

UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

BRIGHT DATA LTD.,

Petitioner

v.

METACLUSTER LT, UAB,

Patent Owner

Case IPR2023-01442

Patent No. 7,020,667

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. 7,020,667
EX. 1002	Prosecution History of U.S. Patent No. 7,020,667
EX. 1003	Declaration of Dr. Tim Arthur Williams
EX. 1004	U.S. Patent No. 6,826,553 to DaCosta et al.
EX. 1005	U.S. Patent Publication No. 2003/0018669 to Kraft
EX. 1006	Appendix of Challenged Claims
EX. 1007	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. 1008	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. 1009	Order (Dkt. 129) in the case of <i>Metacluster LT, UAB v. Bright Data Ltd.</i> , No. 2:22-cv-11 (E.D. Tex. April 14, 2023)
EX. 1010	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)	69
<i>Apple Inc. v. Fintiv, Inc.</i> , IPR2020-00019, Paper 11 (PTAB Mar. 20, 2020) ..	70, 72
<i>General Plastic Co., Ltd. v. Canon Kabushiki Kaisha</i> , IPR2016-01357, Paper 19 (PTAB Sept. 6, 2017)	69

Statutes

35 U.S.C. § 314	1, 69
35 U.S.C. § 325	69

Other Authorities

Director Katherine K. Vidal, ‘Interim Procedure for Discretionary Denials in AIA Post-Grant Proceedings with Parallel District Court Litigation’ (June 21, 2022)	70, 71, 73
--	------------

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 '667 PATENT.....	4
A. Summary of the '667 Patent.....	4
B. Claims of the '667 Patent	6
C. Prosecution History of the '667 Patent.....	7
D. Level of Skill in the Art.....	9
E. Claim Construction.....	9
IV. OVERVIEW OF PRIOR ART REFERENCES	10
A. Applicant Admitted Prior Art.....	10
B. DaCosta (EX.1004)	11
C. Kraft (EX.1005).....	13
V. ANALYSIS OF THE GROUNDS.....	14
A. Ground 1: Anticipation of Claims 1-20 by DaCosta (EX.1004).....	14
1. Claim 1	14
2. Claim 2	26
3. Claim 3	27
4. Claim 4	28
5. Claim 5	29
6. Claim 6	30

7. Claim 7	30
8. Claim 8	31
9. Claim 9	33
10. Claim 10	34
11. Claim 11	35
12. Claim 12	35
13. Claim 13	36
14. Claim 14	37
15. Claim 15	38
16. Claim 16	38
17. Claim 17	41
18. Claim 18	41
19. Claim 19	42
20. Claim 20	42
B. Ground 2: Anticipation of Claims 1-11 and 13-20 by Kraft.....	44
1. Claim 1	44
2. Claim 2	56
3. Claim 3	57
4. Claim 4	57
5. Claim 5	58
6. Claim 6	58
7. Claim 7	59
8. Claim 8	60
9. Claim 9	60
10. Claim 10	61
11. Claim 11	61
12. Claim 13	62
13. Claim 14	63
14. Claim 15	63
15. Claim 16	64
16. Claim 17	66
17. Claim 18	66
18. Claim 19	66
19. Claim 20	67
VI. DISCRETIONARY FACTORS	69

VII. CONCLUSION.....	73
-----------------------------	-----------

I. INTRODUCTION

Petitioner Bright Data Ltd. respectfully requests *inter partes* review and cancellation of claims 1-20 (“Challenged Claims”) of Patent No. 7,020,667 (EX.1001, the “’667 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.1003).

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. 7,020,667 titled “System and Method for Data Retrieval and Collection in a Structured Format.”

1. Real Parties-in-Interest

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

2. Related Matters

There is one related district court case involving the ’667 Patent, styled as *Bright Data Ltd. v. Oxylabs*, UAB, Case No. 2-23-cv-00171 (E.D. Tex.). That case

also involves U.S. Patent No. 7,010,526 assigned to Patent Owner and U.S. Patent No. 11,206,317 assigned to Petitioner.

There are no related administrative matters involving the '667 Patent.

3. Lead and Back-Up Counsel

Lead Counsel	Back-Up Counsel
<p>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</p>	<p>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</p> <p>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</p>

4. Service Information

Electronic Mail	<ol style="list-style-type: none"> 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 '667 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 1-20	§102 Anticipation by DaCosta
Ground 2	Claims 1-11 and 13-20	§102 Anticipation by Kraft

III. OVERVIEW OF THE '667 PATENT

The '667 Patent issued from U.S. Application No. 10/198,827 filed on July 18, 2002. EX.1001, codes (21) and (22).

A. Summary of the '667 Patent

The '667 Patent titled “System and Method for Data Retrieval and Collection in a Structured Format” relates to data collection and retrieval from various Web pages into a database through text parsing. EX.1001, Abstract; EX.1003, ¶32. A user interface allows the user to select the source file that contains data the user wishes to copy, target database that will receive the data, and timing criteria for automatic transfer of data. EX.1001, Abstract; EX.1003, ¶32.

The '667 Patent discloses a data collection and retrieval system that “allows the user to select one or more Web pages containing data of interest; to specify exactly which data within any page is to be captured; to specify how frequently data is to be collected; and to specify the conditions for collection and retrieval”. EX.1001, 2:19-24; EX.1003, ¶33. The system and method include an initialization state and an automatic execution stage in which the initialization stage provides a user interface for the user to select a source file that contains data the user wishes to

copy, a target database to receive the data, and timing criteria for automatically transferring the data. *Id.* at 2:29-36. The automatic execution stage automatically transfers the data from the HTML source file to the target database. EX.1001, 2:38-40; EX.1003, ¶33.

“[T]he system 10 accesses Web pages that have been selected by the user” where “[t]hese Web pages are the source of the data to be transferred to a user defined database.” EX.1001, 4:16-18; EX.1003, ¶34. The ’667 patent describes selecting the source file where “the user opens system 10 in step 405, and the system 10 displays the user input GUI 315” and “[i]n step 410, the user specifies the URL for an HTML page 220 that contains data to be transcribed to the target database 370” where “[i]n so doing, the user selects the Web page 220 as shown by FIG. 2, and saves the Web page URL in system 10.” EX.1001, 4:45-50; EX.1003, ¶34.

The ’667 patent describes parsing the source file where “[i]n step 420 the parser 415 processes the source file 320, tokenizes the HTML format of the Web page 220 and places it in a data array 330” where “[i]deally, the parser 325 divides the data in the source into identifiable units of the Web page 220, exactly replicating the Web page 220.” EX.1001, 4:57-62; EX.1003, ¶35. Further, “[t]he user then selects data and or categories of interest in step 430” and “[i]n step 435, system 10 saves the sequence of data or categories selected by the user to an array locations file 345” where “[s]ystem 10 remembers the location of the data in the array 330 and

retrieves data from the same locations on each successive data transfer.” EX.1001, 5:12-16; EX.1003, ¶35.

The ’667 patent describes setting a schedule for accessing source location where “the user sets the timer 350 by selecting the frequency of data retrieval; and where “[t]he user then specifies in step 445 where to store the retrieved data, such as database 370.” EX.1001, 5:17-25; EX.1003, ¶36.

B. Claims of the ’667 Patent

The ’667 Patent includes 20 claims, with claims 1, 16, and 20 being independent. This Petition challenges claims 1-20. A Claims Appendix is provided in EX.1006.

Exemplary independent claim 1 recites a method as described below:

1. A computer-implemented method of selectively retrieving data from a source file on a public network and automatically collecting retrieved data in a target database, comprising:
 - initializing a data retrieval stage by:
 - selecting the source file;
 - parsing selected data elements in the source file into a list of data elements;
 - selecting at least one of the data elements from the list to be stored in the target database;
 - identifying source locations of the selected at least one of the data elements;
 - selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database;

saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database;
setting a schedule frequency for transferring the at least one of the parsed data elements from the source locations to the target locations in the location file in the target database; and
automatically, selectively parsing new data elements corresponding to the at least one of the parsed data elements previously selected during the initialization of the data retrieval stage, pursuant to the schedule frequency, by transferring the selectively parsed new data elements from the saved source locations in the source file to the saved target locations in the target database.

C. Prosecution History of the '667 Patent

The underlying application that issued as the '667 Patent was filed on July 18, 2002 with 40 claims and titled as “System and Method for Data Retrieval and Collection in a Structured Format.” EX.1002, 154; *see also* EX.1003, ¶¶37-42.

On February 11, 2005, the examiner issued a Non-final Rejection rejecting claims 6-9, 22-25, 32, 34, and 35 under 35 U.S.C. §112 as having insufficient antecedent basis. EX.1002, 154, 126-127. The examiner also rejected claims 1, 3-10, 13-15, 17-28, and 30-36 as being anticipated by U.S. Patent No. 5,532,473 to Niazi et al. (“Niazi”). EX.1002, 154, 127; EX.1003, ¶38. The examiner also rejected claim 2 as being obvious over Niazi in view of U.S. Publication No. 2002/0184170 to Gilbert et al. (“Gilbert”); claims 16 and 29 as being obvious over Niazi in view of U.S. Patent No. 6,636,845 to Chau et al. (“Chau”); and claims 11 and 12 as being

obvious over Niazi in view of U.S. Patent No. 6,532,473 to Bright (“Bright”). EX.1002, 154, 130-133.

On April 23, 2005, Applicant filed a response amending claims 1, 3, 4, 7-10, 14-17, 19, 21, 23-25, 28, 31, 33-35, and 38; canceling claims 2, 5-6, 12, 18, 22, and 32; and adding claim 41. EX.1002, 97-105. Applicant argued that the amendments overcame the rejections under §112 and distinguished Niazi at least because Niazi does not describe the two-stage method and does not select the data elements to be parsed. EX.1002, 106-118.

On July 26, 2005, the examiner issued a Final Rejection rejecting claims 1, 3-4, 7-10, 13-15, 17, 19-21, 23-28, 30-31, 33-38, and 40-41 as being obvious over Niazi in view of M2 Presswire. EX.1002, 81-85. The examiner further rejected claim 16 as being obvious over Niazi in view of M2 Presswire and further in view of Chau. EX.1002, 85. The examiner further rejected claims 11, 29, and 39 as being obvious over Niazi in view of M2 Presswire and further in view of Bright. EX.1002, 85-86.

On September 2, 2005, Applicant filed a request for continued examination and amended independent claims 1, 21, and 31, to recite “saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database.” EX.1002, 59. Applicant also further cancelled claims 26-30 and 33-41. EX.1002, 58-76.

On October 31, 2005, the Examiner issued a Notice of Allowance. EX.1002, 2. The notice also included various Examiner amendments directed to grammatical corrections. EX.1002, 6.

D. Level of Skill in the Art

The relevant technology for the '667 Patent pertains to data retrieval of web sites using crawlers across networks. A person of ordinary skill in the art ("POSA") would have been an individual who, as of July 18, 2002, the filing date of the '667 Patent, had 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.1003, ¶¶30-31.

E. Claim Construction

The claim terms of the '667 Patent have not been construed by any tribunal. Petitioner gives the claim terms their plain and ordinary meaning in light of the specification and prosecution history as would have been read by a POSA.¹ *See also* EX.1003, ¶49.

¹ Petitioner reserves the right to submit proposed claim constructions if construed by a court and/or the Board.

IV. **OVERVIEW OF PRIOR ART REFERENCES**

A. Applicant Admitted Prior Art

The “Background” section of the ’667 Patent states that “there are currently no effective methods for capturing and replicating that data within a database application.” EX.1001, 1:34-36; EX.1003, ¶51. But, the ’667 patent then states that “[t]he user can always manually transfer the data viewed on the Internet into a database.” EX.1001, 1:36-37; EX.1003, ¶51.

The ’667 patent describes the then current state-of-the-art where “[c]urrent systems exist for transferring Internet data to databases for archiving Web content in the form of a source file and visual representation, allowing subsequent viewing of that data at a later date” but that “[w]hile allowing the user to capture data and archive it, these systems do not provide a mechanism for parsing individual data elements, replicating that data in a relational database format, or automatic data collection in a manner specified by the user.” EX.1001, 1:45-52; EX.1003, ¶52.

The ’667 patent background section further describes a system that “retrieves data from a wide range of document formats and converts that data into a common format and location” where “[t]his system includes the capability to collect data automatically in a manner specified by the user” and “[i]t also allows the retrieval of data from a wide range of data formats, including the capability to retrieve data from HTML documents.” EX.1001, 1:54-60; EX.1003, ¶53.

B. DaCosta (EX.1004)

U.S. Patent No. 6,826,553 to DaCosta et al. (“DaCosta,” EX.1004) was filed on November 16, 2000 and issued on November 30, 2004. The ’667 Patent was filed on July 18, 2002 at code (22) and issued on March 28, 2006, at code (45). Thus, DaCosta qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

DaCosta is analogous art to the ’667 Patent at least because DaCosta and the ’667 Patent both relate to the retrieval and manipulation of data through the Internet. EX.1003, ¶55. In addition, DaCosta and the ’667 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, codes (51) and (52); EX.1004, codes (51) and (52); EX.1003, ¶55.

DaCosta is titled “System for Providing Database Functions for Multiple Internet Sources.” EX.1004, code (54). DaCosta presents a system for automatically extracting data from at least one electronic document accessible through the Internet or other computer network. EX.1004, Abstract, code (57) ; EX.1003, ¶56. DaCosta further states that “[t]he present invention relates to acquisition of data and, more particularly, to web browsers for the Internet, as well as to database utilities for data accessible through the Internet” and “to navigate to one or more data sources on the

Internet preferably in an automated manner, extract data irrespective of the format of that data and display, store and/or process the extracted data.” EX.1004, 1:22-29. Further, DaCosta discloses that the “extracted data can be stored in a conventional repository, including Microsoft Exchange 2000 or any other database that constitutes a knowledge repository.” EX.1004, 18:18-21; EX.1003, ¶56.

DaCosta discloses “an architecture for a smart web server to scour data from the Internet for deposit in a data repository.” EX.1004, 5:1-3. “[D]ata feeds Web Folders that comprise a knowledge repository accessible by smart clients.” EX.1004, 18:15-17. DaCosta discloses the use of “navigation and extraction scripts (collectively referred to as Navex) [that] can automatically extract numerous records 260 that match a pattern.” EX.1004, 11:20-22; EX.1003, ¶57.

DaCosta further discloses that “[i]n a preferred embodiment of the present invention, scripts are stored at a centralized repository that is accessible through the Internet.” EX.1004, 18:33-35. DaCosta also discloses that repositories may exist as conventional repositories “including Microsoft Exchange 2000 or any other database that constitutes a knowledge repository for XML tagged data, so that the user can readily access the stored data needed in conjunction with other business processes.” EX.1004, 18:17-23; EX.1003, ¶58.

C. Kraft (EX.1005)

U.S. Publication No. 2003/0018669 to Kraft (“Kraft,” EX.1005) was filed on April 2, 2001 and published on January 23, 2003. The ’667 Patent was filed on July 18, 2002 at code (22) and issued on March 28, 2006, at code (45). Thus, Kraft qualifies as prior art under pre-AIA 35 U.S.C. § 102(e).

Kraft is analogous art to the ’667 Patent at least because Kraft and the ’667 Patent both relate to Internet-based data retrieval and data storage. EX.1003, ¶60. In addition, Kraft and the ’667 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, codes (51) and (52); EX.1005 at codes (51) and (52); EX.1003, ¶60.

Kraft is titled “System and Method for Associating a Destination Document to a Source Document During a Save Process.” EX.1005, code (54). Kraft presents a system and method for use with a web browser and the Internet to save the location and other context data along with the content of the web page when being stored. EX.1005, Abstract, code (57); EX.1003, ¶61. Kraft further discloses a crawler that is run periodically to update previously stored data. EX.1005, ¶[0003].; EX.1003, ¶61. Further, Kraft discloses that the “invention relates to a system for saving the content of a target document bundled with contextual metadata, such as the location

of a source document, as attributes of the target document.” EX.1005, ¶[0001]; EX.1003, ¶61.

V. ANALYSIS OF THE GROUNDS

A. Ground 1: Anticipation of Claims 1-20 by DaCosta (EX.1004)

1. Claim 1

- a. *“A computer-implemented method of selectively retrieving data from a source file on a public network and automatically collecting retrieved data in a target database, comprising: ...”*

To the extent the preamble is limiting, DaCosta discloses this claim limitation. EX.1003, ¶64. DaCosta presents a system for automatically extracting data from at least one electronic document accessible through the Internet or other computer networks. EX.1004, Abstract, code (57). DaCosta further states that “[t]he present invention relates to acquisition of data and, more particularly, to web browsers for the Internet, as well as to database utilities for data accessible through the Internet” and “to navigate to one or more data sources on the Internet preferably in an automated manner, extract data irrespective of the format of that data and display, store and/or process the extracted data.” EX.1004, 1:22-29; EX.1003, ¶64. Further, DaCosta discloses that the “extracted data can be stored in a conventional repository, including Microsoft Exchange 2000 or any other database that constitutes a knowledge repository.” EX.1004, 18:18-21; EX.1003, ¶64.

b. “... initializing a data retrieval stage by: ...”

DaCosta discloses this claim limitation. EX.1003, ¶65. For example, DaCosta discloses initializing a data retrieval stage by defining a navigation path through the Internet. For example, DaCosta discloses “Referring now to FIG. 8, a user needs only to start recording using the navigation recording module 12 and ‘surf’ as usual to a desired destination URL using Internet Explorer, for example” and “[u]pon reaching the destination URL, the user indicates the same to the recording module 12” where “[t]he user can interact with the recording module 12 using buttons 230, for example.” EX.1004, 11:5-11; EX.1003, ¶65.

Further, initializing a data retrieval stage is also shown by performing the steps outlined below. EX.1003, ¶66.

c. “... selecting the source file;...”

DaCosta discloses this claim limitation. EX.1003, ¶67. For example, DaCosta discloses that “[t]he system of the present invention performs the tasks of precisely locating and extracting the select data with a granularity specified by the user from any information source such as search engine results, web pages, other web-accessible documents, e-mail or text feeds in any format, for example, HTML, .txt, pdf, Word, Excel, ppt, ftp text feeds, databases, XML and other standard, as well as non-standard, formats.” EX.1004, 2:31-38; EX.1003, ¶68.

DaCosta further discloses selecting the source file where “[t]he extraction module also has the capability to scrape or harvest the data from the source that is identified by the location procedure so that data can be imported.” EX.1004, 3:21-24; EX.1003, ¶69.

DaCosta further discloses that “[t]he system of the present invention preferably comprises a navigation module that accesses one or more web pages or other web-accessible documents” and “[t]he navigation module provides the capability for a user to specify and store a procedure such as a series of clicks and entries of information, for example, a user name and password, to access a web page or other web-accessible document, as well as the capability to perform the procedure to actually access the web page or other web-accessible document in an automated manner.” EX.1004, 2:55-64; EX.1003, ¶70.

d. “... parsing selected data elements in the source file into a list of data elements;...”

DaCosta discloses this claim limitation. EX.1003, ¶71. For example, DaCosta discloses, as shown in FIG. 4, reproduced below, parsing HTML and text into lists. Specifically, “[r]eferring again to FIG. 4, after an HTML web page, for example, or other web-accessible document is loaded 140, the page is parsed 150 to produce 160 the HTML hierarchy.” EX.1004, 11:53-55; EX.1003, ¶71.

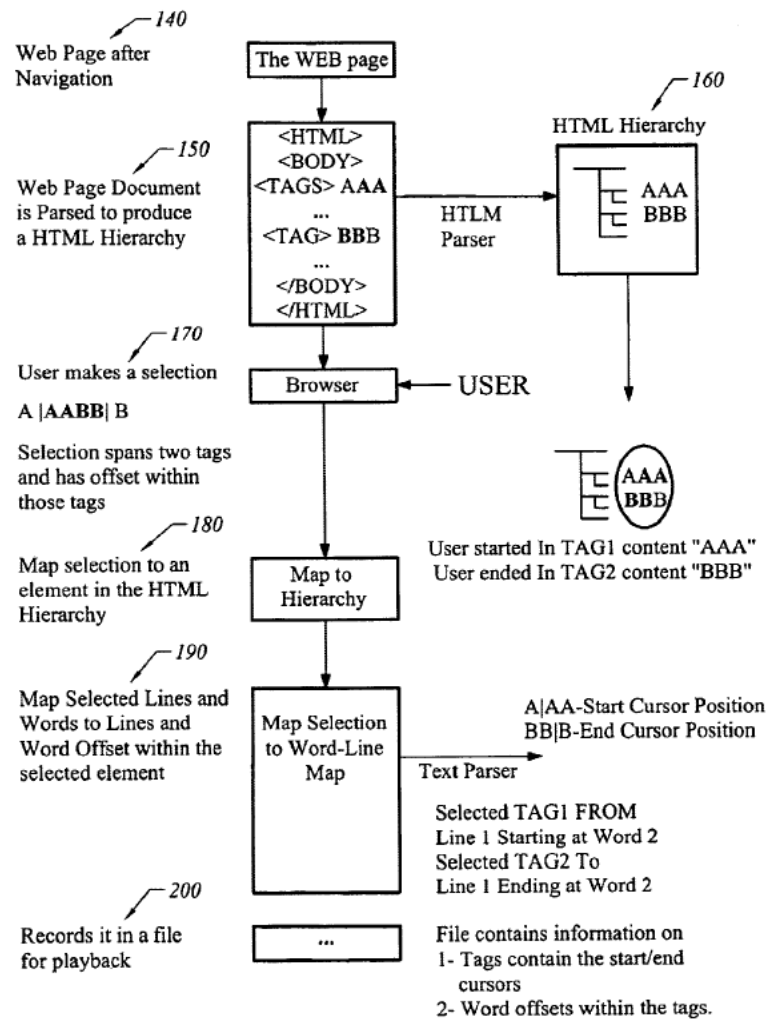


FIG. 4

Further, DaCosta describes receiving the data and parsing it into lists. For example, “when an HTML web page or other web-accessible document is requested, the page is transmitted from the server to an intermediate receptor as a data stream, a well-known example being Winsock” and that “[t]he data stream is separated and arranged into a plurality of elements configured in a well-defined hierarchy so they

may be used to build the actual document displayed to the user by Internet Explorer” and where “[h]ierarchy objects or elements include lists, tables, images and columns, for example.” EX.1004, 7:2-10; EX.1003, ¶72.

e. “... selecting at least one of the data elements from the list to be stored in the target database ...”

DaCosta discloses this claim limitation. EX.1003, ¶73. DaCosta discloses that the process of extracting data includes selecting the data from the list. As stated above, DaCosta discloses in FIG. 4 parsing data into lists. DaCosta further discloses that “[s]imilar to what has been described for the navigation script, the user would then make a selection from the web page or other web-accessible document by highlighting that portion 170, for example” and “[r]egarding the example presented in FIG. 4, a user has selected the starting element portion AA, part of the parent element A, and ended his selection on element portion BB of element B” and then “[t]hat selection is then mapped 180 into the HTML hierarchy that was created 160” and “[i]n other words, the user started in tag 1 of content ‘AAA’ and ended in tag 2 of content ‘BBB.’” EX.1004, 9:56-65; EX.1003, ¶73.

DaCosta further discloses selecting data where “referring now also to FIG. 9, the user then highlights (i.e., using a mouse, drags over) the desired information to be extracted 240, and the highlighted data is automatically recorded and displayed

in lower window 250” and “[t]he user may now exercise the option of editing the extraction script to fine-tune his selection.” EX.1004, 11:12-17; EX.1003, ¶74.

Also, “[r]eferring now also to FIG. 11, a user can utilize a ‘find text’ command 270 to find a particular string such as a stock symbol and extract data according thereto” and “[i]n the case of FIG. 11, the extracted data may include account type, symbol, description, quantity, high, low, change, last price and market value, for example.” EX.1004, 11:25-31; EX.1003, ¶75.

f. “... identifying source locations of the selected at least one of the data elements; ...”

DaCosta discloses this claim limitation. EX.1003, ¶76. As discussed, the extracted data is associated with a URL source location. For example, DaCosta discloses “[i]n a generic example, this could allow one to record ‘metabookmarks,’ i.e., bookmarks that record not only a destination Uniform Resource Locator (‘URL’), but also the required steps to navigate thereto, and play those steps back.” EX.1004, 5:51-55; EX.1003, ¶76.

DaCosta states that “an object of the present invention is to provide a data location and extraction tool capable of automated operation.” EX.1004, 2:2-3; EX.1003, ¶77. DaCosta further discloses that “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate

information from other applications such as Microsoft ('MS') Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶77.

g. “... selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database; ...”

DaCosta discloses this claim limitation. EX.1003, ¶78. DaCosta discloses “an architecture for a smart web server to scour data from the Internet for deposit in a data repository.” EX.1004, 5:1-3; EX.1003, ¶78. DaCosta discloses a repository, *e.g.*, a target database, in which acquired data is stored. For example, “extracted data can be stored in a conventional repository, including Microsoft Exchange 2000 or any other database that constitutes a knowledge repository for XML tagged data, so that the user can readily access the stored data needed in conjunction with other business processes.” EX.1004, 18:18-23; EX.1003, ¶78.

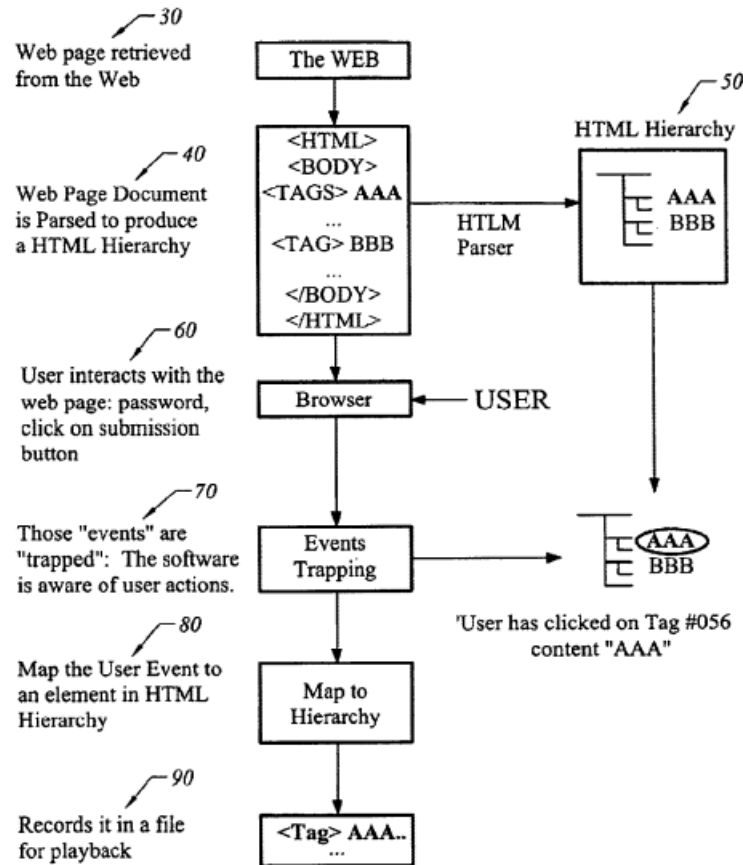
DaCosta further discloses the ability to have distributed repositories or a centralized depository. For example, “[i]n a preferred embodiment of the present invention, scripts are stored at a centralized repository that is accessible through the Internet, for example” and “[i]n this way, if there are multiple users of a script,

should that script fail it is easy to ensure each of the users has a corrected script as soon as possible.” EX.1004, 18:33-38; EX.1003, ¶79.

DaCosta discloses that “[t]he data feeds Web Folders that comprise a knowledge repository accessible by smart clients” and that the “extracted data can be stored in a conventional repository, including Microsoft Exchange 2000 or any other database that constitutes a knowledge repository.” EX.1004, 18:15-21; EX.1003, ¶80.

DaCosta discloses selecting a particular target database and location through the use of scripts. DaCosta discloses the use of navigation scripts and extraction scripts. For example, through a navigation script “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (‘MS’) Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶81.

In addition, the navigation script also controls the storage of the extracted data in a repository, *e.g.*, target database, as shown in element 90 of Figure 2, reproduced below. EX.1003, ¶82.



Recording Navigation Script - Getting to a page

FIG. 2

- h. "... saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database; ..."

DaCosta discloses this claim limitation. EX.1003, ¶83. DaCosta discloses this limitation. As discussed, DaCosta discloses the use of scripts, also referred to as Navex scripts, that include source and destination information, *e.g.*, a location file. *Id.*

For example, “[r]eferring now to FIG. 10, the user upon playing the recorded navigation and extraction scripts (collectively referred to as Navex) can automatically extract numerous records 260 that match a pattern (which may or may not have been fine-tuned) derived from the extracted text by the extraction recording module 22.” EX.1004, 11:19-24; EX.1003, ¶84.

DaCosta also discloses saving the target location in a database, where in FIG. 10 DaCosta discloses the ability to “Save your ‘script’ and run it on demand” and “View output on your Browser, Save it in Excel, Access or Word” and “Do it Every Hour – automatically.” EX.1004, FIG. 10; EX.1003, ¶85. DaCosta further discloses that “[d]ata may also be transferred to a data base for SQL processing.” EX.1004, FIG. 10; EX.1003, ¶85.

As discussed, the target database is a repository where that repository may be a conventional repository or a centralized repository. In an embodiment, DaCosta discloses storing the scripts, *e.g.*, location files, in a centralized repository, *e.g.*, target database. For example, “scripts are stored at a centralized repository that is accessible through the Internet, for example” and “[i]n this way, if there are multiple users of a script, should that script fail it is easy to ensure each of the users has a corrected script as soon as possible,” EX.1004, 18:34-38; EX.1003, ¶86.

However, the scripts may also be stored in a decentralized target database, for example, “[i]n an alternative embodiment of the present invention, extracted data can be stored in a conventional repository, including Microsoft Exchange 2000 or any other database that constitutes a knowledge repository for XML tagged data, so that the user can readily access the stored data needed in conjunction with other business processes.” EX.1004, 18:17-23; EX.1003, ¶87.

- i. *“... setting a schedule frequency for transferring the at least one of the parsed data elements from the source locations to the target locations in the location file in the target database; and ...”*

DaCosta discloses this claim limitation. EX.1003, ¶88. For example, DaCosta discloses repeating its method of extracting, processing, integrating, for example, “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (‘MS’) Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶88.

DaCosta discloses the use of parameters to control setting a schedule frequency, for example, “specifying new form parameters to replace the ones

originally recorded using the navigation recording module 12” and “[t]his enables an application to automatically repeatedly query a web site while introducing variations.” EX.1004, 7:62-65; EX.1003, ¶89. In another example, “[e]ach request by the smart server preferably ‘loops’ after a preset delay to visit the target web site and extract updated information.” EX.1004, 18:23-25; EX.1003, ¶89.

- j. *“... automatically, selectively parsing new data elements corresponding to the at least one of the parsed data elements previously selected during the initialization of the data retrieval stage, pursuant to the schedule frequency, by transferring the selectively parsed new data elements from the saved source locations in the source file to the saved target locations in the target database.”*

DaCosta discloses this claim limitation. EX.1003, ¶90. For example, DaCosta discloses the use of scripts to access new variations of data where “[r]eferring first to form variations, XML encoding of a form provides key-value pairs for form parameters” and “[t]hese can be changed by an application, either at the XML file level by simply replacing text, or at the playback level by accessing the playback module 14 and specifying new form parameters to replace the ones originally recorded using the navigation recording module 12” and “[t]his enables an application to automatically repeatedly query a web site while introducing variations.” EX.1004, 7:57-65; EX.1003, ¶90.

In addition, DaCosta discloses that “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (“MS”) Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶91.

DaCosta further discloses in claim 42 the “method of claim 40, wherein said automatically accessing comprises: periodically returning to said target web page, whereby information of interest is periodically extracted.” EX.1004, 26:20-23. And, in another example, “[e]ach request by the smart server preferably ‘loops’ after a preset delay to visit the target web site and extract updated information.” EX.1004, 18:23-25; EX.1003, ¶92.

Thus, DaCosta closes all the limitations of claim 1. EX.1003, ¶93.

2. Claim 2

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶94. DaCosta further discloses that “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process

the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (‘MS’) Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶95.

Thus, DaCosta discloses all the limitations of claim 2. EX.1003, ¶96.

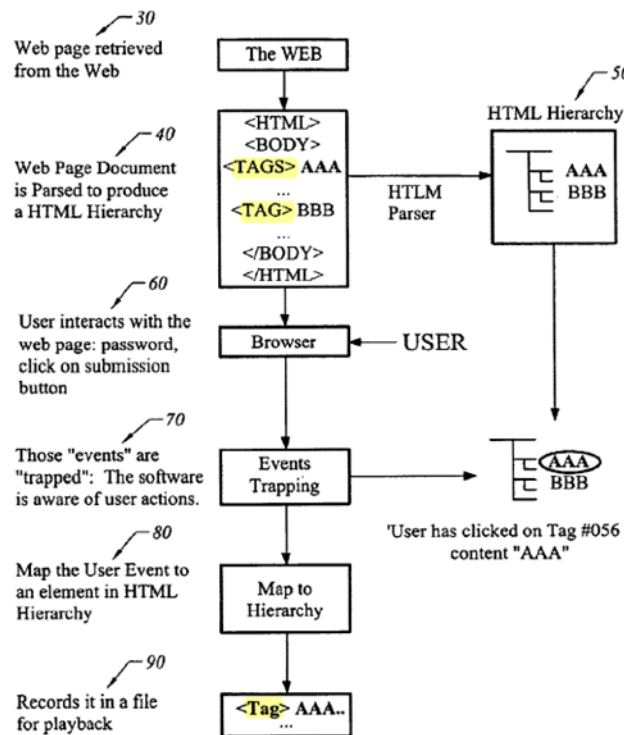
3. Claim 3

Petitioner incorporates by reference its discussion of claims 1 and 2 in Sections V.A.1 and V.A.2. herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶97. DaCosta further discloses that “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (‘MS’) Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶98.

Thus, DaCosta discloses all the limitations of claim 3. EX.1003, ¶99.

4. Claim 4

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶100. DaCosta discloses the use of tags to format data where, for example, “[t]he XML tagged data will in turn be database friendly and in a form that is easily integrated into existing business processes.” EX.1004, 2:52-54; EX.1003, ¶100. Highlighted FIG. 2 is reproduced below illustrating the use of tags.



Recording Navigation Script - Getting to a page

FIG. 2

DaCosta further discloses that “[t]he tag elements that qualify are those that meet all the criteria specified” and “[m]ore particularly, according to a preferred

embodiment of the present invention, the following selection criteria are utilized: (1) tag name-select all tags based on a list of tag names, typically one item; (2) lineage-computed by going up the DOM tree until the BODY (top of the document) is reached (the expression entered must match a portion of the computed lineage or entirely); (3),(4) text-search for text in either HTML or the Innertext; (5),(6) tag inclusions-lineage may not be specified, but instead an include/exclude list may be specified; (7),(8),(9),(10) geometry-Top, Left, Width, Height. Items 1,2,3,4,5,6 are computed by the system based on the DOM tree and Innertext and InnerHTML attributes of the element (available from the DOM tree).” EX.1004, 13:48-62; EX.1003, ¶101.

Thus, DaCosta discloses all the limitations of claim 4. EX.1003, ¶102.

5. Claim 5

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶103. DaCosta further discloses that “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (‘MS’) Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and

automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶104.

Thus, DaCosta discloses all the limitations of claim 5. EX.1003, ¶105.

6. Claim 6

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶106. The ‘667 Patent defines “server” as a “software program or a computer that responds to requests from a Web browser by returning (“serving”) Web documents.” EX.1001, 3:31-33. DaCosta discloses “an engine for accessing data on one or more web pages or other web-accessible documents primarily intended for human readability preferably using a browser, for scraping web page or other web-accessible document data...” EX.1004, 3:53-50; EX.1003, ¶106; *see also, e.g.*, EX.1004, claims 1-4.

Thus, DaCosta discloses all the limitations of claim 6. EX.1003, ¶107.

7. Claim 7

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶108.

DaCosta further discloses the use of the relational database where “[a]ccording to the present invention and referring now to the figures, wherein like reference numerals identify like elements of the various embodiments of the

invention, one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas, process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft ('MS') Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:10-20; EX.1003, ¶108.

DaCosta further discloses a user can “access the information in a database format.” EX.1004; 13:64-14:12; EX.1003, ¶108. DaCosta further discloses that “extracted data can be stored in a conventional repository, including Microsoft Exchange 2000 or any other database that constitutes a knowledge repository for XML tagged data, so that the user can readily access the stored data needed in conjunction with other business processes.” EX.1004, 18:8-25; EX.1003, ¶108.

Thus, DaCosta discloses all the limitations of claim 7. EX.1003, ¶109.

8. Claim 8

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶110.

DaCosta discloses a choice of options by the user where “[r]eferring now to FIG. 5, a user selection (for example, all HTML elements included between a start

position of the cursor and an end position, i.e., a portion highlighted) is expressed in two different spaces: first, the structural space based on the document's hierarchy; and second, the contents space based upon the document's text (individual portions of an object defined within the HTML object hierarchy).” EX.1004, 9:25-51; EX.1003, ¶110.

DaCosta discloses a variety of options to a user when initiating a search, for example, as shown below in highlighted FIG. 8. EX.1003, ¶111.

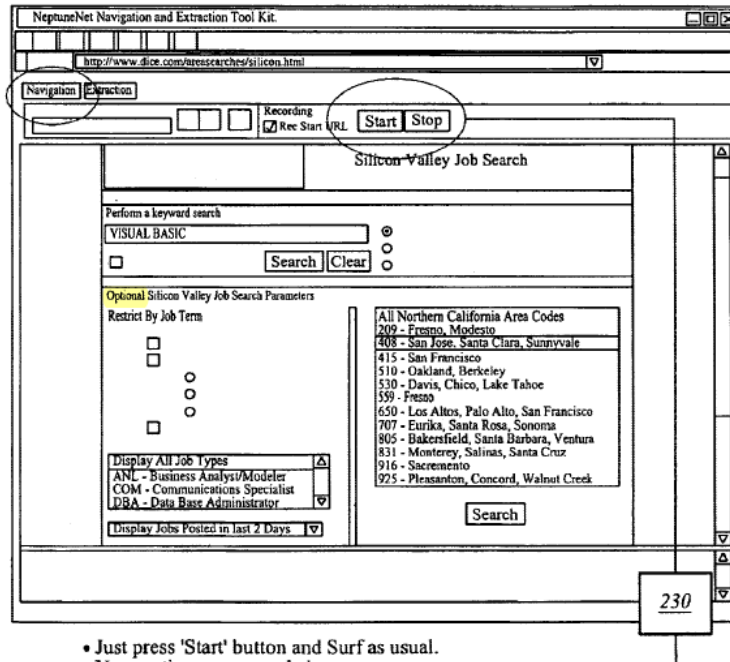
Ease of Use: Build Your Own "JobFinder"

It takes just a few steps to build any Application.

For Example: Looking For a Job? Make your own JOB FINDER!

Step 1: Navigation

- Navigation: Navigate to the URL., Record and Playback the Sequence
- Extraction: How to extract specific information
- Find Text: Use Find Text to find "interesting" information



- Just press 'Start' button and Surf as usual.
- Your actions are recorded.
- When you are done, press 'Stop' button.

FIG. 8

Thus, DaCosta discloses all the limitations of claim 8. EX.1003, ¶112.

9. Claim 9

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶113.

For example, DaCosta discloses the use of smart servers to collect data required by stock analysts such as recent news on more than 1,000 companies. EX.1004, 18:12-14; EX.1003, ¶113. DaCosta discloses “[s]mart servers provide

distributed, multiple dynamic virtual data feeds.” EX.1004, 18:14-15; EX.1003, ¶113.

Further, DaCosta discloses that “[m]oreover, the system in accordance with the present invention is preferably capable of performing some degree of learning when presented with dynamically generated web pages or other web-accessible documents (program generated pages) such as pages containing stock quotes or weather data, for example.” EX.1004, 6:21-26; EX.1003, ¶114.

Thus, DaCosta discloses all the limitations of claim 9. EX.1003, ¶115.

10. Claim 10

Petitioner incorporates by reference its discussion of claims 1 and 9 in Sections V.A.1 and V.A.9 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶116.

DaCosta discloses the use of new scripts to access new data elements, or variations of data where “[r]eferring first to form variations, XML encoding of a form provides key-value pairs for form parameters. These can be changed by an application, either at the XML file level by simply replacing text, or at the playback level by accessing the playback module 14 and specifying new form parameters to replace the ones originally recorded using the navigation recording module 12” and

“[t]his enables an application to automatically repeatedly query a web site while introducing variations.” EX.1004, 7:57-65; EX.1003, ¶117.

Thus, DaCosta discloses all the limitations of claim 10. EX.1003, ¶118.

11. Claim 11

Petitioner incorporates by reference its discussion of claims 1 and 10 in Sections V.A.1 and V.A.10 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶119.

As discussed above with respect to, for example, claim 9, DaCosta discloses “[s]mart servers provide distributed, multiple dynamic virtual data feeds.” EX.1004, 18:14-15; EX.1003, ¶120.

Further, DaCosta discloses that “[m]oreover, the system in accordance with the present invention is preferably capable of performing some degree of learning when presented with dynamically generated web pages or other web-accessible documents (program generated pages) such as pages containing stock quotes or weather data, for example.” EX.1004, 6:21-26; EX.1003, ¶121.

Thus, DaCosta discloses all the limitations of claim 11. EX.1003, ¶122.

12. Claim 12

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶123.

DaCosta discloses the use of “Web Bands [that] is a software system that is preferably installed on the microprocessor-based device of the user and exposed as standard (Vertical and Horizontal) bands housed within the browser, for example, Internet Explorer” and “[t]he Vertical Band, which is the main user interface, contains items (like buttons or hyperlinks) that allow the user to perform specific tasks using the underlying technology according to the present invention (for automated navigation, extraction, display, storage and other processing).” EX.1004, 16:14-23; EX.1003, ¶124.

Thus, DaCosta discloses all the limitations of claim 12. EX.1003, ¶125.

13. Claim 13

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶126.

For example, DaCosta discloses that “[m]oreover, the system in accordance with the present invention is preferably capable of performing some degree of learning when presented with dynamically generated web pages or other web-accessible documents (program generated pages) such as pages containing stock quotes or weather data, for example.” EX.1004, 6:21-26; EX.1003, ¶127.

DaCosta further discloses that “one can automatically navigate to a plurality of web site destinations, extract specified information based upon taught schemas,

process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft ("MS") Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:12-20; EX.1003, ¶128.

Thus, DaCosta discloses all the limitations of claim 13. EX.1003, ¶129.

14. Claim 14

Petitioner incorporates by reference its discussion of claims 1 and 13 in Sections V.A.1 and V.A.13 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶130.

For example, DaCosta discloses the use of scripts to access new variations of data where “[r]eferring first to form variations, XML encoding of a form provides key-value pairs for form parameters” and “[t]hese can be changed by an application, either at the XML file level by simply replacing text, or at the playback level by accessing the playback module 14 and specifying new form parameters to replace the ones originally recorded using the navigation recording module 12” and “[t]his enables an application to automatically repeatedly query a web site while introducing variations.” EX.1004, 7:57-65; EX.1003, ¶130.

In another example, “[e]ach request by the smart server preferably ‘loops’ after a preset delay to visit the target web site and extract updated information.” EX.1004, 18:23-25; EX.1003, ¶131.

Thus, DaCosta discloses all the limitations of claim 14. EX.1003, ¶132.

15. Claim 15

Petitioner incorporates by reference its discussion of claim 1 in Section V.A.1 herein. This additional claim limitation is disclosed in DaCosta. EX.1003, ¶133.

DaCosta discloses the repeating of steps using the script, thus reinitializing the data retrieval stage, for example, “process the extracted data according to customizable scripts, integrate information from other applications such as Microsoft (‘MS’) Word, Excel or Access, view the final output using a browser such as Microsoft Internet Explorer, for example, and automatically repeat these steps in a scheduled manner or when requested, for example.” EX.1004, 5:15-20; EX.1003, ¶133.

Thus, DaCosta discloses all the limitations of claim 15. EX.1003, ¶134.

16. Claim 16

- a. *“A computer program product having instruction codes for selectively retrieving data from a source file on a public network and for automatically collecting retrieved data in a target database, comprising: ...”*

To the extent the preamble is limiting, DaCosta discloses this claim limitation. EX.1003, ¶135. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.a herein

b. “... a first set of instruction codes for initializing a data retrieval stage by: ...”

DaCosta discloses this claim limitation. EX.1003, ¶136. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.b herein.

In addition, as discussed in relation to claim 1, DaCosta discloses a navigation script, e.g., a first set of instructions, and an extraction script, e.g., a second set of instructions. EX.1003, ¶136.

c. “... selecting the source file;...”

DaCosta discloses this claim limitation. EX.1003, ¶137. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.c herein.

d. “... parsing selected data elements in the source file into a list of data elements;...”

DaCosta discloses this claim limitation. EX.1003, ¶138. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.d herein.

e. “... selecting at least one of the data elements from the list to be stored in the target database ...”

DaCosta discloses this claim limitation. EX.1003, ¶139. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.e herein.

- f. “... identifying source locations of the selected at least one of the data elements; ...”*

DaCosta discloses this claim limitation. EX.1003, ¶140. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.f herein.

- g. “... selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database; ...”*

DaCosta discloses this claim limitation. EX.1003, ¶141. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.g herein.

- h. “... saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database; ...”*

DaCosta discloses this claim limitation. EX.1003, ¶142. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.h herein.

- i. “... setting a schedule frequency for transferring the at least one of the parsed data elements from the source locations to the target locations in the location file in the target database; and ...”*

DaCosta discloses this claim limitation. EX.1003, ¶143. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.i herein.

- j. *“... a second set of instruction codes for automatically, selectively parsing new data elements corresponding to the at least one of the parsed data elements previously selected during the initialization of the data retrieval stage, pursuant to the schedule frequency, by transferring the selectively parsed new data elements from the saved source locations in the source file to the saved target locations in the target database.”*

DaCosta discloses this claim limitation. EX.1003, ¶144. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.j herein. In addition, as discussed in relation to claim 1, DaCosta discloses a navigation script, e.g., a first set of instructions, and an extraction script, e.g., a second set of instructions.

17. Claim 17

DaCosta discloses this limitation. EX.1003, ¶145. Petitioner incorporates by reference its discussion of claim 16 in Section V.A.16 herein. Petitioner additionally incorporates by reference its discussion of claim 4 in Section V.A.4 herein.

18. Claim 18

DaCosta discloses this limitation. EX.1003, ¶146. Petitioner incorporates by reference its discussion of claim 16 in Section V.A.16 herein. Petitioner additionally incorporates by reference its discussion of claim 5 in Section V.A.5 herein.

19. Claim 19

DaCosta discloses this limitation. EX.1003, ¶147. Petitioner incorporates by reference its discussion of claim 16 in Section V.A.16 herein. Petitioner additionally incorporates by reference its discussion of claim 6 in Section V.A.6 herein.

20. Claim 20

- a. *“A computer implemented system for selectively retrieving data from a source file on a public network and for automatically collecting retrieved data in a target database, comprising “A computer implemented system for selectively retrieving data from a source file on a public network and for automatically collecting retrieved data in a target database, comprising: ...”*

To the extent the preamble is limiting, DaCosta discloses this claim limitation. EX.1003, ¶148. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.a herein.

- b. *“... means for initializing a data retrieval stage by: ...”*

DaCosta discloses this claim limitation. EX.1003, ¶149. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.b herein.

- c. *“... selecting the source file;...”*

DaCosta discloses this claim limitation. EX.1003, ¶150. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.c herein.

- d. *“... parsing selected data elements in the source file into a list of data elements;...”*

DaCosta discloses this claim limitation. EX.1003, ¶151. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.d herein.

- e. *“... selecting at least one of the data elements from the list to be stored in the target database ...”*

DaCosta discloses this claim limitation. EX.1003, ¶152. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.e herein.

- f. *“... identifying source locations of the selected at least one of the data elements; ...”*

DaCosta discloses this claim limitation. EX.1003, ¶153. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.f herein.

- g. *“... selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database; ...”*

DaCosta discloses this claim limitation. EX.1003, ¶154. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.g herein.

- h. *“... saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database; ...”*

DaCosta discloses this claim limitation. EX.1003, ¶155. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.h herein.

- i. *“... setting a schedule frequency for transferring the at least one of the parsed data elements from the source locations to the target locations in the location file in the target database; and ...”*

DaCosta discloses this claim limitation. EX.1003, ¶156. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.i herein.

- j. *“... means for automatically, selectively parsing new data elements corresponding to the at least one of the parsed data elements previously selected during the initialization of the data retrieval stage, pursuant to the schedule frequency, by transferring the selectively parsed new data elements from the saved source locations in the source file to the saved target locations in the target database.”*

DaCosta discloses this claim limitation. EX.1003, ¶157. Petitioner incorporates by reference the above discussion of claim 1 in Section V.A.1.j herein.

B. Ground 2: Anticipation of Claims 1-11 and 13-20 by Kraft

1. Claim 1

- a. *“A computer-implemented method of selectively retrieving data from a source file on a public network and automatically collecting retrieved data in a target database, comprising: ...”*

To the extent the preamble is limiting, Kraft discloses this claim limitation. EX.1003, ¶159.

Kraft presents a system and method for use with a web browser and the Internet to save the location and other context information along with the content of a web page when being stored. EX.1005, Abstract, code (57). Kraft further discloses a crawler that is run periodically to update previously stored data. EX.1005, ¶[0003]; EX.1003, ¶159. Further, Kraft discloses that the “invention relates to a system for saving the content of a target document bundled with contextual metadata, such as the location of a source document, as attributes of the target document.” EX.1005, ¶[0001]. Then, “[a]t step 340, the save command prompts the system 10 and more specifically the document storage manager 150 to create the destination document 90 by bundling the target document 60 with the context data of the source document 50.” EX.1005, ¶[0060]]; EX.1003, ¶159.

b. “... initializing a data retrieval stage by: ...”

Kraft discloses this claim limitation. EX.1003, ¶160. For example, Kraft discloses initializing the process with a user inputting search parameters, for example “[a]t step 305 of method 300, the user inputs query parameters 172 (FIG. 2) using the browser 140.” EX.1005, ¶[0057]; EX.1003, ¶160; *see also* EX.1005, FIG. 3.

Further, initializing a data retrieval stage is also shown by performing the steps outlined below. EX.1003, ¶161.

c. “... selecting the source file;...”

Kraft discloses this claim limitation. EX.1003, ¶162. For example, Kraft discloses that a user selects the source file, for example, “[w]hen a user wishes to access the source document, the system uses the saved context metadata to link the user to the source document.” EX.1005, ¶[0014] ; EX.1003, ¶162. Kraft defines the source document as “[a]n initial document or web page that points, whether directly or indirectly, to a target document and/or to a destination document.” EX.1005, ¶[0033] ; EX.1003, ¶162.

As will be further discussed below, Kraft further discloses that “at step 320 the service provider 100 returns the search results to the user's web browser 140 as the source document, and establishes a connection with the system 10” and “[t]he user reviews the search results at step 330, and, at step 335, the user navigates the Internet using the hyperlinks 70, 72, 74, 76 in the source document 50 and the intermediate document or documents 55 (FIG. 1).” EX.1005, ¶[0058]; EX.1003, ¶163. Kraft further discloses that “[t]he user continues his or her navigation until he or she detects the desired target document 60 (FIG. 1). At which point, the user identifies such target document 60, issues a save command, and enters the destination address (URL) 176 of the destination documents repository 170 where

he or she desires to store the destination document 90 (FIG. 1).” EX.1005, ¶[0059]; EX.1003, ¶163. FIG. 1 of Kraft is reproduced below.

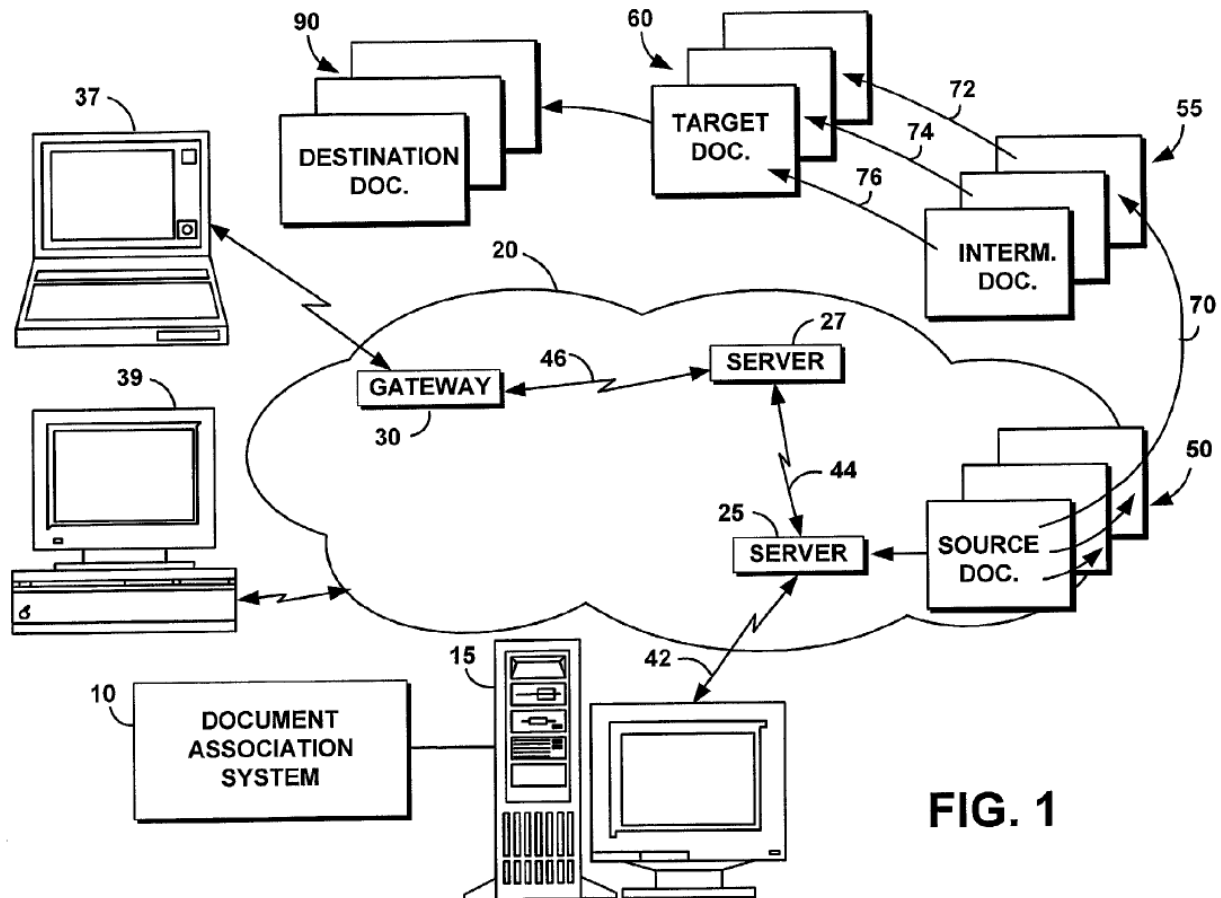


FIG. 1

Kraft further discloses that: “FIG. 4 shows an exemplary, partial screen shot of a source document such as an HTML page 400 that contains hyperlinks to various other documents in the form of underlined and highlighted text, i.e., 405 and 410. In this example, a target document titled “White Paper” is referenced by an embedded hyperlink 405 pointing to <http://time/pdfNWhitePaper.pdf>.” EX.1005, ¶¶[0061]-

[0062]; EX.1003, ¶164. Kraft further discloses that the target document is bundled with context attributes and saved as a destination document with its attributes, such that, when the URL 630 of the source document will return the user to the source document. EX.1005, ¶¶[0063]-[0064]; EX.1003, ¶164.

d. “... parsing selected data elements in the source file into a list of data elements;...”

Kraft discloses this claim limitation. EX.1003, ¶165. For example, Kraft discloses that “[d]uring a document authoring stage, HTML ‘tags’ are embedded within the informational content of the document” and “[w]hen the web document (or ‘HTML document’) is subsequently transmitted by a web server to a web browser, the tags are interpreted by the browser and used to parse and display the document” and “[i]n addition to specifying how the web browser is to display the document, HTML tags can be used to create hyperlinks to other web documents.” EX.1005, ¶[0028]; EX.1003, ¶165.

Kraft further discloses that the list of data elements may include such hyperlinks, for example, “at step 320 the service provider 100 returns the search results to the user's web browser 140 as the source document, and establishes a connection with the system 10” and “[t]he user reviews the search results at step 330, and, at step 335, the user navigates the Internet using the hyperlinks 70, 72, 74, 76

in the source document 50 and the intermediate document or documents 55 (FIG. 1).” EX.1005, ¶[0058]; EX.1003, ¶166.

e. “... selecting at least one of the data elements from the list to be stored in the target database ...”

Kraft discloses this claim limitation. EX.1003, ¶167. Kraft discloses that “[t]he user continues his or her navigation until he or she detects the desired target document 60 (FIG. 1)” and “[a]t which point, the user identifies such target document 60, issues a save command, and enters the destination address (URL) 176 of the destination documents repository 170 where he or she desires to store the destination document 90 (FIG. 1)” and “[t]he destination documents repository 170 can be located on the user's computer, on the network 20, and/or within the service provider 100.” EX.1005, ¶[0059]; EX.1003, ¶167.

f. “... identifying source locations of the selected at least one of the data elements; ...”

Kraft discloses this claim limitation. EX.1003, ¶168. As discussed, Kraft discloses “at step 320 the service provider 100 returns the search results to the user's web browser 140 as the source document, and establishes a connection with the system 10” where “[t]he user reviews the search results at step 330, and, at step 335, the user navigates the Internet using the hyperlinks 70, 72, 74, 76 in the source

document 50 and the intermediate document or documents 55 (FIG. 1).” EX.1005, ¶[0058]; EX.1003, ¶168.

Kraft further discloses that “[a]t step 340, the save command prompts the system 10 and more specifically the document storage manager 150 to create the destination document 90 by bundling the target document 60 with the context data of the source document 50, as explained earlier.” EX.1005, ¶[0060]; EX.1003, ¶169.

Kraft discloses that “document association system and method of the present invention satisfy this need by bundling or associating a target document (i.e., a web page) and the context of a source document as metadata to the target document during a save process.” EX.1005, ¶[0009]; EX.1003, ¶170. Kraft further describes the context data where “[t]he context of the source document may include, for example, one or more of the following parameters: [t]he location or address, such as the URL, of the source document; the path, such as pages examined to navigate from the source document to the target document; and the input parameters required to generate the target document, such as the search query inputted by the user.” EX.1005, ¶¶[0010]-[0013]; EX.1003, ¶170.

Kraft further defines context data, *e.g.*, the location file, where “[s]pecific exemplary attributes (or context data) added by the system 10 are shown under the Origin folder 620 as Source, Author, Revision Number, and Target, where the

Source refers to the URL 630 of the source document, and Target refers to the URL 640 of the target document.” EX.1005, ¶[0064]; EX.1003, ¶171.

g. “... selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database; ...”

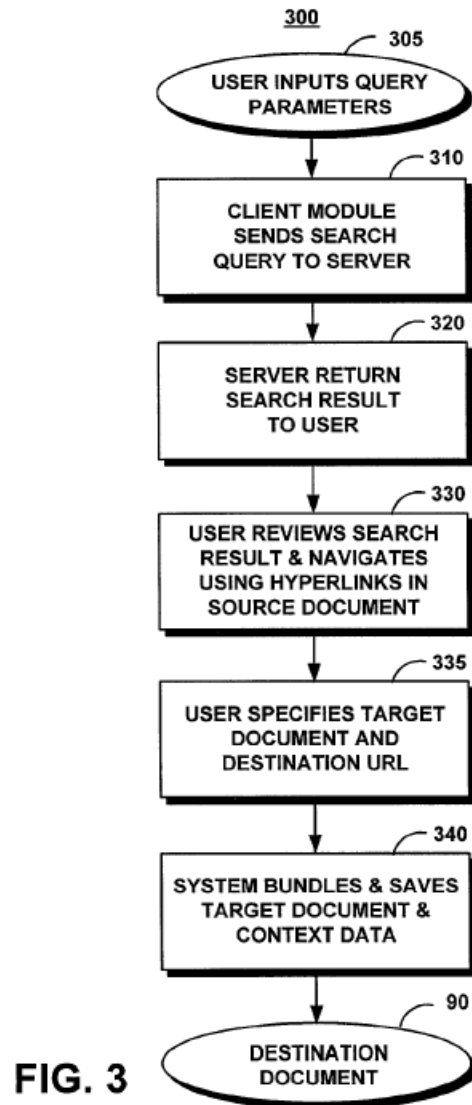
Kraft discloses this claim limitation. EX.1003, ¶172. Kraft discloses the target database as the “destination documents repository 170 where destination documents 90 (FIG. 1) are stored.” EX.1005, ¶[0045]; EX.1003, ¶172.

Kraft discloses selecting the target data and target locations where “[t]he user continues his or her navigation until he or she detects the desired target document 60 (FIG. 1)” and “[a]t which point, the user identifies such target document 60, issues a save command, and enters the destination address (URL) 176 of the destination documents repository 170 where he or she desires to store the destination document 90 (FIG. 1).” EX.1005, ¶[0059]; EX.1003, ¶173.

Kraft further discloses that the “query and query results can be stored, for example in the destination documents repository 170 or in any other data storage system, whether on the user's side, the service provider's 100, or an independent network storage repository for later use by the document storage manager 150.” EX.1005, ¶[0051]; EX.1003, ¶174.

- h. “... saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database; ...”*

Kraft discloses this claim limitation. EX.1003, ¶175. As discussed, Kraft discloses context data, *e.g.*, location file, that is stored with the destination document in the target database, for example, “document association system and method of the present invention satisfy this need by bundling or associating a target document (i.e., a web page) and the context of a source document as metadata to the target document during a save process.” EX.1005, ¶[0009]; EX.1003, ¶175. The bundled context data and destination document are saved at the destination URL, *e.g.*, target database, as shown in FIG. 3, reproduced below. EX.1003, ¶175.



Kraft describes the context data where “[t]he context of the source document may include, for example, one or more of the following parameters: [t]he location or address, such as the URL, of the source document; the path, such as pages examined to navigate from the source document to the target document; and the input parameters required to generate the target document, such as the search query inputted by the user.” EX.1005, ¶¶[0010]-[0013]; EX.1003, ¶176.

Kraft further describes, with respect to FIGS. 6 and 7 that the document properties 600 for the destination document are saved, along with additional attributes (context data). EX.1005, ¶¶[0062]-[0063]; EX.1003, ¶177.

- i. *“... setting a schedule frequency for transferring the at least one of the parsed data elements from the source locations to the target locations in the location file in the target database; and ...”*

Kraft discloses this claim limitation. EX.1003, ¶178. For example, Kraft discloses setting a schedule frequency where “users will be able to return to the source document, and optionally to use applications for automatically synchronizing a destination document to the target document.” EX.1005, ¶[0009]; EX.1003, ¶178.

Kraft further discloses that “[w]hen coupled with a synchronization application, the system of the invention allows the user to update the destination document to reflect changes in the target document, allowing a convenient mechanism for updating saved documents” where “[t]he synchronization application performs a comparison of the destination document with the target document to detect changes and to automatically update the destination document.” EX.1005, ¶[0016]; EX.1003, ¶179.

Kraft also discloses that the “system 10, transparently to the user, continuously or periodically operates in the background.” EX.1005, ¶[0044]; EX.1003, ¶180. Kraft, in the background section also notes that data can be updated

periodically where “[t]he crawler is run periodically to update previously stored data and to append information about newly created web pages” and that “[t]he information compiled by the crawler is stored in a Metadata repository or database.” EX.1005, ¶[0003]; EX.1003, ¶180.

- j. *“... automatically, selectively parsing new data elements corresponding to the at least one of the parsed data elements previously selected during the initialization of the data retrieval stage, pursuant to the schedule frequency, by transferring the selectively parsed new data elements from the saved source locations in the source file to the saved target locations in the target database.”*

Kraft discloses this claim limitation. EX.1003, ¶181. For example, Kraft further discloses that “[w]hen coupled with a synchronization application, the system of the invention allows the user to update the destination document to reflect changes in the target document, allowing a convenient mechanism for updating saved documents” where “[t]he synchronization application performs a comparison of the destination document with the target document to detect changes and to automatically update the destination document.” EX.1005, ¶[0016]; EX.1003, ¶181.

Kraft further discloses the ability to identify input parameters to identify new dynamic content, for example, “[i]nput parameters 172, such as query parameters for a dynamic document, from a query transformer 230” and “URL 174 of a source

document (i.e., 50 in FIG. 1) from the dynamic query matcher 160 (or from the service provider 100).” EX.1005, ¶¶[0046]-[0047]; EX.1003, ¶182.

Thus, Kraft closes all the limitations of claim 1. EX.1003, ¶183.

2. Claim 2

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶184. Kraft further discloses that “[t]he documents storage manager 150 is responsible for bundling the content 178 of the target document 60 and the contextual data related to the source document 50, and to save the newly bundled document as destination document 90 in the destination documents repository 170 ... [and] [i]n use, the client session query, including the input parameters 172, is forwarded to the service provider 100 for normal query processing, whereupon the service provider 100 forwards the search results to the system 10 for further processing.” EX.1005, ¶¶[0050]-[0051]; EX.1003, ¶185.

Kraft further discloses automatically updating the destination document, for example, if the target document was deleted from its original location or relocated. EX.1005, ¶[0016]; EX.1003, ¶186.

Thus, Kraft discloses all the limitations of claim 2. EX.1003, ¶187.

3. Claim 3

Petitioner incorporates by reference its discussion of claims 1 and 2 in Sections V.B.1 and V.B.2 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶188. Kraft further discloses that “the service provider 100 is generally comprised of a web crawler 200.” EX.1005, ¶[0052]; EX.1003, ¶189. And, that the definition of a crawler is “[a] program that automatically explores the World Wide Web by retrieving a document and recursively retrieving some or all the documents that are linked to it.” EX.1005, ¶[0026]; EX.1003, ¶189.

Kraft further discloses automatically updating the destination document, for example, if the target document was deleted from its original location or relocated. EX.1005, ¶[0016]; EX.1003, ¶190.

Thus, Kraft discloses all the limitations of claim 3. EX.1003, ¶191.

4. Claim 4

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶192.

Kraft discloses “[a] standard language for attaching presentation and linking attributes to informational content within documents” and “[d]uring a document authoring stage, HTML ‘tags’ are embedded within the informational content of the document” such that “[w]hen the web document (or ‘HTML document’) is

subsequently transmitted by a web server to a web browser, the tags are interpreted by the browser and used to parse and display the document.” EX.1005, ¶[0028]; EX.1003, ¶193.

Thus, Kraft discloses all the limitations of claim 4. EX.1003, ¶194.

5. Claim 5

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶195.

Kraft further discloses that a document may be generated by a document application, *e.g.*, where “FIG. 6 illustrates the document properties 600 for the destination document” and that “[t]he document Description folder 615 remains the same as provided by the operating system and the document application.” EX.1005, ¶[0063]; EX.1003, ¶196.

Thus, Kraft discloses all the limitations of claim 5. EX.1003, ¶197.

6. Claim 6

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶198.

Kraft discloses that documents may be derived from a web server, for example, FIG.4 shows an exemplary, partial screen shot of a source document such as an HTML page 400. EX.1005, ¶[0061]; EX.1003, ¶198. Kraft further discloses

“[w]hen the web document (or “HTML document”) is subsequently transmitted by a web server to a web browser, the tags are interpreted by the browser and used to parse and display the document.” EX.1005, ¶[0028]; EX.1003, ¶198.

Thus, Kraft discloses all the limitations of claim 6. EX.1003, ¶199.

7. Claim 7

Petitioner incorporates by reference of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶200.

Kraft discloses the use of databases, for example, “[t]he information compiled by the crawler is stored in a Metadata repository or database.” EX.1005, ¶[0003]; EX.1003, ¶201. Kraft also discloses that “the system 10 can be used with a stand alone database of documents that may have been derived from the WWW and/or other sources.” EX.1005, ¶[0040]; EX.1003, ¶201. Kraft also discloses the use of documents repository 170, or any other data storage system, for later use by the document storage manager 150. EX.1005, ¶[0050]; EX.1003, ¶201.

Kraft further discloses that the disclosed invention is applicable to database “and other tables with indexed entries.” EX.1005, ¶[0056]; EX.1003, ¶202. Kraft further discloses, for example, that “[t]he search engine repository 210 maintains information or metadata from previously encountered web pages, which metadata is

used by the abstract/indexing engine 220 to prepare the abstracts.” EX.1005, ¶[0054]; EX.1003, ¶202.

Thus, Kraft discloses all the limitations of claim 7. EX.1003, ¶203.

8. Claim 8

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶204.

For example, Kraft discloses initializing the process with a user inputting search parameters, for example “[a]t step 305 of method 300, the user inputs query parameters 172 (FIG. 2) using the browser 140.” EX.1005, ¶[0057]; EX.1003, ¶204.

Thus, Kraft discloses all the limitations of claim 8. EX.1003, ¶205.

9. Claim 9

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶206.

For example, Kraft discloses that a web page can be dynamic or static. EX.1005, ¶[0037]; EX.1003, ¶206. Kraft further discloses that “[t]he system 10 includes the following components: a user module also referred to herein as a document storage manager 150, a server module also referred to herein as dynamic query matcher 160, and a destination documents repository 170 where destination documents 90 (FIG. 1) are stored.” EX.1005, ¶[0045]; EX.1003, ¶206.

Kraft further discloses where “[i]nput parameters 172, such as query parameters for a dynamic document, from a query transformer 230; URL 174 of a source document (i.e., 50 in FIG. 1) from the dynamic query matcher 160 (or from the service provider 100).” EX.1005, ¶¶[0046]-[0047]; EX.1003, ¶207.

Thus, Kraft discloses all the limitations of claim 9. EX.1003, ¶208.

10. Claim 10

Petitioner incorporates by reference its discussion of claims 1 and 9 in Sections V.B.1 and V.B.9 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶209.

Kraft discloses where “[t]he crawler is run periodically to update previously stored data and to append information about newly created web pages.” EX.1005, ¶[0003]; EX.1003, ¶210.

Thus, Kraft discloses all the limitations of claim 10. EX.1003, ¶211.

11. Claim 11

Petitioner incorporates by reference its discussion of claims 1 and 10 in Sections V.B.1 and V.B.10 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶212.

As discussed above with respect to, for example, claim 10, Kraft discloses where “[t]he crawler is run periodically to update previously stored data and to

append information about newly created web pages.” EX.1005, ¶[0003]; EX.1003, ¶213. In addition, Kraft discloses the ability to identify input parameters to identify new dynamic content, for example, “[i]nput parameters 172, such as query parameters for a dynamic document, from a query transformer 230” and “URL 174 of a source document (i.e., 50 in FIG. 1) from the dynamic query matcher 160 (or from the service provider 100).” EX.1005, ¶¶[0046]-[0047]; EX.1003, ¶213.

Thus, Kraft discloses all the limitations of claim 11. EX.1003, ¶214.

12. Claim 13

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶215.

Kraft discloses that “users will be able to return to the source document, and optionally to use applications for automatically synchronizing a destination document to the target document.” EX.1005, ¶[0009]; EX.1003, ¶216. Kraft further discloses that “[t]he synchronization application performs a comparison of the destination document with the target document to detect changes and to automatically update the destination document” and “[i]f the target document were deleted from its original location or relocated, the destination document is marked as orphaned.” EX.1005, ¶[0016]; EX.1003, ¶216.

Thus, Kraft discloses all the limitations of claim 13. EX.1003, ¶217.

13. Claim 14

Petitioner incorporates by reference its discussion of claims 1 and 13 in Sections V.B.1 and V.B.12 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶218.

For example, Kraft discloses in its background section that “[t]he crawler is run periodically to update previously stored data and to append information about newly created web pages” and that “[t]he information compiled by the crawler is stored in a Metadata repository or database.” EX.1005, ¶[0003]; EX.1003, ¶218.

Thus, Kraft discloses all the limitations of claim 14. EX.1003, ¶219.

14. Claim 15

Petitioner incorporates by reference its discussion of claim 1 in Section V.B.1 herein. This additional claim limitation is disclosed in Kraft. EX.1003, ¶220.

Kraft discloses the use of a crawler to retrieve documents where a crawler is defined as “[a] program that automatically explores the World Wide Web by retrieving a document and recursively retrieving some or all the documents that are linked to it.” EX.1005, ¶[0026]; EX.1003, ¶220.

Thus, Kraft discloses all the limitations of claim 15. EX.1003, ¶221.

15. Claim 16

- a. *“A computer program product having instruction codes for selectively retrieving data from a source file on a public network and for automatically collecting retrieved data in a target database, comprising: ...”*

To the extent the preamble is limiting, Kraft discloses this claim limitation. EX.1003, ¶222. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.a herein

- b. *“... a first set of instruction codes for initializing a data retrieval stage by: ...”*

Kraft discloses this claim limitation. EX.1003, ¶223. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.b herein.

- c. *“... selecting the source file;...”*

Kraft discloses this claim limitation. EX.1003, ¶224. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.c herein.

- d. *“... parsing selected data elements in the source file into a list of data elements;...”*

Kraft discloses this claim limitation. EX.1003, ¶225. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.d herein.

- e. *“... selecting at least one of the data elements from the list to be stored in the target database ...”*

Kraft discloses this claim limitation. EX.1003, ¶226. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.e herein.

- f. *“... identifying source locations of the selected at least one of the data elements; ...”*

Kraft discloses this claim limitation. EX.1003, ¶227. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.f herein.

- g. *“... selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database; ...”*

Kraft discloses this claim limitation. EX.1003, ¶228. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.g herein.

- h. *“... saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database; ...”*

Kraft discloses this claim limitation. EX.1003, ¶229. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.h herein.

- i. *“... setting a schedule frequency for transferring the at least one of the parsed data elements from the source locations to the target locations in the location file in the target database; and ...”*

Kraft discloses this claim limitation. EX.1003, ¶230. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.i herein.

- j. *“... a second set of instruction codes for automatically, selectively parsing new data elements corresponding to the at least one of the parsed data elements previously selected during the initialization of the data retrieval stage, pursuant to the schedule frequency, by transferring the selectively parsed new data elements from the saved source locations in the source file to the saved target locations in the target database.”*

Kraft discloses this claim limitation. EX.1003, ¶231. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.j herein.

16. Claim 17

Kraft discloses this limitation. EX.1003, ¶232. Petitioner incorporates by reference its discussion of claim 16 in Sections V.B.15 herein. Petitioner additionally incorporates by reference its discussion of claim 4 in Section V.B.4 herein.

17. Claim 18

Kraft discloses this limitation. EX.1003, ¶233. Petitioner incorporates by reference its discussion of claim 16 in Section V.B.15 herein. Petitioner additionally incorporates by reference its discussion of claim 5 in Section V.B.5 herein.

18. Claim 19

Kraft discloses this limitation. EX.1003, ¶234. Petitioner incorporates by reference its discussion of claim 16 in Section V.B.15 herein. Petitioner additionally incorporates by reference its discussion of claim 6 in Section V.B.6 herein.

19. Claim 20

- a. *“A computer implemented system for selectively retrieving data from a source file on a public network and for automatically collecting retrieved data in a target database, comprising “A computer implemented system for selectively retrieving data from a source file on a public network and for automatically collecting retrieved data in a target database, comprising: ...”*

Kraft discloses this claim limitation. EX.1003, ¶235. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.a herein.

- b. *“... means for initializing a data retrieval stage by: ...”*

Kraft discloses this claim limitation. EX.1003, ¶236. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.b herein.

- c. *“... selecting the source file;...”*

Kraft discloses this claim limitation. EX.1003, ¶237. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.c herein.

- d. *“... parsing selected data elements in the source file into a list of data elements;...”*

Kraft discloses this claim limitation. EX.1003, ¶238. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.d herein.

- e. *“... selecting at least one of the data elements from the list to be stored in the target database ...”*

Kraft discloses this claim limitation. EX.1003, ¶239. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.e herein.

- f. *“... identifying source locations of the selected at least one of the data elements; ...”*

Kraft discloses this claim limitation. EX.1003, ¶240. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.f herein.

- g. *“... selecting the target database and target locations in the target database where the at least one of the data elements is to be stored in the target database; ...”*

Kraft discloses this claim limitation. EX.1003, ¶241. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.g herein.

- h. *“... saving the source locations and target locations of the at least one of the parsed data elements in a location file in the target database; ...”*

Kraft discloses this claim limitation. EX.1003, ¶242. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.h herein.

- i. *“... setting a schedule frequency for transferring the at least one of the parsed data elements from the source locations to the target locations in the location file in the target database; and ...”*

Kraft discloses this claim limitation. EX.1003, ¶243. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.i herein.

- j. “... means for automatically, selectively parsing new data elements corresponding to the at least one of the parsed data elements previously selected during the initialization of the data retrieval stage, pursuant to the schedule frequency, by transferring the selectively parsed new data elements from the saved source locations in the source file to the saved target locations in the target database.”

Kraft discloses this claim limitation. EX.1003, ¶244. Petitioner incorporates by reference the above discussion of claim 1 in Section V.B.1.j herein.

VI. DISCRETIONARY FACTORS

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

First, the prior art presented in this petition for the independent claims is not the same or substantially the same as the art previously presented during prosecution of the '667 Patent. As such, none of the factors set forth in *Advanced Bionics* apply to this petition. See *Advanced Bionics, LLC v. MED-EL Elektromedizinische Geräte GmbH*, IPR2019-01469, Paper 6 at 7-9 (PTAB Feb. 13, 2020).

Second, the '667 Patent has not been challenged in any prior administrative proceeding. Thus, none of the factors set forth in *General Plastic* apply to this petition. See *General Plastic Co., Ltd. v. Canon Kabushiki Kaisha*, IPR2016-01357, Paper 19 at 15-16 (PTAB Sept. 6, 2017).

Third, the '667 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”), but as explained below, 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.1007.

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.1008 at 35. The case was opened on April 14, 2023.² Thus, the trial would statistically occur around September 2024.³

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.

² 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.1009. Oxylabs’ counterclaims for infringement of U.S. Patent No. 7,010,526 and the ’667 Patent were filed on 4/28/2023. See EX.1010.

³ 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.1007.

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 been issued by the court. Accordingly, this factor weighs strongly against discretionary denial.

Fintiv Factor 4:

There are 20 claims in the '667 Patent.

Claims 1, 4, 7, 8, 9, 13, and 14 of the '667 Patent are at issue in the Bright Data Litigation. That totals 7 asserted claims, with only claim 1 being independent.

Claims 1-20 are challenged in this petition. That totals 20 challenged claims, with claims 1, 16, and 20 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 '667 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.

VII. CONCLUSION

Petitioner respectfully requests cancellation of the Challenged Claims of the '667 Patent.

Respectfully submitted,

Date: October 4, 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 13,453 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: October 4, 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:

Intellectual Property Law Department
T.J. Watson Research Center
1101 Kitchawan Road, Route 134, PO Box 218
Yorktown, NY 10598

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: October 4, 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.