibel in combination with a POSA's general knowledge teaches the limitations of claim 2. *Id.*

Notably, the '526 Patent does not disclose any specific means or methods for billing the customer, which supports that it would have been within the general knowledge of a POSA to utilize known means or methods for doing so. EX.1008 at ¶239.

It would have been obvious for a POSA to incorporate the additional step of billing the user for responding to the user request for information on a per request basis into the methods of Seibel. EX.1008 at ¶240. For example, Seibel expressly discloses administrative system 14 manages user accounts and billing. EX.1004 at 3:47-52. A POSA would have been motivated to bill the user for responding to the user request for information, *e.g.*, bill the user per chart. EX.1008 at ¶240. It would also have been obvious to bill the user for responding to the user request for information on a subscription basis, for example, if the user frequently used the system of Seibel. *Id.*

Billing the user for responding to the user request for information would allow Seibel's services to make a profit in order to maintain the system so that it continues to provide services and products to users. EX.1008 at ¶241. This would have been merely combining known methods (data mining and known methods of

billing customers) to yield predictable results (profit to continue provision of services). *Id.*

For at least these reasons, Seibel combined with the general knowledge of a POSA teaches the additional limitations of claim 2. EX.1008 at ¶242.

2. Claim 35

For at least the reasons discussed with respect to claim 2 in Ground 6, Seibel combined with the general knowledge of a POSA teaches the additional limitations of claim 35. EX.1008 at ¶¶243-244.

VI. DISCRETIONARY DENIAL WOULD NOT BE APPROPRIATE

The factors considered under 35 U.S.C. §§314 and 325 do not weigh in favor of exercising discretion to deny institution.

A. The factors under §314 weigh against discretionary denial

As noted above, the ‘526 Patent is involved in a related district court case, *Bright Data Ltd. v. Oxylabs, UAB*, No. 2:23-cv-171 (E.D. Tex.) (“Bright Data Litigation”). The *Fintiv* factors weigh against discretionary denial. *See Apple Inc. v. Fintiv, Inc.*, IPR2020-00019, Paper 11 at 5-16 (PTAB Mar. 20, 2020) (“*Fintiv*”); *see also* Director Katherine K. Vidal, ‘Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation’ (June 21, 2022) (“Vidal Memorandum”).

In addition to *Fintiv* factors 1, 3 and 4, which weigh against discretionary denial, the petition presents compelling evidence of unpatentability and that determination alone demonstrates that the Board should not exercise discretionary denial. See Vidal Memorandum at 4-5.

***Fintiv* Factor 1:**

At the time of filing this petition, neither party has moved to stay the Bright Data Litigation. Determining whether Judge Gilstrap would grant a motion to stay the Bright Data Litigation is highly speculative. Accordingly, this factor weighs against discretionary denial.

***Fintiv* Factor 2:**

The Docket Control Order (Dkt. 28) sets jury selection for December 9, 2024. EX.1013.

Based on the most recent statistics (dated 6/30/2023), the median time-to-trial for civil cases in E.D. Tex. is 16.7 months. EX.1014 at 35. The case was

opened on April 14, 2023.¹⁹ Thus, the trial would statistically occur around September 2024.²⁰

The final written decision in this proceeding is expected around March 2025. Accordingly, this factor weighs in favor of discretionary denial. However, Director Vidal has instructed that “when other relevant factors weigh against exercising discretion to deny institution or are neutral, the proximity to trial should not alone outweigh all of those other factors.” Vidal Memorandum at 8.

Fintiv Factor 3:

There have been no substantive proceedings in the Bright Data Litigation, which weighs strongly against discretionary denial.

¹⁹ Bright Data’s counterclaims for infringement of U.S. Patent No. 11,206,317 were initially filed in Case No. 2:22-cv-11 (E.D. Tex.) on 2/14/2023. The court severed the counterclaims into a new action, referred to herein as the Bright Data Litigation. See EX.1015. Oxylabs’ counterclaims for infringement of the ‘526 Patent and U.S. Patent No. 7,020,667 were filed on 4/28/2023. See EX.1016.

²⁰ The median time-to-trial for civil cases in E.D. Tex. in 2022 was 24.5 months, which would put the trial date around April 2025.

Petitioner has been diligent and filed this petition while the Bright Data Litigation is still in its infancy. *See Fintiv* at 11. Additionally, Petitioner was not aware of the prior art presented in the petition prior to the Bright Data Litigation. Petitioner uncovered the prior art by conducting a reasonably diligent search in anticipation of serving invalidity contentions in the Bright Data Litigation.

Pursuant to the Docket Control Order, the Claim Construction Hearing is scheduled for 6/25/2024, the deadline to complete fact discovery is 7/22/2024, the deadline to complete expert discovery is 8/26/2024, the deadline to file dispositive motions is 9/3/2024, and the pretrial conference is scheduled for 11/4/2024. *See generally* EX.1013.

An institution decision in this proceeding is expected around March 2024. By that time, the parties will not yet have filed any claim construction briefing and fact discovery will still be ongoing. No substantive orders will have issued by the court. Accordingly, this factor weighs strongly against discretionary denial.

***Fintiv* Factor 4:**

There are 54 claims in the ‘526 Patent.

Claims 1-3, 6, 8, 9, 16, and 19 of the ‘526 Patent are at issue in the Bright Data Litigation. That totals 8 asserted claims, with only claim 1 being independent.

Claims 1-3, 6-9, 11, 13, 16-17, 19-20, 34-36, 39-42, 44, 46, 49-50, and 52-53 are challenged in this petition. That totals 26 challenged claims, with claims 1 and 34 being independent.

Accordingly, this factor weighs strongly against discretionary denial.

Additionally, Petitioner hereby stipulates that, if instituted, the petitioner will not pursue the same grounds presented in this petition against the '526 Patent in the Bright Data Litigation. This stipulation further weighs against discretionary denial.

***Fintiv* Factor 5:**

The parties are the same in the Bright Data Litigation. As such, this factor weighs in favor of discretionary denial.

***Fintiv* Factor 6:**

On balance, the *Fintiv* factors weigh against discretionary denial. Additionally, this petition presents compelling evidence of unpatentability. Institution would strengthen the patent system and serve the public interest. See Vidal Memorandum at 5.

B. The factors under §325 weigh against discretionary denial

This petition relies on primary references Jacobsen (EX.1003) and Seibel (EX.1004) which were not previously considered by the Office. These primary prior art references and arguments are not the same or substantially the same as those previously presented during prosecution of the ‘526 Patent. *See Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6 at 7-9 (PTAB Feb. 13, 2020)(“*Advanced Bionics*”). Additionally, the Office did not have the benefit of the expert testimony of Dr. Williams regarding the arguments set forth in the petition.

As to certain dependent claims, this petition relies on the combination of Jacobsen and Jaramillo (EX.1005) in Ground 3 and the combination of Jacobsen and Hegli (EX.1006) in Ground 4. These particular combinations were not considered by the Office during prosecution of the ‘526 Patent. As such, the Office has not considered the same or substantially the same art and arguments in Grounds 3 and 4. *See Advanced Bionics* at 7-9; *see also Thorne Res., Inc. v. Trs. of Dartmouth Coll.*, IPR2021-00491, Paper 18 at 8 (PTAB Aug. 12, 2021)(“The fact that one piece of art from the combination was previously presented and/or argued to the Office alone is insufficient to satisfy the first prong of *Advanced Bionics* two-part test...”). The new arguments in the petition warrant consideration by the Board.

The Office previously considered a **different** primary reference, Chipman (EX.1010), in combination with Jaramillo (EX.1005) and in combination with Hegli (EX.1006). The Office considered the combination of Chipman and Jaramillo with respect to challenged claims 2 and 35. EX.1002 at 94. The Office considered the combination of Chipman and Hegli with respect to challenged claims 6-9, 11, 16-17, 39-42, 44, and 49-50. EX.1002 at 95.

During prosecution, the applicant argued: (1) that Chipman did not disclose “...creating at least a second data miner in response to the customer request for information...” with respect to challenged claims 1 and 34 (EX.1002 at 34 and 83); (2) that there was no proper motivation to combine Chipman and Jaramillo with respect to challenged claims 2 and 35 (EX.1002 at 36); (3) that Jaramillo did not disclose billing a customer on a per request basis, as opposed to a subscription basis, with respect to challenged claims 2 and 35 (EX.1002 at 36); and (4) that there was no proper motivation to combine Chipman and Hegli (EX.1002 at 36 and 84).

The examiner ultimately accepted the applicant’s arguments to point (1) regarding the independent claims and the disclosure of Chipman. EX.1002 at 10.

With respect to point (1), the petition establishes that Jacobsen discloses “...creating at least a second data miner in response to the customer request for information...” regarding challenged claims 1 and 34. *See also* EX.1008 at ¶¶79-81

and ¶¶97-99. The Jacobsen reference and arguments regarding the Jacobsen reference were not previously considered by the examiner. Additionally, this “creating” limitation was the sole basis on which the applicant distinguished independent claims 1 and 34 from Chipman during prosecution. The examiner had initially responded that Chipman’s parsing and storing of a search request satisfied this “creating” limitation. EX.1002 at 50 (citing EX.1010 at ¶[0063]); *see also* EX.1002 at 34. The disclosure of Jacobsen relied on in the petition is not substantially the same as the disclosure of Chipman relied on by the examiner. Unlike Chipman, Jacobsen expressly discloses a customized site ranking feature based on the words in the user query. *E.g.*, EX.1008 at ¶¶79-81. If the examiner had considered Jacobsen, as opposed to Chipman, then independent claims 1 and 34 would not have been allowed.

With respect to point (2), the petition establishes that there is proper motivation to combine Jacobsen and Jaramillo. As one non-limiting example, both Jacobsen and Jaramillo disclose cataloguing/storing documents in a computer database. *See* EX.1008 at ¶¶116-120. These arguments were not previously considered by the examiner.

With respect to point (3), the petition establishes that, contrary to the applicant’s argument, Jaramillo does indeed disclose billing on a per request basis as well as on a subscription basis. EX.1005 at ¶[0026]. The examiner did not

substantively respond to the applicant's argument and only allowed claims 2 and 35 based on their respective dependency from claims 1 and 34. EX.1002 at 10.

With respect to point (4), the petition establishes that there is proper motivation to combine Jacobsen and Hegli. As one non-limiting example, both Jacobsen and Hegli disclose analysis of lexical elements in order to filter/classify web pages. See EX.1008 at ¶¶130-133. These arguments were not previously considered by the examiner. Additionally, the applicant did not argue that Hegli does not disclose the additional limitations recited in challenged dependent claims 6-9, 11, 16-17, 39-42, 44, and 49-50.

At minimum, the new primary reference Jacobsen, relied on in this petition, warrants consideration by the Board. See also *Becton, Dickinson & Co. v. B. Braun Melsungen AG*, IPR2017-01586, Paper 8 at 17-18 (PTAB Dec. 15, 2017). Jacobsen was not previously considered by the Office.

VII. CONCLUSION

Petitioner respectfully requests cancellation of the Challenged Claims of the '526 Patent.

Respectfully submitted,

Date: September 27, 2023

By: /s/ Thomas M. Dunham

Thomas M. Dunham

Reg. No. 39,965

Cherian LLP

1901 L Street NW, Suite 700

Washington, D.C. 20036

(202) 838-1567

ATTORNEY FOR PETITIONER,
BRIGHT DATA LTD.

CERTIFICATE OF COMPLIANCE WITH TYPE-VOLUME LIMITS

This petition consists of 13,809 words, which is no more than 14,000 words, excluding table of contents, table of authorities, table of exhibits, mandatory notices, certificate of service, and this certificate. The petition complies with the type-volume as mandated in 37 C.F.R. §42.24. In preparing this certificate, counsel has relied on the word count of the word-processing system used to prepare the paper (Microsoft Word).

Respectfully submitted,

Date: September 27, 2023

By: /s/ Thomas M. Dunham
Thomas M. Dunham
Reg. No. 39,965

Cherian LLP
1901 L Street NW, Suite 700
Washington, D.C. 20036
(202) 838-1567

ATTORNEY FOR PETITIONER,
BRIGHT DATA LTD.

CERTIFICATE OF SERVICE

Pursuant to 37 C.F.R. §42.6(e) and 37 C.F.R. §42.105(a), the undersigned hereby certifies the PETITION FOR *INTER PARTES* REVIEW and all exhibits thereto were served on the undersigned date via USPS Priority Mail Express®, at the following addresses:

Patent File Correspondence Address:

John L. Rogitz
Rogitz & Associates
750 B Street, Suite 3120
San Diego, CA 92101

Assignee Address:

Metacluster LT UAB
A. Gostauto G. 40A
Vilnius, Lithuania LT-03163

Courtesy copies were also served via e-mail to Litigation Counsel at the following addresses:

scallahan@ccrglaw.com

cbovenkamp@ccrglaw.com

ctolliver@ccrglaw.com

jscott@ccrglaw.com

msibley@ccrglaw.com

jheuton@ccrglaw.com

brett.govett@nortonrosefulbright.com

daniel.leventhal@nortonrosefulbright.com

Respectfully submitted,

Date: September 27, 2023

By: /s/ Thomas M. Dunham
Thomas M. Dunham
Reg. No. 39,965

Cherian LLP
1901 L Street NW, Suite 700
Washington, D.C. 20036
(202) 838-1567

ATTORNEY FOR PETITIONER,
BRIGHT DATA LTD.